Citation: Neben, Abraham R. et al. "Foreground and Sensitivity Analysis for Broadband (2D) 21 cm–Lyα and 21 cm–Hα Correlation Experiments Probing the Epoch of Reionization." The Astrophysical Journal 849, 1 (October 2017): 50 © 2017 The American Astronomical Society

As Published: http://dx.doi.org/10.3847/1538-4357/AA8F9C

Publisher: IOP Publishing

Persistent URL: http://hdl.handle.net/1721.1/114972

Version: Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

Massachusetts Institute of Technology

Foreground and Sensitivity Analysis for Broadband (2D) 21 cm–Lyα and 21 cm–Hα Correlation Experiments Probing the Epoch of Reionization

Abraham R. Neben[1,2] ⓘ, Brian Stalder[3] ⓘ, Jacqueline N. Hewitt[1,2], and John L. Tonry[3] ⓘ

[1] MIT Kavli Institute, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[2] Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[3] Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, HI 96822, USA

Abstract

A detection of the predicted anticorrelation between 21 cm and either Lyα or Hα from the epoch of reionization (EOR) would be a powerful probe of the first galaxies. While 3D intensity maps isolate foregrounds in low-$k_{\parallel}$ modes, infrared surveys cannot yet match the field of view and redshift resolution of radio intensity mapping experiments. In contrast, 2D (i.e., broadband) infrared intensity maps can be measured with current experiments and are limited by foregrounds instead of photon or thermal noise. We show that 2D experiments can measure most of the 3D fluctuation power at $k < 0.2$ Mpc^{-1} while preserving its correlation properties. However, we show that foregrounds pose two challenges: (1) simple geometric effects produce percent-level correlations between radio and infrared fluxes, even if their luminosities are uncorrelated; and (2) radio and infrared foreground residuals contribute sample variance noise to the cross spectrum. The first challenge demands better foreground masking and subtraction, while the second demands large fields of view to average away uncorrelated radio and infrared power. Using radio observations from the Murchison Widefield Array and near-infrared observations from the Asteroid Terrestrial-impact Last Alert System, we set an upper limit on residual foregrounds of the 21 cm–Lyα cross-power spectrum at $z \sim 7$ of $\Delta^2 < 181$ (kJy sr^{-1} mK) (95%) at $\ell \sim 800$. We predict levels of foreground correlation and sample variance noise in future experiments, showing that higher-resolution surveys such as LOFAR, SKA-LOW, and the Dark Energy Survey can start to probe models of the 21 cm–Lyα EOR cross spectrum.

Key words: cosmology: observations – dark ages, reionization, first stars – infrared: diffuse background

1. Introduction

Deep radio and infrared observations are nearing detection of the first stars and galaxies from the cosmic dawn. As such sources form, they are thought to blow out ionized bubbles, eventually merging and reionizing the universe. See Furlanetto et al. (2006), Morales & Wyithe (2010), Pritchard & Loeb (2012), and Mesinger (2016) for reviews. First-generation 21 cm observatories such the Murchison Widefield Array (MWA; Bowman et al. 2013; Tingay et al. 2013) and the Precision Array for Probing the Epoch of Reionization (PAPER; Parsons et al. 2014; Ali et al. 2015; Jacobs et al. 2015; Pober et al. 2015) are setting ever tighter limits on redshifted neutral hydrogen emission from the neutral regions between these bubbles, and the now-underway Hydrogen Epoch of Reionization Array (HERA; Ewall-Wice et al. 2016; Neben et al. 2016; DeBoer et al. 2017; Patra et al. 2017) is expected to detect and characterize the EOR power spectrum in the coming years. Similar efforts are under way by the Low Frequency Array (van Haarlem et al. 2013; Yatawatta et al. 2013). Ultimately, the Square Kilometer Array (SKA; Hall 2005; Dewdney & Lazio 2008; Dewdney et al. 2009; bij de Vaate et al. 2011) will image the EOR over redshift, revealing the detailed hydrogen reionization history of the universe.

At the same time, new galaxy surveys are beginning to constrain the reionizing sources themselves. Deep galaxy surveys (Bouwens et al. 2011, 2014, 2015; Grogin et al. 2011; Dunlop et al. 2013; Ellis et al. 2013; Illingworth et al. 2013; Oesch et al. 2013; Robertson et al. 2013; Bowler et al. 2015, 2017; Roberts-Borsani et al. 2016; Wilkins et al. 2016) and cluster lensing surveys (Atek et al. 2015; Coe et al.

2015; McLeod et al. 2015; Bouwens et al. 2016a, 2016b; Huang et al. 2016; McLeod et al. 2016; Repp et al. 2016; Livermore et al. 2017) are finding hundreds of galaxy candidates at $6 < z < 10$ down to UV magnitudes of $M_{AB} \sim -17$ (Finkelstein 2016), and extremely wide surveys are searching for the rare bright galaxies (Trenti et al. 2011; Bradley et al. 2012; Schmidt et al. 2014; Bernard et al. 2016; Calvi et al. 2016) from the reionization epoch. However, current models require the ionizing contribution of far fainter galaxies down to $M_{AB} \sim -13$ (Alvarez et al. 2012; Bouwens 2016) in order for reionization to be complete by the time that cosmic microwave background (CMB) optical depth measurements (Planck Collaboration et al. 2016) say it must be. Deeper observations with the *James Webb Space Telescope* (*JWST*; Gardner et al. 2006), expected to probe down to $M_{AB} \sim -15.5$ in order 100 hr integrations (Finkelstein 2016), will be needed to begin to probe this crucial faint population directly.

Infrared intensity mapping offers several advantages compared to galaxy surveys. Power spectrum analyses can be sensitive to an EOR component even if the signal-to-noise ratio (S/N) in individual pixels is small, and instead of being limited to the brightest galaxies, intensity mapping is sensitive to the cumulative light from *all* sources. The expected bright Lyα (e.g., Amorín et al. 2017) and Hα (e.g., Smit et al. 2014) radiation from EOR galaxies at $z \sim 6 - 8$ is motivating intensity mapping at micron-scale wavelengths. Working around foregrounds is challenging, though. While early studies suggested that angular fluctuations in infrared intensity maps traced EOR galaxies (e.g., Kashlinsky et al. 2005, 2007, 2012), Helgason et al. (2016) find that, given current constraints on the EOR, this is unlikely. Intrahalo light (Cooray et al. 2012; Zemcov et al. 2014) and Galactic dust (Yue et al. 2016) have

been proposed as more likely explanations for the larger-than-expected fluctuations, though Mitchell-Wynne et al. (2015) show that much of this excess can be removed with higher-resolution measurements using the *Hubble Space Telescope* (*HST*). In any case, all these components are likely present at some level, and degeneracies coupled with imperfect foreground knowledge make isolating the EOR contribution difficult.

For these reasons, cross-correlation with 21 cm maps may be the cleanest way to extract the EOR component of the near-infrared background. The synergy is clear: the galaxies sourcing reionization generate strong Lyα emission, while the neutral regions between them glow at rest frame 21 cm. On typical ionized bubble scales, bright spots in IR maps likely correspond to ionized regions, and thus 21 cm dark spots, and vice versa. This effect is expected to source an anticorrelation, seen in simulations by Silva et al. (2013) and Heneka et al. (2016) and modeled analytically by Feng et al. (2017) and Mao (2014).

A similar large-scale anticorrelation is found by Lidz et al. (2009) and Park et al. (2014) in simulations of 21 cm cross-correlation with galaxy redshift surveys. However, conducting redshift surveys both wide and deep enough to cross-correlate with 21 cm maps is challenging owing to the hugely different spatial scales probed by 21 cm experiments and spectroscopic galaxy surveys. For instance, the $\sim 3'$ angular resolution of the MWA is nearly equal to the field of view of the Hubble Deep Field and the *JWST*. In a source-by-source manner Beardsley et al. (2015) show that this anticorrelation may be studied by inspecting the 21 cm brightness temperatures at the locations of *JWST* sources, but as discussed above, source detections will be limited to the brightest and rarest objects.

Intensity mapping thus holds great promise to extend EOR science, and 3D intensity mapping (i.e., with a redshift dimension) has the advantage of avoiding the majority of continuum emission from intermediate-redshift galaxies. This foreground emission is expected to contaminate only the lowest few line-of-sight Fourier modes (Gong et al. 2017), while line interlopers at intermediate redshifts are easily masked (Gong et al. 2014, 2017; Pullen et al. 2014; Comaschi et al. 2016). 3D power spectra are also easier to understand theoretically, as they quantify emission from a fundamentally 3D volume, and early demonstrations of this type of analysis are given by Chang et al. (2010) and Masui et al. (2013), who detected the cross-correlation between 21 cm emission and a galaxy redshift survey at $z \sim 1$. However, near-infrared intensity mapping in 3D likely requires space-based observations to avoid atmospheric OH lines (e.g., Sullivan & Simcoe 2012), as well as fine spectral resolution to match the redshift resolution of typical 21 cm experiments.

For instance, Pober et al. (2014) show that with moderate foreground assumptions, HERA should achieve $>5\sigma$ detections of the 21 cm power spectrum over $0.2 < k_{\parallel} < 0.5$ at $z \sim 8$, corresponding to redshift scales of $0.03 < \Delta z < 0.1$. Resolving these same line-of-sight modes of the Lyα field at the same redshift requires a spectral resolution[4] of $80 < R < 250$. Achieving this high spectral resolution over a large enough field of view to match a 21 cm survey is challenging. The proposed *SPHEREx* mission (Doré

et al. 2014, 2016) would image the entire sky in the near-infrared with $R = 40$ spectroscopy for a cost of roughly \$100 million, and the concept *Cosmic Dawn Intensity Mapper* (Cooray et al. 2016) mission would achieve $R = 200$–300 over 10 square degrees for nearly 10 times the cost.

In contrast, 2D (i.e., broadband) intensity mapping enables similar science with far shallower and cheaper observations (Fernandez et al. 2014; Mao 2014), though here the main challenge is imperfect foreground removal. Even if radio and infrared foreground residuals are uncorrelated, they leak power into the cross-correlation analysis, which averages down only over sufficiently large fields of view. As OH emission from the atmosphere is relatively smooth over few-degree scales (High et al. 2010), there is hope that these observations can be conducted from the ground for a further reduction in cost. A number of new ground-based wide-field surveys are coming online such as the Dark Energy Survey (Dark Energy Survey Collaboration et al. 2016), Pan-STARRS (Tonry et al. 2012), and the Asteroid Terrestrial-impact Last Alert System (ATLAS; Tonry 2011). Further, the *Transiting Exoplanet Survey Satellite* (Ricker et al. 2014) will survey the entire sky over the 600–1000 nm band, and the *Wide-field Infrared Survey Telescope* (*WFIRST*; Spergel et al. 2013) will observe instantaneous deep fields $100\times$ larger than those of *HST* and *JWST*. It is important to note that a large, uniform focal plane greatly facilitates intensity mapping, lest structures on relevant angular scales be lost in the calibration of many independent regions of a segmented focal plane, such as that of Pan-STARRS.

In this paper we study the foregrounds in broadband 21 cm–Lyα and 21 cm–Hα intensity mapping correlation experiments targeting the EOR. We begin in Section 2 with a review of our Fourier transform and power spectrum conventions. In Section 1 we present the MWA and ATLAS observations we use and discuss processing these data into images. In Section 2 we characterize the bright radio and infrared point-source foregrounds and demonstrate that geometric effects introduce slight positive correlations that will overpower the cosmological signal unless significant masking and subtraction are conducted. In Section 3, we study how best to mask and subtract radio and infrared foregrounds in real-world images and quantify the foreground residuals. We set the first limits on residual foregrounds of the broadband 21 cm–Lyα cross spectrum at $z \sim 7$ using data from the MWA and ATLAS and predict the sensitivities of future experiments and compare them with the expected levels of geometric foreground correlation, illustrating what it will take to realize this measurement.

2. Power Spectrum and Correlation Conventions

2.1. Power Spectrum Definitions

We define the 3D power spectrum $P(\boldsymbol{k})$ of the image cube $I(\boldsymbol{x})$ as

$$P(\boldsymbol{k}) = \frac{\langle |\tilde{I}(\boldsymbol{k})|^2 \rangle}{V}, \qquad (1)$$

where $\tilde{I}(\boldsymbol{k})$ is given by

$$\tilde{I}(\boldsymbol{k}) = dV \sum_{\boldsymbol{x}} I(\boldsymbol{x}) e^{i\boldsymbol{k}\cdot\boldsymbol{x}}, \qquad (2)$$

---

[4] The redshift resolution of a spectral line intensity mapping experiment observing emission at redshift z is given by $\Delta z = (1 + z)/R$, where R is the spectral resolving power.

Here V is the survey volume, and dV is the voxel size. Note that $P(\boldsymbol{k})$ has units of $[I]^2 \cdot \mathrm{Mpc}^3$ and that we often plot instead the more intuitive quantity $\Delta(k) = [k^3 P(k)/2\pi^2]^{1/2}$, where $P(k)$ is the average of $P(\boldsymbol{k})$ within the 1D power spectrum bin k.

Similarly, over narrow fields of view, the angular power spectrum $C(\ell)$ of a 2D (e.g., broadband) image $I(\boldsymbol{\theta})$ can be shown to be approximately

$$C(\ell) = \frac{\langle |\tilde{I}(\ell)|^2 \rangle}{\Omega}, \tag{3}$$

where $\tilde{I}(\ell)$ is given by

$$\tilde{I}(\ell) = d\Omega \sum_{\boldsymbol{\theta}} I(\boldsymbol{\theta}) e^{i\ell \cdot \theta}, \tag{4}$$

where Ω is the survey solid angle and $d\Omega$ is the pixel size. Thus, over a narrow field of view, we need only evaluate a Fourier transform to estimate the angular power spectrum. Writing this out in detail gives[5]

$$C(\ell(a, b)) = \left| \sum_{m,n} I(m, n) \exp\left(-\frac{2\pi i}{N}(am + bn) \right) \right|^2 \frac{d\theta^2}{N^2}, \tag{5}$$

where $d\theta = d\theta_x = d\theta_y$ is the pixel size, $N \equiv N_x = N_y$ is the number of pixels on a side of a square image, and $-N/2 \leqslant a$, $b \leqslant N/2$ are integers. Further, $\ell^2 = \ell_x^2 + \ell_y^2$, and

$$\ell_x = 2\pi a/N d\theta, \tag{6}$$

$$\ell_y = 2\pi b/N d\theta. \tag{7}$$

Note that $C(\ell)$ has the units of $[I]^2 \cdot d\theta^2$, and we often work with $\Delta(\ell) = [\ell(\ell + 1)C(\ell)/2\pi]^{1/2}$, which has the same units[6] as I. Here $C(\ell)$ is the average of $C(\ell)$ within the 1D power spectrum bin ℓ.

The 3D 21 cm power spectrum is often cylindrically binned from 3D $\boldsymbol{k}$ space to 2D $(k_\perp, k_\parallel)$ space, where $k_\perp^2 \equiv k_x^2 + k_y^2$ represents modes perpendicular to the line of sight and $k_\parallel = k_z$ represents modes along the line of sight. We show in Appendix B that this cylindrically binned power spectrum is related to the angular power spectrum of a broadband image (over a narrow field of view) as

$$P(k_\perp, k_\parallel = 0) = D_c^2 \Delta D_c C(\ell(k_\perp)). \tag{8}$$

Here $\ell = D_c k_\perp$, where D_c is the comoving line-of-sight distance to the center of the cube and ΔD_c is the comoving depth of the cube.

2.2. Cross Spectrum versus Coherence

The 3D and 2D cross spectra are defined, extending Equations (1) and (3) to the cross spectrum, as

$$P_{12}(\boldsymbol{k}) = \frac{\langle \tilde{I}_1^*(\boldsymbol{k}) \tilde{I}_2(\boldsymbol{k}) \rangle}{V}, \tag{9}$$

$$C_{12}(\ell) = \frac{\langle \tilde{I}_1^*(\ell) \tilde{I}_2(\ell) \rangle}{\Omega}, \tag{10}$$

where 1 and 2 denote the 21 cm and the IR fields, respectively. The cross spectrum is a quantity that ranges between $\pm(C_1(\ell)C_2(\ell))^{1/2}$ in the 2D case, depending on how correlated, uncorrelated, or anticorrelated the two fields are. It is thus often renormalized as

$$c_{12}(\ell) \equiv \frac{C_{12}(\ell)}{\sqrt{C_1(\ell)C_2(\ell)}}, \tag{11}$$

where c is known as the coherence and is insensitive to a simple rescaling of either field. However, uncorrelated foreground residuals in either field will substantially bias the coherence toward zero (Furlanetto Lidz 2007; Lidz et al. 2009), whereas they merely contribute a zero mean noise to the cross spectrum. Slight foreground correlations, of course, will bias the cross spectrum as well, as we explore later.

3. Observations and Imaging

3.1. 21 cm Observations

The MWA is a low-frequency radio interferometer in Western Australia consisting of 128 phased array tiles, each with $\sim 30° \times (150\,\mathrm{MHz}/f)$ beams (FWHM) and steerable in few-degree increments with a delay line beamformer. We use low-frequency observations of a quiet field centered at (R.A., decl.) = $(0°, -27°)$ J2000 recorded over 30.72 MHz bandwidth centered at 186 MHz, corresponding to $z = 6.0$–7.3 for rest-frame 21 cm.

We use MWA image products produced by Beardsley et al. (2016). The MWA observations are recorded as 2-minute "snapshots" that are flagged for radio frequency interference using COTTER (Offringa et al. 2015) and then calibrated and imaged using Fast Holographic Deconvolution (FHD).[7] Model visibilities are simulated from a foreground model of diffuse and point-source (Carroll et al. 2016) emission in the field and used for both calibration and foreground subtraction. For each snapshot, FHD produces naturally weighted data and model image cubes, as well as primary and synthesized beam cubes. FHD outputs these "cubes" in HEALPix format per frequency. Note that this processing is performed in parallel on "odd" and "even" data cubes whose data are interleaved at a 2 s cadence for the purpose of avoiding a thermal noise bias in autospectrum analyses. In cross-spectrum analyses, the noise between the radio and IR images is independent, so in principle we should average the odd and even cubes together to achieve the lowest noise. However, for simplicity, we use only the even cube in this work given that thermal noise is much smaller than residual foregrounds.

Following Dillon et al. (2015), we rotate these HEALPix maps so that the MWA field center lies at the north pole, and then we project the pixels onto the xy plane to obtain naturally weighted image space cubes of the raw data (I_{nat}), the model data ($I_{\mathrm{nat,mod}}$), the synthesized beam (I_w) (i.e., the Fourier transform of the uv weights), and the primary beam, all in orthographic projection. We flag the upper and lower 80 kHz channels in each of 24 coarse channels across the band to mitigate aliasing, average each cube over frequency to make
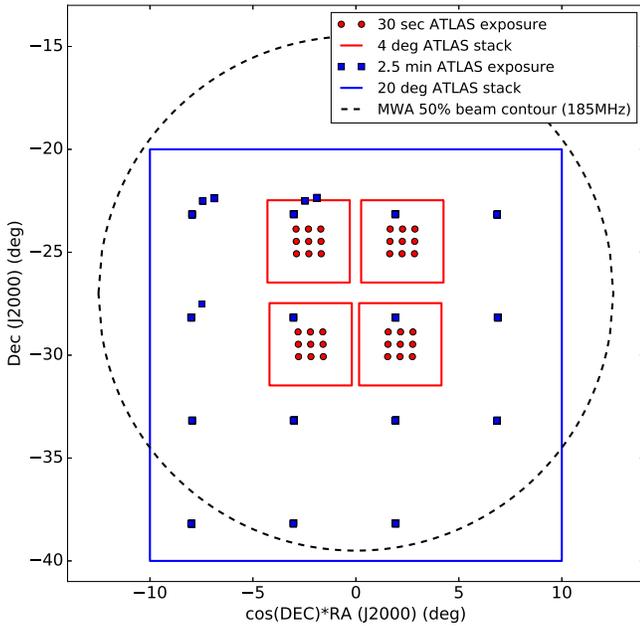

Figure 1. MWA deep integration field (black dashed circle) shown relative to our two ATLAS surveys. Blue filled squares show the observation centers of our wide ATLAS survey aimed at studying foregrounds, and the large blue square outline shows the stacked image. Red filled circles show observation centers for our slightly deeper survey, and red square outlines show the four stacked images we generate. Note that the ATLAS field of view is $5°.5$.

broadband images, and then apply uniform weighting using

$$I_{uni}(\boldsymbol{\theta}) = \frac{10^{-26}\lambda^2}{k_B} \sum_{\boldsymbol{u}} \frac{\tilde{I}_{nat}(\boldsymbol{u})}{\tilde{I}_w(\boldsymbol{u})} e^{-2\pi i \boldsymbol{\theta} \cdot \boldsymbol{u}} d^2 u, \qquad (12)$$

where $\tilde{I}_i(\boldsymbol{u}) = \sum_{\boldsymbol{\theta}} I_i(\boldsymbol{\theta}) e^{2\pi i \boldsymbol{\theta} \cdot \boldsymbol{u}} d\Omega$ for $i = $ nat, w. The units of I_{nat} and I_w are Jy sr^{-1} and sr^{-1}, respectively, as seen from their approximate definitions of $I_{nat}(\boldsymbol{\theta}) \approx \sum_j V(\boldsymbol{u}_j) e^{-2\pi i \boldsymbol{u}_j \cdot \boldsymbol{\theta}} du^2$ and $I_w(\boldsymbol{\theta}) \approx \sum_j e^{-2\pi i \boldsymbol{u}_j \cdot \boldsymbol{\theta}} du^2$, where the sums are over all measured visibilities $V(\boldsymbol{u}_j)$ in units of Jy. These definitions are only approximate because FHD performs corrections to account for wide-field effects.

3.2. IR Observations

ATLAS is a system of multiple 0.5 m $f/2$ wide-field telescopes (Tonry 2011) designed for near-Earth asteroid detection and tracking. Two telescopes are currently in operation located on the islands of Maui and Hawaii. Each telescope has a single $10,560 \times 10,560$ STA1600 CCD sensor, with a pixel scale of $1''.86$, giving a field of view of $5°.5$ on a side. We observe in the Johnson I band (810 nm center with 150 nm FWHM) with the Maui telescope, corresponding to $z = 5.1–6.3$ for Lyα. While this redshift range does not exactly match that of our radio observations, it overlaps sufficiently for our purpose of characterizing the noise and foregrounds in 21 cm–Lyα cross-correlation experiments. The automatic processing pipeline provides single sky-flattened (using a median of the nightly science exposures) images registered to nominal Two Micron All Sky Survey astrometry.

We perform two separate observing campaigns, which we illustrate in Figure 1. We first perform a wide survey to best characterize bright foregrounds. We raster-scan a roughly $20° \times 20°$ grid with $5°$ spacing over the MWA field (dashed black circle), integrating for 2.5 minutes at each pointing (blue

filled squares). The observations were conducted between 2016 September 07 22:00 and 2016 September 08 00:50 Hawaii–Aleutian Standard Time, when the Moon was 36% illuminated. We use `swarp`[8] (Bertin et al. 2002) to stack all these frames over a $20°$ orthographic field centered on (R.A., decl.) = $(0°, -30°)$ (large blue square) with $1''.86$ resolution, using the default background subtraction settings to mitigate temporal and spatial background variation.

Our second campaign is a slightly deeper survey designed to better mitigate airglow fluctuations and CCD systematics for the purpose of studying faint foregrounds below the detection limit. We select four 5 deg fields positioned around the MWA beam peak for best cross-correlation precision: (R.A., decl.) = $(-2°.5, -24°.5)$, $(2°.5, -24°.5)$, $(-2°.5, -29°.5)$, $(2°.5, -29°.5)$ (J2000). We raster-scan a 3×3 grid of 30 s observations within each field (red filled circles) intended to mitigate slight amplifier nonuniformities across the CCD array. The observations were conducted on 2016 November 02 between 21:47 and 23:11 Hawaii–Aleutian Standard Time, when the Moon was 5% illuminated.

We stack the frames in each of the four deep fields using `swarp` over only the central $4° \times 4°$ region, over which all nine 30 s frames overlap (red squares). Otherwise, slight background discontinuities would be introduced by the different temporal coverage of different regions of the stack. In this stacking we disable background subtraction for the purpose of studying the effects of airglow-induced diffuse backgrounds.

4. Point-source Foregrounds

In this section we show with data and simulations that geometric effects can introduce slight correlations between radio and infrared foreground fluxes in broadband surveys, and we quantify how these correlations vary with source masking depth. As the foregrounds are so much brighter than the EOR emission, even slight foreground correlations can bury the predicted EOR anticorrelation, though the expected sign difference will help to identify which effect has been detected.

4.1. Catalogs

To characterize the bright sources relevant to broadband 21 cm–Lyα and 21 cm–Hα intensity mapping correlation measurements, we calculate the correlations between catalogs at 185 MHz, 850 nm, and 4.5 μm as a function of mask depth. These bands correspond roughly to 21 cm, Lyα, and Hα, respectively, from $z \sim 6–7$.

We use the 185 MHz catalog reduced from deep observations of the MWA field depicted in Figure 1 by Carroll et al. (2016). Figure 2 (left panel) shows a histogram of source fluxes in this field. The survey depth varies somewhat over the MWA field owing to the varying primary beam, resulting in a catalog that is 50% complete down to 150 mJy and 95% complete down to 250 mJy. These completeness levels are shallower than the 70 mJy completeness quoted by Carroll et al. (2016) owing to our large rectangular field. The MWA's intrinsic astrometry is at the $2'–3'$ level, though Carroll et al. (2016) cross-match with higher-frequency catalogs to achieve order $10''$ astrometry.
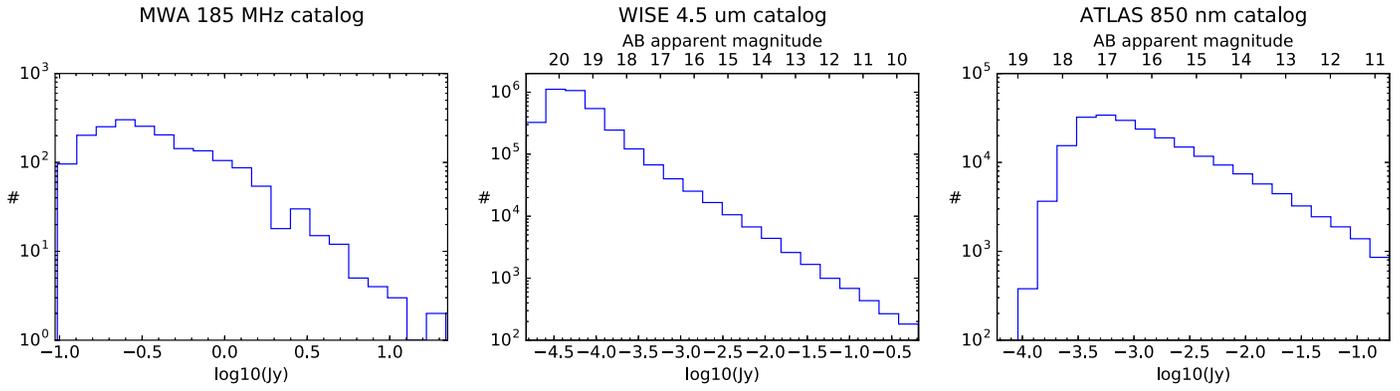
---

[8] http://www.astromatic.net/software/swarp

Figure 2. Histogram of source fluxes in the 185 MHz catalog (left), the 4.5 μm catalog (center), and the 850 nm catalog (right). The catalogs are complete to roughly 250, 0.09, and 0.5 mJy, respectively.

We use the W2 band of ALLWISE (Wright et al. 2010; Cutri et al. 2013) as our 4.5 μm catalog. We download the list of sources within the MWA field using the All Sky Search on the NASA/IPAC Infrared Science Archive[9] and plot the histogram of source fluxes in Figure 2 (middle panel). This ALLWISE band is specified to be 95% complete down to 88 μJy (15.7 AB mag), though it has slight sky coverage nonuniformities due to satellite coverage.

Lastly, we run SExtractor[10] (Bertin & Arnouts 1996) on our wide 20° ATLAS composite image to generate an 850 nm catalog. We allow local background bias and noise estimation, set pixel saturation at 20,000 counts to avoid artifacts, and use the AUTO aperture profile. We extract sources down to 3σ above the background, in order to achieve the most complete point-source mask. Given that our ATLAS observations have been calibrated and imaged through a preliminary pipeline, we cross-match these sources with sources closer than $1''$ in the AAVSO[11] Photometric All Sky Survey[12] (Henden et al. 2016), which is complete to $\sim$3 mJy. We find matches for $\sim$20% of ATLAS detections, not unreasonable given the deeper flux limit of ATLAS compared to APASS. The top panel of Figure 3 shows a 2D histogram of APASS versus ATLAS magnitude as a function of ATLAS magnitude. We fit a Gaussian to the relative magnitude for sources brighter than 13 mag and find that our roughly calibrated ATLAS sources are too bright by 0.279 ± 0.003 mag. Applying this correction, we plot a histogram of ATLAS source fluxes in the right panel of Figure 2, finding that our survey is complete to roughly 0.5 mJy, a factor of 6 deeper than the APASS survey.

4.2. Catalog Radio–Infrared Flux Correlations

Having prepared catalogs of point-source foregrounds in our three bands, we proceed to study how they manifest in intensity mapping correlation experiments. Traditionally, radio/infrared correlations have been studied by cross-matching high-frequency radio detections with infrared sources coincident within a few arcseconds and then plotting radio versus infrared luminosity. Such studies have revealed the well-known radio–far-infrared correlation thought to be due to massive star formation (e.g., de Jong et al. 1985; Helou et al. 1985; Xu et al. 1994; Yun et al. 2001; Willott
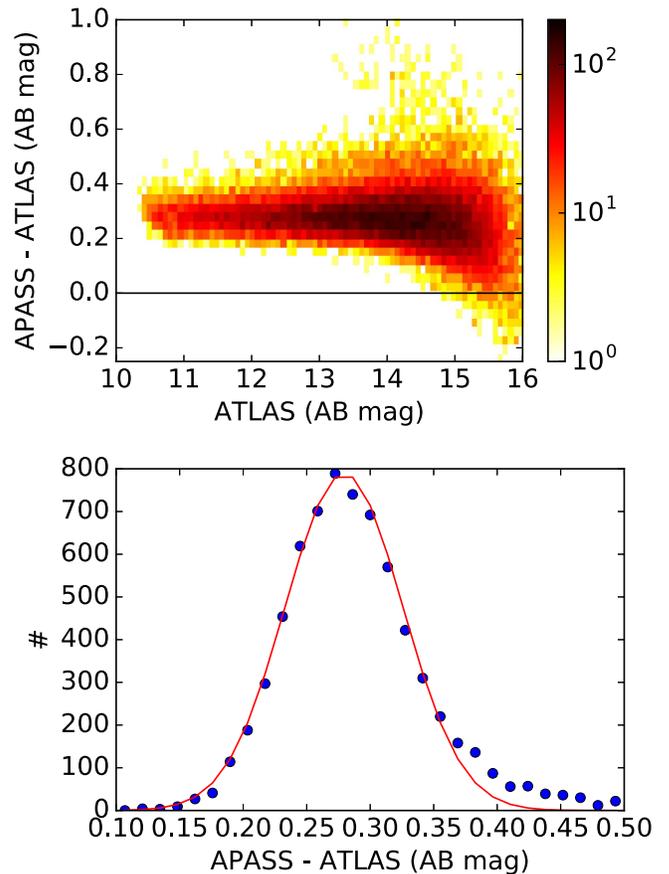
_____

[9] http://irsa.ipac.caltech.edu
[10] http://www.astromatic.net/software/sextractor
[11] American Association of Variable Star Observers.
[12] https://www.aavso.org/download-apass-data

Figure 3. To improve the rough initial ATLAS calibration, we cross-match ATLAS sources with those from APASS, plot a 2D histogram (top) of the relative magnitude as a function of ATLAS magnitude, and then fit a Gaussian to the magnitude offset for sources brighter than 13 mag. We find that our roughly calibrated ATLAS sources are too bright by 0.279 ± 0.003 mag.

et al. 2003; Mauch & Sadler 2007). Massive stars blow out ionized bubbles, generating radio free–free emission correlated with the ionizing flux. Some fraction of these ionizing photons are absorbed by dust clouds and reprocessed into far-infrared emission (Xu et al. 1994). At radio frequencies lower than $\sim$10 GHz, synchrotron dominates over free–free emission, and the correlation is thought to arise from the acceleration of cosmic-ray electrons in these stars' supernovae.

Our approach is different. For all the advantages of broadband intensity mapping, sources cannot be localized to specific redshifts, meaning that it is foreground fluxes, not luminosities, whose correlations are of interest. Of course, compact foregrounds may be masked or subtracted to some residual level, but any correlation of these residual foreground fluxes could bury the EOR correlation. We begin in this section by analyzing foreground fluxes as a function of masking depth, and in the next section we turn to the foregrounds in residual images below the detection limit of these catalogs.

We begin by gridding all three catalog fluxes in Jy to the $20° × 20°$ grid centered at (R.A., decl.) = $(0, -30°)$ depicted in Figure 1 at $5'$ resolution and calculating the zero-lag correlations between the images as

$$c = \frac{\langle I_{\mathrm{rad}} I_{\mathrm{IR}} \rangle - \langle I_{\mathrm{rad}} \rangle \langle I_{\mathrm{IR}} \rangle}{\sqrt{(\langle I_{\mathrm{rad}}^2 \rangle - \langle I_{\mathrm{rad}} \rangle^2)(\langle I_{\mathrm{IR}}^2 \rangle - \langle I_{\mathrm{IR}} \rangle^2)}}, \qquad (13)$$

where $\langle \rangle$ denotes an average over pixels, and the uncertainty due to sample variance is $\Delta c \approx N_{\mathrm{pix}}^{-1/2}$, where N_{pix} is the total number of pixels in the image. Between MWA and *WISE* catalogs we find $c = -0.003 \pm 0.005$, and between MWA and ATLAS catalogs we find $c = 0.001 \pm 0.005$. Both are consistent with zero, as expected, as the brightest sources in both infrared catalogs are likely stars, whose radio emission is expected to be negligible. As a first experiment, we recalculate these correlations after excluding the brightest 10% of sources in all three catalogs, effectively masking down to $10^{-3.75}$ Jy at $4.5\,\mu\mathrm{m}$ and 10^{-2} Jy at 850 nm, and find $c_{\mathrm{MWA-WISE}} = 0.031 \pm 0.005$ and $c_{\mathrm{MWA-ATLAS}} = 0.0086 \pm 0.005$. The former is a 6σ detection and merits some investigation. How does this apparent correlation depend on the flux cut? What is it due to? And what does it mean for broadband correlation experiments? Further, does the MWA–ATLAS correlation remain consistent with zero at stricter flux cuts?

To begin to answer these questions, we plot in Figure 4 the 185 MHz–4.5 $\mu\mathrm{m}$ correlation (top left) and 185 MHz–850 nm correlation (top right) as a function of the masking depth (i.e., the maximum flux of remaining sources). We plot the S/Ns of these correlation measurements, taking the noise to be $N_{\mathrm{pix}}^{-1/2}$ as described above, in the next row. Note that adjacent cells in the correlation matrix plots are somewhat correlated, so a consistent positive sign is not in and of itself evidence of significance. We assess significance by comparing each correlation measurement individually with the expected noise (the S/N), as well as by checking that the correlation vanishes when the 185 MHz image is flipped (bottom two rows).

The 185 MHz and 4.5 $\mu\mathrm{m}$ catalogs exhibit a positive correlation peaking at 0.0332 ± 0.005 after masking infrared sources down to 10^{-4} Jy (18.9 mag) and radio sources down to 1 Jy, and remain significant down to the completeness limits of these catalogs. There is no significant correlation detection after flipping the 185 MHz image, indicating that this detection is not an artifact of the analysis or of primary beam or vignetting effects. The 185 MHz and 850 nm catalogs exhibit a marginal 3σ correlation after masking infrared sources down to 10^{-3} Jy (16.4 mag) and radio sources down to 0.3 Jy, though it does not appear significant in comparison to the level of correlation noise in the flipped image.

To understand these findings, we begin by investigating which 4.5 $\mu\mathrm{m}$ sources are responsible for this correlation. We select the subset of sources detected in the *WISE* 3.4 $\mu\mathrm{m}$, 4.5 $\mu\mathrm{m}$, and 12 $\mu\mathrm{m}$ bands and plot them (Figure 5, left panel) in the $W_{23} \equiv [4.6\,\mu\mathrm{m}]$–$[12\,\mu\mathrm{m}]$ versus $W_{12} \equiv [3.4\,\mu\mathrm{m}]$–$[4.6\,\mu\mathrm{m}]$ color–color space used by Wright et al. (2010) to illustrate the separation between different types of sources. Nikutta et al. (2014) study more quantitatively how sources separate in this space, finding that stars are isolated in the region $W_{12} = -0.04 \pm 0.03$, $W_{23} = 0.05 \pm 0.04$ (1σ). In the right panel, we plot the faintest 90% of sources (fainter than 18.25 mag at 4.6 $\mu\mathrm{m}$) in the same color–color space and observe that this cut effectively cleanly excludes nearly all the stars. This explains the detection of a 185 MHz–4.5 $\mu\mathrm{m}$ correlation only after masking the brightest 10% of sources.

To further probe which mid-infrared sources are responsible for this correlation, we make a rough cut to separate quasars and active galactic nuclei (AGNs; $W_{12} > 0.6$) from star-forming (SF) galaxies ($W_{12} < 0.6$; Mingo et al. 2016). In Figure 6 we plot the power spectrum of 185 MHz sources (left panel), 4.5 $\mu\mathrm{m}$ AGNs (middle panel, blue points), and 4.5 $\mu\mathrm{m}$ SF galaxies (middle panel, red points). In the right panel we plot the coherence (i.e., normalized cross spectrum) of the 185 MHz catalog with AGNs (blue) and with SF galaxies (red). The AGN cut exhibits no significant correlation with the 185 MHz sources, while the SF galaxy cut exhibits a significant correlation rising from a few percent at $\ell \sim 7000$ to 16% at $\ell \sim 300$.

The fall of the correlation toward high ℓ is likely due to the MWA's $3'$ resolution at 185 MHz, corresponding to a maximum ℓ of roughly 4000. Both the falling 4.5 $\mu\mathrm{m}$ catalog power spectrum and the relatively flat 185 MHz power spectrum are functions of the detailed properties of these surveys. Tegmark et al. (2002) and Dodelson et al. (2002) show that the galaxy angular power spectrum $C(\ell)$ is approximately equal to the 3D matter power spectrum $P(k(\ell))$ convolved with a window function that depends on the redshift coverage and flux limit of the sample. The matter power spectrum is known to rise as k^1 for $k \lesssim 0.02$ h Mpc^{-1} before falling as k^{-3}. Galaxy surveys typically probe the regime just after the turnover where the slope is transitioning from 0 to -3 (Tegmark & Zaldarriaga 2002). In order to maximize its sensitivity to low surface EOR 21 cm emission, the MWA was designed as a relatively compact array in comparison to higher-resolution radio interferometers such as the Very Large Array. This low resolution makes the MWA catalog severely flux limited (Carroll et al. 2016), which in turn effectively masks many galaxies that would otherwise be seen. This large masked volume translates into a wide Fourier convolution kernel, explaining why the MWA catalog power spectrum is so flat.

Lastly, we hypothesize that the absence of an observed 185 MHz–850 nm correlation is due to the larger fraction of stars in 850 nm images than in 4.5 $\mu\mathrm{m}$ images. To check this, we study the fraction of stars and galaxies in a similar Galactic field observed by the Sloan Digital Sky Survey (SDSS; Eisenstein et al. 2011) DR13 (SDSS Collaboration et al. 2016). SDSS catalog[13] objects are classified as either stars or galaxies using *ugriz* photometry, though unfortunately SDSS does not reach the declination of our MWA field at (R.A., decl.) = $(0, -27°)$. We thus use a $5°$ field centered at (R.A.,
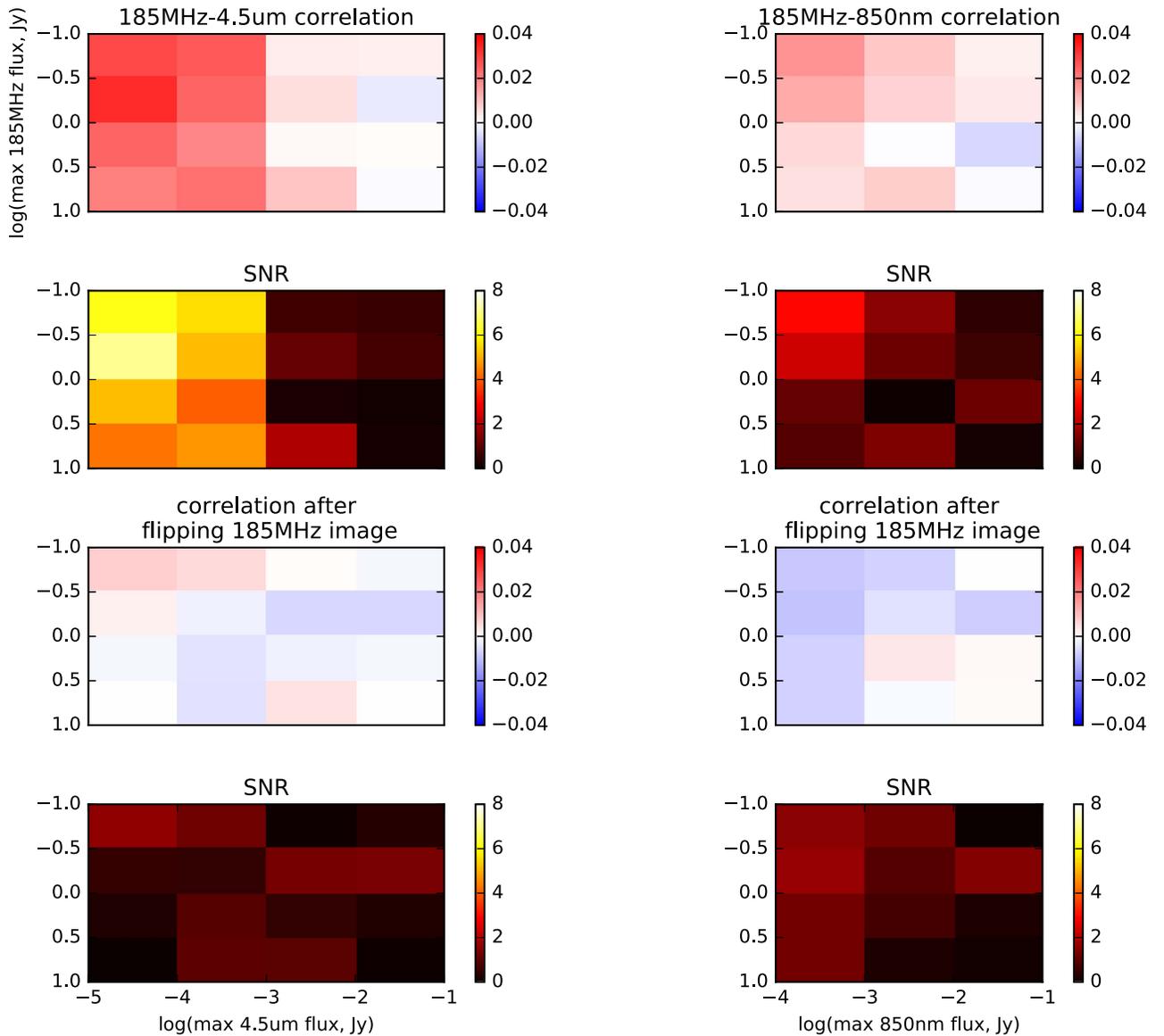
---

[13] http://skyserver.sdss.org/dr13/en/tools/search/rect.aspx

Figure 4. Image space correlation coefficient between 185 MHz and 4.5 μm sources (top left) and between 185 MHz and 850 nm sources (top right), both as a function of radio and infrared flux cuts. In the second row we calculate the S/N in each bin. The bottom two rows show the correlations and S/Ns after flipping the 185 MHz image about its vertical axis before the correlation calculation, giving an independent estimate of the noise. A significant correlation between 185 MHz and 4.5 μm sources appears after masking infrared sources down to 1 mJy, which we show in Figure 5 cleanly removes the stars from the sample. The 185 MHz and 850 nm sources exhibit a marginal correlation after masking down to 1 mJy, though it does not appear significant in comparison to the noise correlation seen after flipping the radio image. We show in Figure 7 that this is due to increased stellar contamination at 850 nm compared to that at 4.5 μm.

decl.) = (205°, 22°), which has the same Galactic longitude but is flipped to the other side of the Galactic plane, giving a similar line of sight through the Galaxy.

In Figure 7 we plot a histogram of the fluxes of the objects marked as stars (red) and galaxies (blue) in this field at the SDSS i band (762 ± 65) nm. We observe that stars dominate the field above 10^{-4} Jy; this is a factor of a few below the survey depth of our wide ATLAS 850 nm image, and thus none of the flux cuts explored above were deep enough to reveal the extragalactic sources. This is consistent with the fact that the 185 MHz–4.6 μm correlation only appeared after the stars were removed, a procedure easier in the mid-infrared than the near-infrared.

4.3. Simulations of Distance-induced Flux Correlation

Let us now consider why the 4.5 μm SF sample is 5%–15% correlated with the 185 MHz catalog, while the AGN sample is

not. Of course, some slight correlation is expected simply because brighter AGNs typically reside in more massive galaxies, which are typically brighter in stars (see, e.g., Figure 1 in Seymour et al. 2007, or Figure 4 in Willott et al. 2003), but Mauch & Sadler (2007) find no strong correlation between radio and *near*-infrared luminosities. As discussed above, though, broadband correlation intensity mapping experiments are affected not only by luminosity correlations but by flux correlations as well. We show in this section that fluxes in two different bands may appear correlated owing to geometric effects even when their intrinsic luminosities are completely independent of each other. By geometric effects we refer to to the fact that more distant objects are generally weaker in all bands than nearer objects.

We first make a few approximations to build intuition and then simulate the effect as a function of source masking depth.
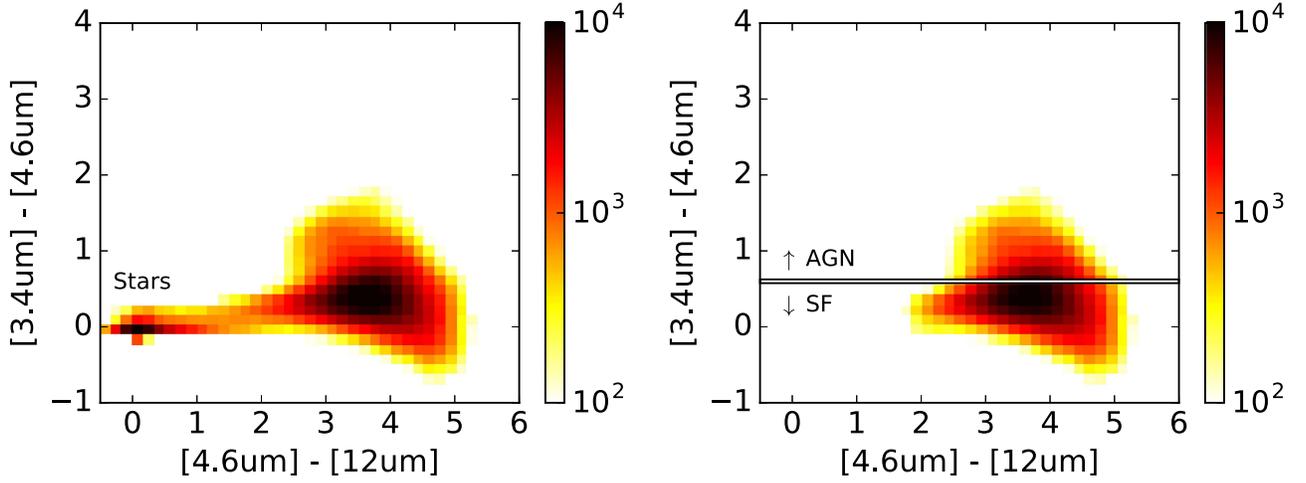
Figure 5. The ALLWISE sources are plotted in the color–color space of Wright et al. (2010) prior to any flux cuts (left panel), showing a cluster of stars near (0, 0). The color bar shows the number of sources in each cell. Then after cutting out the brightest 10% of sources (fainter than 18.25 AB mag), the stars are cleanly removed. We roughly split the remaining sources into AGNs ($W_{12} > 0.6$) and SF galaxies ($W_{12} < 0.6$; Mingo et al. 2016), where $W_{12} \equiv [3.4\,\mu\mathrm{m}]$—$[4.6\,\mu\mathrm{m}]$.

Consider a sky survey with fixed field of view of a set of objects with uncorrelated infrared and radio emission. By uncorrelated we mean that the infrared and radio luminosities are independent random variables determined by the infrared and radio luminosity functions, respectively. Assume that the objects are uniformly distributed in space out to $z \sim 0.5$, and work in Cartesian space for simplicity. We are interested in the effective correlation between radio and infrared fluxes in the same sky pixels, but let us approximate this by calculating the correlation between source fluxes in the two bands. Starting from Equation (13), we have

$$c = \frac{\langle F_{\mathrm{rad}} F_{\mathrm{IR}} \rangle - \langle F_{\mathrm{rad}} \rangle \langle F_{\mathrm{IR}} \rangle}{\sqrt{(\langle F_{\mathrm{rad}}^2 \rangle - \langle F_{\mathrm{rad}} \rangle^2)(\langle F_{\mathrm{IR}}^2 \rangle - \langle F_{\mathrm{IR}} \rangle^2)}}, \quad (14)$$

where $\langle \rangle$ denote an average over sources in the catalog. Then making the approximation that the radio and infrared luminosity functions are independent of line-of-sight distance, we can rewrite this equation in terms of moments of these luminosity functions and the distribution of object distances using $F_i = L_i / 4\pi D^2$ for $i =$ radio, IR

$$c = \frac{\beta - 1}{\sqrt{(\beta \alpha_{\mathrm{rad}} - 1)(\beta \alpha_{\mathrm{IR}} - 1)}} \approx \frac{1}{\sqrt{\alpha_{\mathrm{rad}} \alpha_{\mathrm{IR}}}}, \quad (15)$$

where $\beta \equiv \langle D^{-4} \rangle / \langle D^{-2} \rangle^2$ and $\alpha_i = \langle L_i^2 \rangle / \langle L_i \rangle^2$, and again $\langle \rangle$ denote an average over sources in the catalog. The approximation in the last part of this equation results from the fact that for typical parameters and luminosity functions (see below) $\beta > > 1$ and $\alpha > 1$.

For a survey of a fixed angular field of view, uniform spatial distribution of objects, and Cartesian spacetime, the distribution of object distances is $\rho(D) = \rho_0 D^2$. Assuming that all the objects are between distances D_{min} and D_{max}, the normalization constant is $\rho_0 = 3/(D_{\mathrm{max}}^3 - D_{\mathrm{min}}^3)$, and

$$\beta = \frac{\int_{D_{\mathrm{min}}}^{D_{\mathrm{max}}} D^{-4} \rho(D) dD}{\left(\int_{D_{\mathrm{min}}}^{D_{\mathrm{max}}} D^{-2} \rho(D) dD \right)^2} \quad (16)$$

$$= \frac{D_{\mathrm{max}}^2 + D_{\mathrm{max}} D_{\mathrm{min}} + D_{\mathrm{min}}^2}{3 D_{\mathrm{max}} D_{\mathrm{min}}}. \quad (17)$$

We observe that the radio–infrared flux correlation c for some type of objects is a function of their radio and infrared luminosity functions. In fact, we can see immediately that if the luminosity distributions are wide, their α values are large, and c is small. Conversely, if the luminosity functions are narrow, then the distance to the sources plays a more significant role in determining their fluxes, and so c is larger.

To quantify whether this effect can explain our measured radio–infrared correlation in SF galaxies and the lack of one in AGNs, we use AGN and SF luminosity functions at 1.4 GHz from Mauch & Sadler (2007) (Figure 8, right panel) and at 8 μm from Fu et al. (2010) (left panel). The former describe galaxies at $z < 0.3$, while the latter describe galaxies at $z \sim 0.6$. In principle, we should use luminosity functions at our actual radio and infrared bands, 185 MHz and 4.5 μm, and of course this analysis could be extended using proper redshift-dependent luminosity functions, though we find that our simplified analysis suffices to explain our earlier correlation measurements. We leave a more detailed study for future work. Indeed, Prescott et al. (2016) find that the AGN and SF galaxy radio luminosity functions at lower frequencies, specifically at 325 MHz, closely follow those at 1.4 GHz up to an overall scaling that cancels out of our correlation coefficient. We use approximately the same range of luminosities used by Mauch & Sadler (2007) and Fu et al. (2010) and adjust the minimum luminosities slightly to achieve the same number density of each type of object in both radio and infrared surveys. Above these limiting magnitudes (see Figure 8), the number density of AGNs is $\sim 0.0020 \, \mathrm{Mpc}^{-3}$ and that of SF is $\sim 0.00011 \, \mathrm{Mpc}^{-3}$. In the end we find that our results are only weakly sensitive to these luminosity minima, as their faint ends become less and less significant in real, flux-limited samples.

We pick fiducial survey parameters of $d_{\mathrm{min}} = 20$ Mpc and $d_{\mathrm{max}} = 3000$ Mpc ($z_{\mathrm{max}} = 0.75$), giving $\beta \approx 50$. Using the above luminosity functions, we find $\alpha_{\mathrm{SF,IR}} = 1.474$, $\alpha_{\mathrm{SF,rad}} = 14.56$, $\alpha_{\mathrm{AGN,IR}} = 22.97$, and $\alpha_{\mathrm{AGN,rad}} = 257.5$. These values agree with qualitative observation that the AGN luminosity function is wider than the SF luminosity function in both radio and infrared bands (Figure 8). These values give a predicted
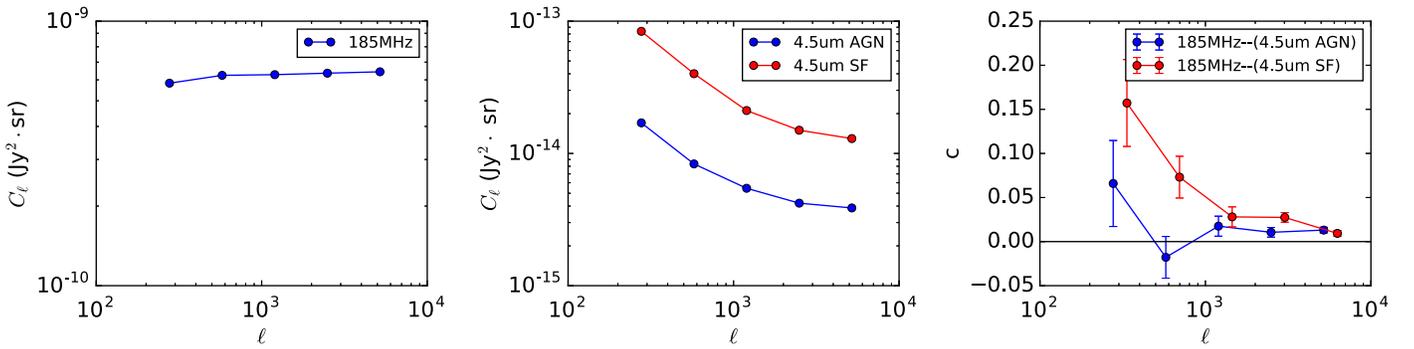
Figure 6. Power spectrum of 185 MHz sources (left panel), 4.5 μm sources (middle panel), and coherence between 185 MHz and 4.5 μm sources (right panel). We roughly separate the 4.5 μm sources into SF galaxies and AGNs as illustrated in Figure 5. We find that the 185 MHz–4.5 μm correlation observed in Figure 4 holds only for the star-forming galaxies in the infrared sample.
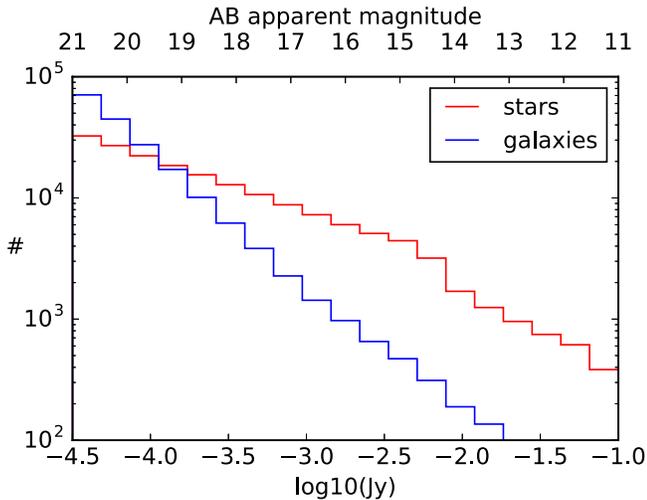

Figure 7. Histogram of fluxes of stars and galaxies in a 5° SDSS field at the same Galactic longitude as the MWA field, but flipped to the other side of the Galactic plane. These fluxes are in the SDSS i band (762 ± 65) nm.

radio–infrared correlation of 0.21 for SF galaxies and 0.01 for AGNs, agreeing with our finding of a significant radio–infrared correlation for SF and near-zero correlation for AGNs. The exact values deviate from our measurements for a number of reasons. The MWA and *WISE* catalogs are not matched in depth or redshift coverage and thus do not survey an exactly overlapping set of radio and infrared sources. Further, these calculations assume a volume-limited survey, in contrast to our flux-limited radio and IR surveys. Additionally, real-world luminosity functions can exhibit redshift evolution. Of course, our measurement in Section 4.2 did not even split up the MWA catalog into separate AGN and SF subsets, as such detailed characterization of low-frequency radio foregrounds remains an active area of research. Lastly, the limited MWA resolution pushes the observed correlation with infrared images to zero at high ℓ, suppressing the overall correlation computed in image space. Future work will be needed to quantify these effects in greater detail and assess their significance in EOR cross-spectrum measurements.

Can this unwanted radio–infrared foreground correlation be mitigated by masking the brightest sources? Using the luminosity functions presented above, we simulate radio and infrared surveys for each of AGNs and SF galaxies. We begin by generating the mock radio catalogs of AGNs and SF galaxies, choosing a Poisson random number of each in each of

400 logarithmic luminosity bins. Using logarithmic bins best samples the large dynamic range of the luminosity functions. We distribute the objects uniformly over a volume $D_{max} = cz_{max}/H_0 = 3212$ Mpc deep and $\theta_{FOV}D_{max} = 1121$ Mpc wide and then pick a random infrared luminosity for each radio object from the appropriate infrared luminosity function. Finally, we plot in Figure 9, along the lines of Figure 4, the predicted 1.4 GHz–8 μm correlation of our mock AGN and SF catalogs after masking down to a maximum radio and infrared flux. As we saw above, without any flux cut we find a roughly 20% radio–infrared correlation for SF galaxies and negligible correlation for AGNs. As we mask fainter and fainter sources, the AGN correlation generally increases to the 5%–10% level, while the SF correlation first increases and then decreases after masking down to 10^{-4} Jy. With increasing mask depth, these correlations do not necessarily approach zero monotonically, and more detailed modeling of effective foreground flux correlations will be necessary in real-world intensity mapping correlation experiments probing the EOR. In the next section we move beyond the bright sources and study the magnitudes and correlation properties of the residual radio and infrared foregrounds in our MWA and ATLAS observations.

5. Residual Foregrounds and Cross-spectrum Limits

In this section we characterize the power spectra and correlation properties of the residual 185 MHz and 850 nm foregrounds after subtracting and masking the bright sources identified by the surveys discussed in the previous section.

5.1. Residual 21 cm Foregrounds

We begin by quantifying the 185 MHz foreground residuals in angular power spectrum measurements. In broadband (i.e., multifrequency synthesis) images, thermal noise quickly integrates below foreground residuals. Reaching the cosmological signal should therefore be a matter of foreground mitigation and not the long time averages needed to measure the 3D power spectrum (e.g., Beardsley et al. 2013; Pober et al. 2014). We check this hypothesis, asking how much observation time is required to achieve the best foreground subtraction.

Working with the MWA image products presented in Section 3.1, we compute the angular power spectrum as

$$C(\ell) = \frac{\sum_\ell |\tilde{I}_{uni}(\ell)\tilde{I}_w(\ell)|^2}{\sum_\ell |\tilde{I}_w(\ell)|^2} \frac{d\theta^2}{N^2} \qquad (18)$$
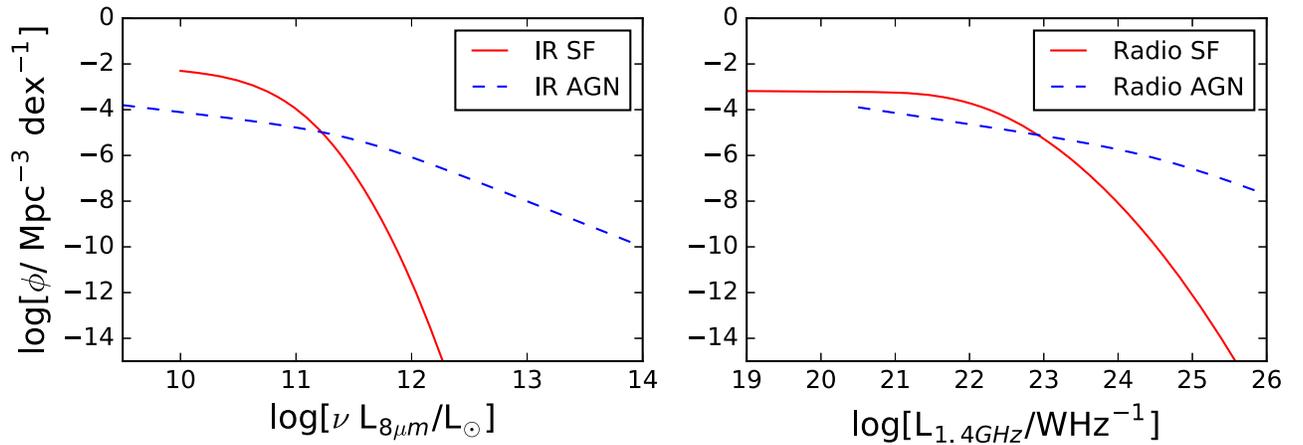
Figure 8. AGN and SF luminosity functions at 8 μm from Fu et al. (2010) (left panel) and at 1.4 GHz from Mauch & Sadler (2007) (right panel).
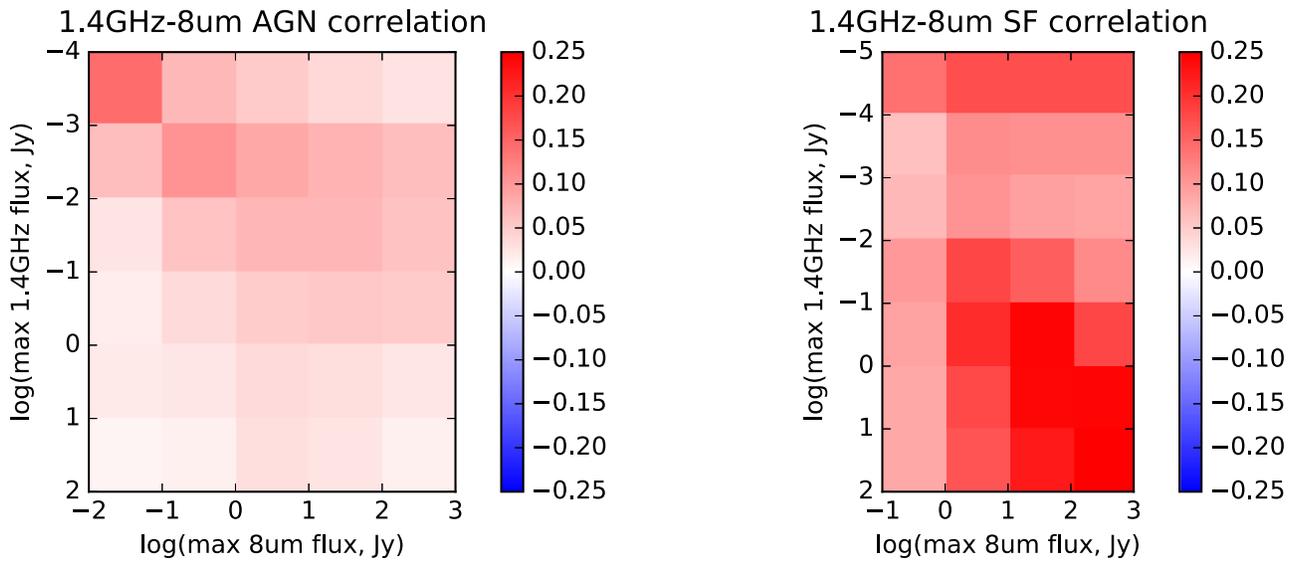

Figure 9. Apparent correlation of radio and mid-infrared fluxes from a mock volume-limited survey with independent radio and mid-infrared luminosities using the luminosity functions from Figure 8 for AGNs and SF galaxies. Without any flux cuts (lower right corner of each plot) we observe a significant correlation between radio and mid-infrared fluxes for SF objects and a near-zero correlation for AGNs. This agrees with our measurements on real 185 MHz and 4.5 μm sources in Figure 4. As fainter and fainter objects are masked, we observe that the AGN correlation gradually strengthens, while the SF correlation weakens somewhat.

summing over all the ℓ values in each ℓ bin, where N is the number of pixels of size $d\theta$ on each side of the square image. Note that $\ell = 2\pi u$, where u is the Fourier dual to angle from the field center assuming the small-angle approximation. We estimate the thermal noise power spectrum by computing the power spectrum of the difference between the interleaved odd and even cubes discussed earlier, which contains only thermal noise.

We plot in Figure 10 the power spectra of 185 MHz broadband images from 3 hr data selections, spread over 1 (red solid), 5 (green solid), and 19 (blue solid) nights. We also make a broadband image of the same 32 hr data set used by Beardsley et al. (2016) (black dashed). We plot the power spectra of the raw (pre-foreground subtraction), residual (post-foreground subtraction), and noise (difference between successive integrations) images out to $\ell = 2600$, corresponding to a maximum baseline length of $\sim$700 m. Beyond this, the uv coverage becomes sparse, introducing artifacts in the application of gridding and uniform weighting, though a more sophisticated analysis could likely use these longer baselines.

We cross-check our imaging and power spectrum analysis by comparing the $k_\parallel = 0$ bin of the 3D power spectrum of Beardsley et al. (2016) converted to an angular power spectrum using Equation (8) (magenta dashed).

We find that the raw power spectra of the different 3 hr data sets agree with each other and with that of the 32 hr data set, as expected given that foregrounds overwhelm thermal noise in a broadband image. Interestingly, the residual power spectrum decreases as the 3 hr are spread over more and more nights until it reaches the level of the deep 32 hr integration, a factor of $\sim$4 lower in power than in the single-night analysis. These findings could be explained by slight ionosphere-related errors that limit the accuracy of each night's calibration and thus of its foreground subtraction. Further work would be needed to understand this effect in detail, but for now our conclusion is that the 32 hr integration has the best foreground subtraction, and we use this deep cube in our later analyses. Note that, as expected, the thermal noise of the deep integration is 10 times lower in power than the 3 hr integrations. Because these are band-averaged images, even
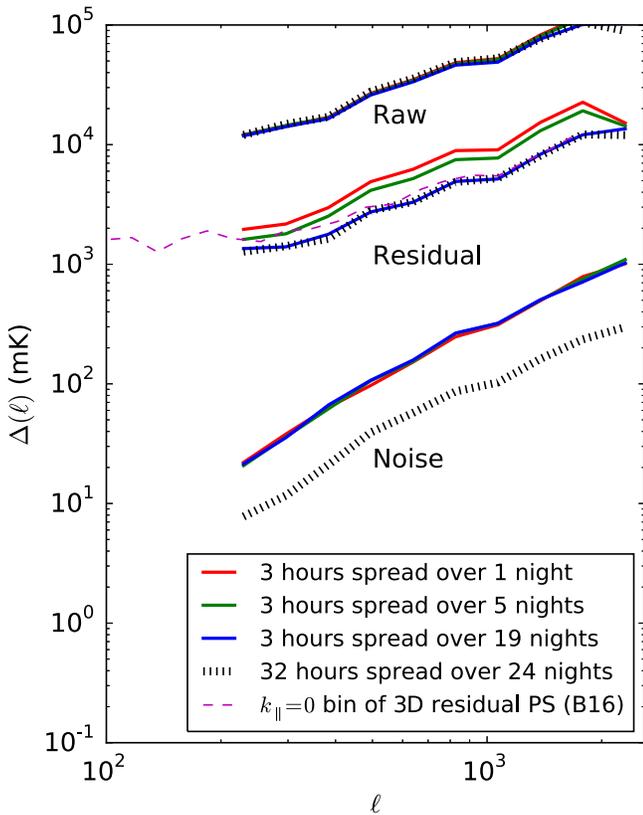
Figure 10. Raw, residual (post-foreground subtraction), and noise power spectra of 185 MHz broadband images from various 3 hr data selections (solid lines) and from the 32 hr data selection used by Beardsley et al. (2016) (black dashed). The uv plane is nearly filled after only 3 hr, yet we find that spacing these ∼100 2-minute integrations over many independent nights reduces the foreground residuals by a factor of up to ∼4 in power. As a cross-check on our analysis, we plot the $k_\parallel = 0$ bin of the 3D power spectrum of Beardsley et al. (2016) converted to an angular power spectrum using Equation (8) (magenta dashed).

the 3 hr thermal noise is at least 100 times lower than its foreground residuals.

5.2. Residual IR Foregrounds

We proceed to generate foreground-masked 850 nm images of each of the four deep ATLAS integration fields shown in Figure 1. Each of these fields is a stack of nine 30 s exposures with 5° field of view, dithered such that the overlap is a ∼4° region. By confining ourselves to this overlap region, we avoid the background discontinuities that affect many nominally wide-field infrared image data sets whose mosaicking introduces significant background patchiness. Mitchell-Wynne et al. (2015) have demonstrated that complex fitting (Fixsen et al. 2000) can help reduce such patchiness in small mosaics (∼10 arcminutes), but further work is required to study whether such techniques can be applied to much larger fields.

Our approach is to mask each image at ATLAS's native $1.''86$ resolution and then coarse grid down to $3.'5$, approximately the resolution of the MWA, taking each coarse pixel's value as the average of all unmasked fine pixels within it. If fewer than 10% of its fine pixels remain unmasked, we consider the whole coarse pixel masked to avoid introducing too much noise variation between different coarse pixels. For illustrative purposes, we proceed in the following four stages,

which we illustrate in Figure 11 for the ATLAS field centered at (R.A., decl.) = (2°.74, −24°.79) (the top right red box in Figure 1). Each row shows the result of an additional masking stage, as outlined below. The left column shows a typical 9′ field to illustrate the masking up close, the middle column shows the resulting coarse binned image with $3.'5$ resolution, and the right column shows the fast Fourier transform (FFT) of the center image (plotted as $\log[\Delta(\ell)/(\mathrm{kJy\ sr}^{-1})]$) in order to identify detector systematics.

1. Mask saturated regions (Figure 11, row 1). Nearly saturated pixels are associated with nearby bright stars that would dominate fluctuation measurements; thus, we mask 4′ around all pixels within 30% of saturation (white regions in left and middle columns). This wide mask removes the broad wings of the point-spread function (PSF) revealed by these extremely bright sources. Roughly 92% of fine pixels remain unmasked after this step, and 96% of coarse pixels remain.

2. Mask sources to 5σ (Figure 11, row 2). We mask circular regions with radius equal to five times the profile rms along the minor axis of each source as measured by SExtractor (see Section 4.1). The reason we use the minor-axis rms is that vertical charge leakage in the CCD array results in unrealistically large major-axis rms measurements for bright sources. After this stage ∼87% of fine pixels remain unmasked, and the fraction of coarse pixels remaining is unaffected.

3. Mask sources to 12σ and mask other emission above $70\ \mathrm{kJy\ sr}^{-1}$ over the background (Figure 11, row 3). We use a larger masking radius to remove the PSF wings around the 5σ source masks. We also determine that the vertical CCD charge leakage can be isolated by looking for emission brighter than $70\ \mathrm{kJy\ sr}^{-1}$ over the background, and that it can be flagged without cutting into the shot noise between sources. A total of 58% of fine pixels remain unmasked after this step, and again the fraction of coarse pixels remaining is unaffected.

4. Mask horizontal and vertical Fourier modes (Figure 11, row 4). The previous two stages revealed detector artifacts within $\Delta\ell \sim 100$ of $\ell_x = 0$ and $\ell_y = 0$. These compact Fourier systematics correspond to slight horizontal and vertical discontinuities in the center image due to imperfect gain matching between 16 different amplifiers that process different rectangular regions of the CCD array. We conservatively mask Fourier modes with $|\ell_x| < 200$ or $|\ell_y| < 200$ to eliminate this effect.

Note that these 850 nm deep observations were recorded during near new moon conditions, and we find that the mean airglow in source-free regions is $\sim 3 \times 10^3\ \mathrm{kJy\ sr}^{-1}$, of order 19 AB mag arcsec^{-2}. For comparison, Sullivan & Simcoe (2012) measure a 1020 nm continuum airglow brightness (i.e., after spectrally masking the OH lines) of 20 ± 0.5 AB mag arcsec^{-2} far away from the Moon.

We proceed to characterize the residual infrared fluctuations in power spectrum space, using the optimal quadratic estimator to properly account for the masking of coarse pixels. This estimator was introduced to astronomy by Tegmark (1997) to recover CMB power spectra from maps with arbitrary survey geometries and noise properties and has recently been revived for 3D power spectrum analysis of 21 cm data by Dillon et al.
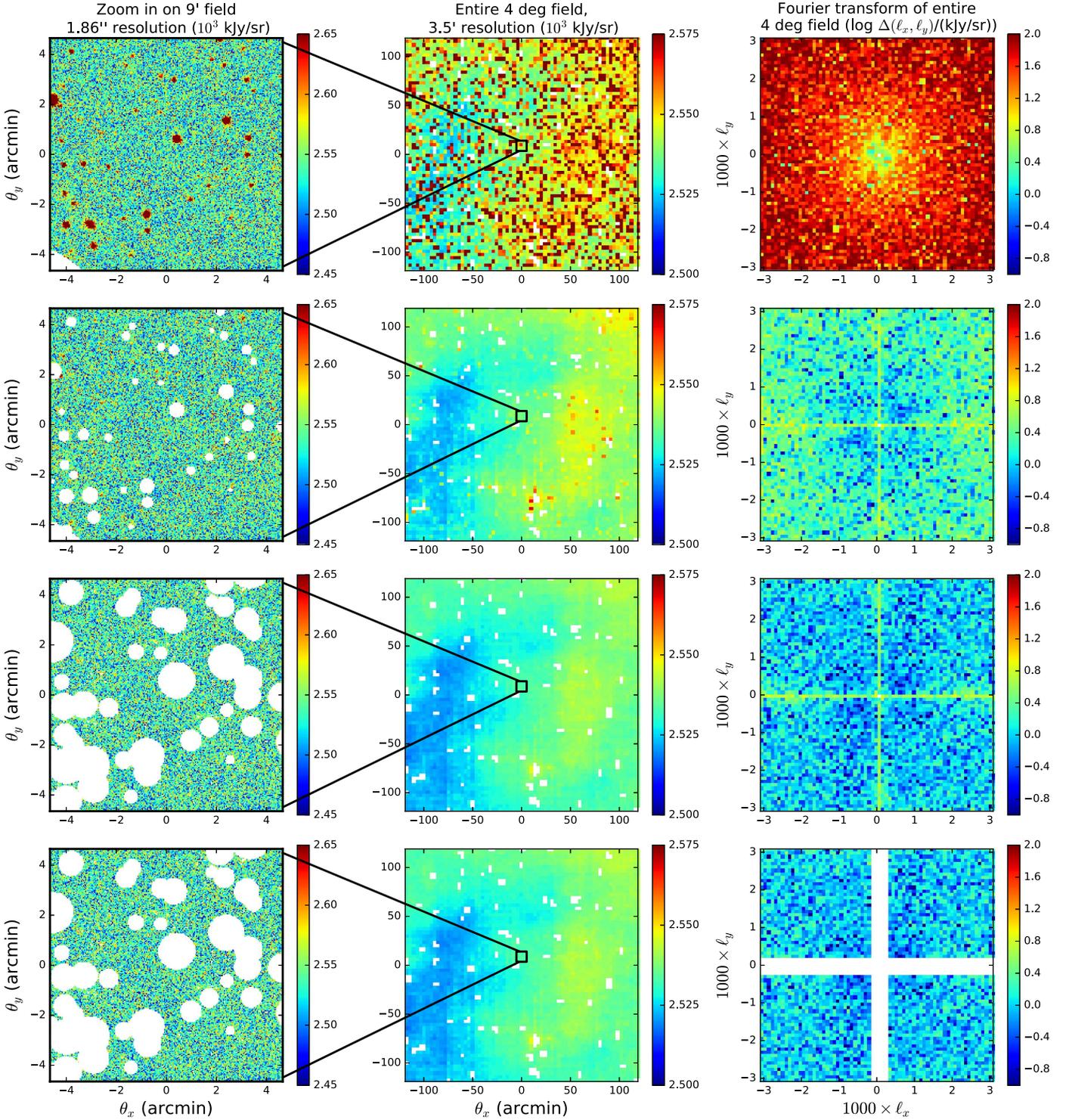
Figure 11. The rows of this figure illustrate our four stages of foreground removal on ATLAS images, as detailed in Section 5.2. We apply this process to all four 4° deep ATLAS fields shown in Figure 1, but here we show the one centered at (R.A., decl.) = 2°74, −24°79) for illustration. The first row shows the results of masking 4′ around all nearly saturated regions, the second row shows the results of masking out to 5σ, the third row shows the results after masking out to 12σ and all pixels above 70 kJy sr^{-1} over the background, and the last row shows the results after masking the nearly horizontal or vertical Fourier modes. In each row, we show the central 9′ field at 1″86 resolution (left), the entire 4° field after coarse gridding to 3′5 (middle), and the FFT of the coarse-gridded field to highlight systematics (right). Note that the left and middle panels of the bottom row are identical to those in the row above it, illustrating that the image space mask is the same; only the Fourier mask is different.

(2014, 2015), Liu & Tegmark (2011), Dillon et al. (2013), and Ali et al. (2015). We employ this estimator in a manner more similar to the original CMB case, using it to account for pixel masking in broad bandpower spectrum estimation. We briefly summarize the estimator here and refer to Dillon et al. (2014) for a more detailed description.

We label the normalized estimator of the power in bin α as p_α, related to the unnormalized estimator q_α as $\boldsymbol{p} = \boldsymbol{Mq}$. Bold

lowercase letters are vectors, while bold uppercase letters are matrices. The unnormalized estimator is given by

$$q_\alpha = \frac{1}{2}(\boldsymbol{x} - \langle\boldsymbol{x}\rangle)^t \boldsymbol{C}^{-1}\boldsymbol{C}_{,\alpha}\boldsymbol{C}^{-1}(\boldsymbol{x} - \langle\boldsymbol{x}\rangle), \quad (19)$$

where $\boldsymbol{x}$ is a column vector containing all N_{pix} pixel measurements, $\boldsymbol{C}$ is the pixel–pixel covariance matrix, and $\boldsymbol{C}_{,\alpha} \equiv d\boldsymbol{C}/dp_\alpha$ is the derivative of the covariance with respect to the power in bin α. Note that $\boldsymbol{C}_{,\alpha} = \boldsymbol{A}^\dagger\boldsymbol{A}$, where $\boldsymbol{A}$ is an $N_\alpha \times N_{pix}$ with elements $A_{ij} = \exp(i\boldsymbol{\theta}_j \cdot \boldsymbol{\ell}_i)$. Here $\boldsymbol{\ell}_i$ refers to the ith out of N_α ℓ modes in bin α. Note that ^{t} denotes a transpose and $\dagger$ donates a conjugate transpose.

The matrix of window functions (i.e., horizontal error bars) of the bandpowers p_α, defined such that $\boldsymbol{p}_{estimated} = \boldsymbol{W}\boldsymbol{p}_{true}$, is given by $\boldsymbol{W} = \boldsymbol{M}\boldsymbol{F}$, where $\boldsymbol{F}$ is the Fisher matrix and $\boldsymbol{M}$ is an arbitrary invertible normalization function encoding the compromise between horizontal and vertical error bars. The covariance between the measured p_α values is given by $\boldsymbol{\Sigma} = \boldsymbol{M}\boldsymbol{F}\boldsymbol{M}^t$. Dillon et al. (2014) argue that taking $\boldsymbol{M} \propto \boldsymbol{F}^{-1/2}$ is a good compromise between small horizontal error bars and small vertical error bars. For simplicity, we take $\boldsymbol{M} = \boldsymbol{F}^{-1/2}$ and correct the normalization at the end by dividing each element of $\boldsymbol{p}$ by the peak of the appropriate row of $\boldsymbol{W}$. In this case, the bandpower variances are given by the reciprocals of the peaks of the window functions used to normalize the bandpowers. We find that using the sum instead of the peak significantly biases the recovered bandpowers downward when the power spectrum is nonflat, as in our case. Lastly, the elements of the Fisher matrix are given by

$$\boldsymbol{F}_{\alpha\beta} = \frac{1}{2}\text{tr}(\boldsymbol{C}^{-1}\boldsymbol{C}_{,\alpha}\boldsymbol{C}^{-1}\boldsymbol{C}_{,\beta}). \quad (20)$$

Now we turn to application of this formalism to power spectrum estimation from our masked IR images. Later we will adapt it to estimation of the 21 cm–IR cross spectrum. We take $\boldsymbol{x}$ to be a vector of all IR coarse (3$'$.5) pixel values, with masked pixels set to zero. After gridding the high-resolution images to reach this resolution, photon shot noise is negligible, and the image space covariance is the sum of the sample variance $\boldsymbol{C}_{signal}$ and the masking covariance $\boldsymbol{C}_{mask}$. $\boldsymbol{C}_{mask}$ is a diagonal matrix with ∞ for masked pixels and 0 otherwise. In practice, we replace ∞ with a number 10^7 times larger than the largest eigenvalue of $\boldsymbol{C}_{signal}$, finding that the results are not sensitive to this parameter. The sample variance is easily obtained by writing it in Fourier space, $\boldsymbol{C}_{ft}$, where it is a diagonal matrix with a guess of the true power spectrum on the diagonal, and then Fourier transforming it into image space with Fourier transform matrix $\mathcal{F}$. Putting these together gives

$$\boldsymbol{C} = \mathcal{F}^\dagger\boldsymbol{C}_{ft}\mathcal{F} + \boldsymbol{C}_{mask}. \quad (21)$$

Note that $\mathcal{F}$ is an $N_{pix} \times N_{pix}$ matrix with elements $\mathcal{F}_{ij} = \exp(-i\boldsymbol{\theta}_i \cdot \boldsymbol{\ell}_j)$, where i runs over all pixels and j runs over all Fourier cells. Said differently, a guess of the power spectrum is necessary to optimally downweight the sample variance noise on the estimated power spectrum. If the accuracy of the guess were in question, one could always iterate by feeding the estimated power spectrum back into the quadratic estimator, though in practice we find that this is not necessary in this work.

Using this estimator, we calculate the power spectrum after each stage of masking and plot the mean spectrum over all four

deep ATLAS fields in Figure 12 (left panel). Instead of predicting the bandpower errors from the input covariances, we conservatively bootstrap the error bars by computing the standard deviation of each bandpower over the four fields. The power spectrum of all 850 nm sources, masking only saturated regions, rises proportionally to ℓ (red dots), as expected for Poisson source counts when the power spectrum is plotted as $\Delta = \sqrt{\ell^2 C_\ell/2\pi}$. Masking sources out to 5σ removes two orders of magnitude in power (green dots), and masking out to 12σ and above 70 kJy sr^{-1} over the background removes a factor of a few more in power (blue dots). Lastly, excluding modes with $|\ell_x| < 200$ or $|\ell_y| < 200$ gives our final 850 nm anisotropy spectrum (black dots). We note that more stringent masking does not significantly alter this final result. After the first masking stage, we use a flat power spectrum in $C(\ell)$ as our guess in the quadratic estimator formalism, and after the other three masking steps we use the 1.1 μm power spectrum from Zemcov et al. (2014).

In Figure 12 (right panel) we compare our final residual power spectrum measurement with other measurements[14] in the literature. Our 850 nm ATLAS spectrum (black dots) agrees very well with the 1.1 ± 0.25 μm CIBER spectrum (Zemcov et al. 2014) (blue dots), with much smaller error bars at the ℓ modes we can access ($\ell \lesssim 4000$) due to ATLAS's larger field of view.

Zemcov et al. (2014) argue that their spectrum is limited by foregrounds at all ℓ: by diffuse Galactic light (DGL) (i.e., dust) at $\ell \lesssim 1000$, by intrahalo light (IHL; cyan dashed line) at $\ell \sim 1000$, and by galaxy number counts below the flux limit at larger ℓ. ATLAS's 2$''$ resolution is only a factor of 3 finer than CIBER's 7$''$ resolution, so perhaps it is not surprising that the level of galaxy number counts is not appreciably different. The ATLAS points at $\ell > 1000$ are in fact slightly above CIBER's, though the measurements were conducted over slightly different bands. Mitchell-Wynne et al. (2015) demonstrate a dramatic improvement in foreground reduction $\ell \gtrsim 10^3$ in 10$'$ HST mosaics with 0$''$.1 resolution. Their power spectrum (red points) is a factor of $\sim$10 in amplitude below the previous IHL model, and their new fitting suggests that Zemcov et al. (2014) were in fact limited almost exclusively by DGL (red solid line). For comparison, we plot their *new* IHL model (red dashed) and EOR model (gray area). These results show that finer angular resolution can help reduce the foregrounds in the $\ell \sim 10^2$–10^4 modes that can be cross-correlated with 21 cm EOR intensity mapping experiments.

5.3. Modeling the 21 cm–Lyα Cross Spectrum

Before turning to 21 cm–Lyα cross-spectrum measurements with our 185 MHz and 850 nm images, we generate optimistic and pessimistic theoretical cross spectra for comparison. We simulate 21 cm and Lyα cubes using 21 cm power spectra from Pober et al. (2014), Lyα power spectra from Gong et al. (2014), and the coherence between the two fields from Heneka et al. (2016). Combining simulations from all these sources allows us to better estimate the modeling uncertainty. Future work is needed to more self-consistently model these fields and their correlation over a range of possible reionization scenarios and to infer astrophysical parameters from observed cross spectra.

Gong et al. (2014) model the Lyα cross spectrum and plot an uncertainty region over a range of likely values of escape

---

[14] Note that $I_f/(\text{kJy sr}^{-1}) \approx 0.3\lambda I_\lambda/(\text{nW m}^{-2}\text{ sr}^{-1})$ at $\lambda = 1$ μm.
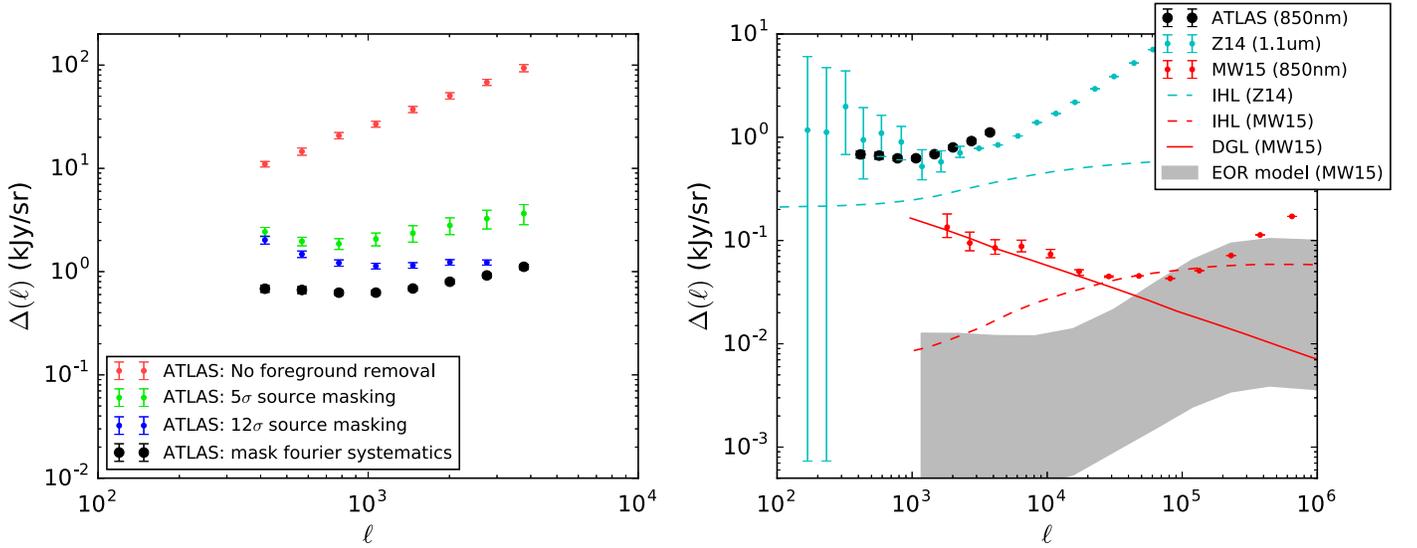
Figure 12. Left panel: power spectra of our broadband ATLAS images after various stages of foreground removal as described in Section 5.2. The error bars show the sample variance noise estimated by the standard deviation over our four 4° fields (see Figure 1). Black dots show our final ATLAS 850 nm power spectrum after all stages of foreground removal. Right panel: we compare our ATLAS power spectrum (black dots) to the 1.1 μm power spectrum of Zemcov et al. (2014) (Z14) (cyan dots) using the CIBER experiment and the 850 nm power spectrum of Mitchell-Wynne et al. (2015) (MW15) using 10′ *HST* mosaics. We show the intrahalo light (IHL) models of Z14 (cyan dashed line) and MW15 (red dashed line), as well as the diffuse galactic light (DGL) model of the latter authors. These results show that finer angular resolution can help reduce the foregrounds in the $\ell \sim 10^2$–10^4 modes that can be cross-correlated with 21 cm EOR intensity mapping experiments. The gray area shows the predicted EOR contribution to the infrared anisotropies from MW15.
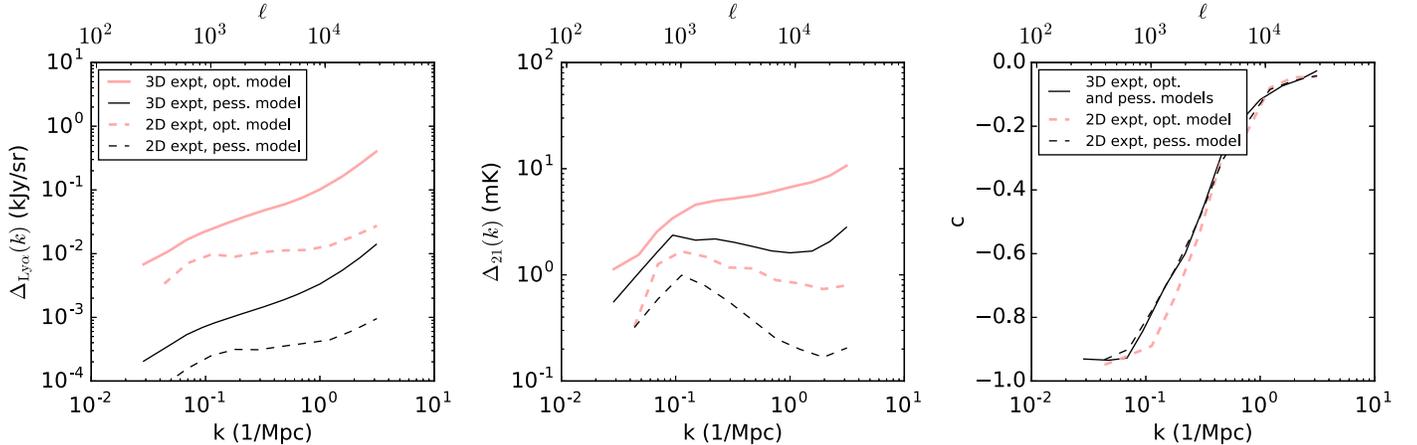

Figure 13. As described in Section 5.3, we identify optimistic (red solid lines) and pessimistic (black solid lines) Lyα (left panel) and 21 cm (middle panel) 3D power spectra from Gong et al. (2014) and from Pober et al. (2014) and Mesinger et al. (2011), respectively, and use the coherence function from Heneka et al. (2016) (right panel, black solid) for both scenarios. We then simulate approximate cubes assuming Gaussian statistics, average them over frequency, and plot the angular power spectra of Lyα (left panel, dashed lines) and 21 cm emission (middle panel, dashed lines) and their coherence functions (right panel, dashed lines). Power on short angular scales ($\ell \gtrsim 10^4$) is suppressed by a factor of $\sim$100 (in power units), while power on longer angular scales ($\ell \lesssim 10^2$) is suppressed by only a factor of $\sim$2 compared to the 3D power spectrum. This indicates that there remains significant observable signal in broadband 21 cm–Lyα cross-spectrum observations.

fraction of ionizing photons, fraction of radiation emitted at Lyα, SF rate, and intergalactic medium (IGM) clumping factor (their Figure 1). We take the upper and lower edges of their uncertainty region as our optimistic and pessimistic power spectra, respectively.

Similarly, Pober et al. (2014) simulate 21 cm power spectra over a range of reionization scenarios using Mesinger et al. (2011), with various values of ionizing efficiency (ζ) (including the escape fraction), the minimum halo virial temperature (T_{vir}), and the ionizing photon mean free path in the IGM (R_{mfp}). Whether the signal at our redshift of interest ($z = 7$) is large or not depends mostly on whether reionization is already largely finished by that time or not. So we take as our pessimistic

model the $z = 8$ power spectrum of the ($\zeta = 31.5$, $T_{vir} = 1.5 \times 10^4$ K, $R_{mfp} = 30$ Mpc) scenario, whose reionization midpoint is $z = 9.5$. We take as our optimistic model a late reionization scenario with $T_{vir} = 3 \times 10^5$ K (other parameters unchanged), whose reionization midpoint is $z = 5.5$.

From these power spectra and coherence functions, we generate approximate 21 cm and Lyα cubes assuming Gaussian statistics. This is an approximation, given that the 21 cm field is expected to become less and less Gaussian as reionization proceeds (Wyithe & Morales 2007), and more work is needed to understand this effect. The simulated cubes have $(1 \text{ Mpc})^3$ resolution over a $(218 \text{ Mpc})^3$ volume at $z = 7$, corresponding to $\Delta z = 0.6$ and 0′.4 angular resolution. Our 21 cm and Lyα

cubes have units of mK and kJy sr^{-1}, respectively, and we average them in the line-of-sight direction to produce broad-band images.

We plot the original 3D spherically averaged power spectra and coherence function as solid lines in Figure 13, and we plot the computed 2D power spectra from the line-of-sight averaged cubes, as well as their coherence, as dashed lines. The 3D and 2D power spectra are plotted as $\Delta(k)$ and $\Delta(\ell)$, respectively, as defined in Section 2.1. As expected, line-of-sight averaging tends to remove power on short spatial scales (compared to the line-of-sight depth of the cube), but it acts similarly on both cubes, so the coherence function is preserved. Note that the cross spectrum Δ_{12} is defined in terms of the coherence as $\Delta_{12}^2 = c\,[\Delta_1^2\Delta_2^2]^{1/2}$ for both the 3D and 2D cases, and we use the same coherence function from Heneka et al. (2016) in the optimistic and pessimistic scenarios. These results show that 2D experiments can detect much of the 3D fluctuation power at $k < 0.2$ Mpc^{-1}, and motivate more complete and self-consistent simulations of the 21 cm and Lyα fields throughout the EOR over a range of optimistic and pessimistic scenarios.

5.4. Limits on the 21 cm–Lyα Cross Spectrum and an Experimental Design Study

Having characterized the residual sky power spectrum at 185 MHz and 850 nm after applying the best foreground masking and subtraction permitted by our data sets, we now search for a correlation between these radio and foreground residuals. To account for the nonuniform uv sampling of the MWA image, as well as for the image space masking of the ATLAS image, we again use the optimal quadratic estimator. In Appendix C we show that with the approximation that the correlation between the two images is small, the proper extension of the optimal quadratic estimator from the autospectrum case to the cross-spectrum case is given by

$$q_\alpha \approx (\boldsymbol{x}_{21} - \langle\boldsymbol{x}_{21}\rangle)^t \boldsymbol{C}_{21}^{-1} \boldsymbol{C}_{,\alpha} \boldsymbol{C}_{\mathrm{IR}}^{-1} (\boldsymbol{x}_{\mathrm{IR}} - \langle\boldsymbol{x}_{\mathrm{IR}}\rangle), \tag{22}$$

$$\boldsymbol{F}_{\alpha\beta} \approx \mathrm{tr}(\boldsymbol{C}_{21}^{-1} \boldsymbol{C}_{,\alpha} \boldsymbol{C}_{\mathrm{IR}}^{-1} \boldsymbol{C}_{,\beta}), \tag{23}$$

where $\boldsymbol{C}_{,\alpha}$ is the same matrix used in Section 5.2, and $\boldsymbol{x}_{21}$ and $\boldsymbol{x}_{\mathrm{IR}}$ are column vectors of 21 cm and IR pixel values. Observe that in this approximation we model the radio covariance separately from the IR covariance. This separation is also ideal for our current study, where we seek to *characterize* any cross spectrum between the two, rather than *model* a cross spectrum and *downweight* by it.

As before, we use normalization $\boldsymbol{M} = \boldsymbol{F}^{-1/2}$ in $\boldsymbol{p} = \boldsymbol{M}\boldsymbol{q}$ and perform a final normalization using the peaks of the window functions $\boldsymbol{W} = \boldsymbol{M}\boldsymbol{F}$. We use the same IR covariance matrix given in Equation (21) and model the radio covariance matrix as solely due to sample variance. We construct this covariance using the first term on the right-hand side of Equation (21), taking the power spectrum of the residual MWA broadband images as our guess. We emphasize that in both cases random noise is subdominant to the foreground residuals, so we do not include photon shot noise or thermal noise in these covariances.

We apply this formalism to each of the four $4°$ deep ATLAS fields shown in Figure 1, pairing each IR image with the overlapping region of the MWA image cropped from the naturally weighted image. We crop the image space synthesized beam over the same field of view and then use it to apply uniform weighting to the cropped MWA image using Equation (12). In Figure 14 we plot the resulting cross
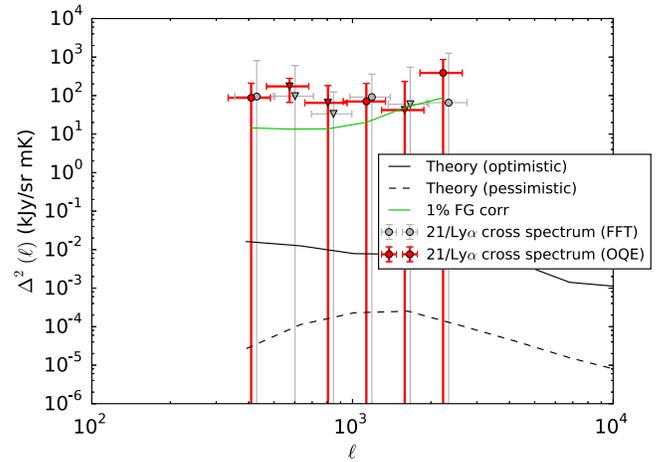

Figure 14. Measured cross spectrum between 185 MHz and 850 nm images after our best foreground subtraction and masking. Our optimal quadratic estimator results are shown as red markers, while our FFT-based cross-spectrum results are shown as gray markers. Most are consistent with zero, and roughly half are positive (circles) and negative (triangles), as expected if there is no measured correlation. Error bars show 2σ uncertainties, and we take the tops of these error bars as upper limits on the cross spectrum of residual foregrounds of 21 cm–Lyα emission from the EOR at $z \sim 7$. We achieve an upper limit of $\Delta^2 < 181$ (kJy sr^{-1} mK) (95%) at $\ell \sim 800$. We also plot the radio–infrared foreground cross spectrum that a 1% flux correlation would produce (green line) and find that it is below the sensitivity of our MWA–ATLAS experiment, which is limited by the sample variance noise of uncorrelated foregrounds.

spectrum (red markers) averaged over all four fields with 2σ error bars. Most are consistent with zero, and roughly half the estimated bandpowers are negative (triangles) and positive (circles), as expected if there is no measured correlation. Our bins are evenly spaced in $\log\ell$, and it is beneficial to overlap them slightly since our normalization matrix $M \sim F^{-1/2}$ ensures that bin errors are always uncorrelated. We take the tops of the error bars as 95% upper limits on the cross spectrum of residual foregrounds of 21 cm and Lyα emission from the EOR, achieving a tightest upper limit of $\Delta^2 < 181$ (kJy sr^{-1} mK) (95%) at $\ell \sim 800$. For comparison, we plot optimistic (black solid) and pessimistic (black dashed) model cross spectra as discussed in the previous section.

We check these quadratic estimator results (red points) against FFT-based power spectrum results on the same images (gray points). In fact, the two spectra are quite similar. In most bins the quadratic estimator limits are lower but are not significantly closer to the fiducial 1% radio/IR foreground correlation. To understand why the FFT results are only slightly worse, consider that after our IR foreground masking only $\sim$5% of coarse pixels end up masked. With different masking parameters more or fewer coarse pixels would get masked, in some cases necessitating the quadratic estimator to deal with large masked regions. But with the best masking parameters we settled on, the improvement is modest.

Also in Figure 14 we plot the radio–infrared foreground cross spectrum that a 1% flux correlation would produce (green line), as motivated by our results in Sections 4.3 and 4.2. Such a slight geometric flux correlation is thus still predicted to be slightly below the sensitivity of our MWA–ATLAS experiment, but is it significant for other classes of experiments? To study this, we predict for a range of current and future radio–IR surveys both the foreground sample variance noise and the foreground cross spectrum for a slight geometric correlation.

Radio Survey	Resolution	S_{max}
MWA	6′	150 mJy
MWA (GMRT Catalog)	6′	25 mJy
HERA (GMRT Catalog)	12′	25 mJy
GMRT	40″	25 mJy
LOFAR	20″	3 mJy
SKA	10″	0.1 mJy
IR Survey	Field of View	f_{IR} [a]
ATLAS	8°	1.0
WideATLAS	40°	1.0
WFIRST Mosaic	4°	0.01
DES Mosaic	40°	0.01

Note.
[a] Fraction of IR foreground power remaining relative to our ATLAS analysis.

We describe the radio and IR surveys we study here and summarize them side by side in Table 1.

Radio surveys

1. *MWA.* We use the 6′ resolution permitted by our MWA analysis, corresponding to a maximum baseline length of 700 m and a survey depth at 50% completeness of 150 mJy (see Section 4.1).
2. *HERA.* Primarily designed for 3D power spectrum measurements that are severely thermal noise limited, HERA is a compact redundant array, and even the outriggers only improve the resolution to 12′ (DeBoer et al. 2017). We assume that the GMRT catalog is used for source subtraction (see below).
3. *GMRT.* We assume a maximum baseline length of 10 km, within which the *uv* plane is filled relatively uniformly; this corresponds to a 40″ resolution. Intema et al. (2017) demonstrate a catalog depth of 25 mJy at 50% completeness.
4. *LOFAR.* Similarly, we use a maximum baseline length of 20 km, to maintain a relatively filled *uv* plane, corresponding to a 20″ resolution. Yatawatta et al. (2013) report a catalog depth of 3 mJy.
5. *SKA.* Prandoni & Seymour (2015) report a nominal SKA-Low confusion-limited survey depth of 0.1 mJy at 10″ resolution.

IR surveys

1. *ATLAS.* Our four ATLAS stacking fields give a total field of view of 8°, though we are planning a future wider survey with a 40° field of view.
2. *WFIRST mosaic.* Mitchell-Wynne et al. (2015) have demonstrated high-fidelity mosaicking using the technique of Fixsen et al. (2000), producing 10′ wide mosaics from the ∼2′ *HST* field of view. Along these lines, we assume that a 4° *WFIRST* field can be used through some combination of mosaicking (making larger IR images) and tiling (averaging the cross spectrum over many small fields). The instantaneous field of view of *WFIRST* is roughly 100 times that of *HST* with comparable resolution. Motivated by the results of Mitchell-Wynne et al. (2015), we assume a 100× improvement in IR foreground removal over our ATLAS analysis.

3. *DES mosaic.* Along these lines, we assume that a similar analysis can be conducted using the Dark Energy Survey (Dark Energy Survey Collaboration et al. 2016) over a field of view that is 10× larger on a side, due to the instrument's correspondingly larger field of view. We assume the same improvement in foreground removal as in the *HST* mosaic discussed previously.

Neglecting the small geometric foreground correlation, the foreground sample variance noise due to uncorrelated foregrounds is

$$\Delta_{noise}^2(\ell) = \sqrt{\frac{\Delta_{IR}^2(\ell)\,\Delta_{rad}^2(\ell)}{2N_\ell}}$$

$$= \sqrt{\frac{\Delta_{ATLAS}^2(\ell)\,\Delta_{MWA}^2(\ell)\,f_{rad}\,f_{IR}}{2N_\ell}}$$

$$\times \left(\frac{\theta_{FOV,ATLAS}}{\theta_{FOV}}\right)\left(\frac{d\theta}{d\theta_{MWA}}\right), \qquad (24)$$

where f_{IR} and f_{rad} are the fractions of IR and radio foreground power remaining relative to those in our ATLAS and MWA analyses. Using the empirical formula for radio source counts given in Di Matteo et al. (2002), f_{rad} scales as $f_{rad} \propto S_{max}^{1.25}$, where S_{max} is the flux limit of the subtraction. Scaling this relative to our MWA analysis gives $f_{rad} = (S_{max}/S_{max,MWA})^{1.25}$.

Similarly, the foreground cross spectrum for a correlation c is given by

$$\Delta_{FG}^2(\ell) = c\sqrt{\Delta_{IR}^2(\ell)\,\Delta_{rad}^2(\ell)}$$

$$= c\sqrt{\Delta_{ATLAS}^2(\ell)\,\Delta_{MWA}^2(\ell)\,f_{rad}f_{IR}}. \qquad (25)$$

In Figure 15 we plot the foreground sample variance noise versus foreground cross spectrum for a range of current and future radio and infrared survey pairs. We calculate these for the $\ell = 400$ bin and assume a 1% geometric foreground correlation. The diagonal black line separates the regions within which each of these effects dominates. We observe that our MWA/ATLAS analysis (black circle) is the only experiment limited by sample variance noise of uncorrelated foregrounds. An improved analysis using the GMRT catalog for radio source subtraction from MWA images and widening the ATLAS survey to a wide 40° (black star) falls barely in the foreground correlation-limited regime. A similar analysis using HERA instead of the MWA (yellow star) produces similar results, as HERA is not optimized for the wide-field, high-resolution radio surveys that broadband correlation experiments require.

LOFAR reaches within a factor of a few in power of the optimistic EOR cross-spectrum model when correlated against the *HST* (red triangle) or DES (red square) mosaics, and the SKA coupled with the same IR surveys achieves a near detection of the optimistic model (white square and white triangle). None of the experiments reach the pessimistic model, but beginning to probe optimistic models already starts to exclude late reionization scenarios, and performing the same correlation experiment at higher redshifts can likely give more leverage on earlier ones.

Interestingly, we note that virtually all the intermediate experiments between this work and the ultimate detection will
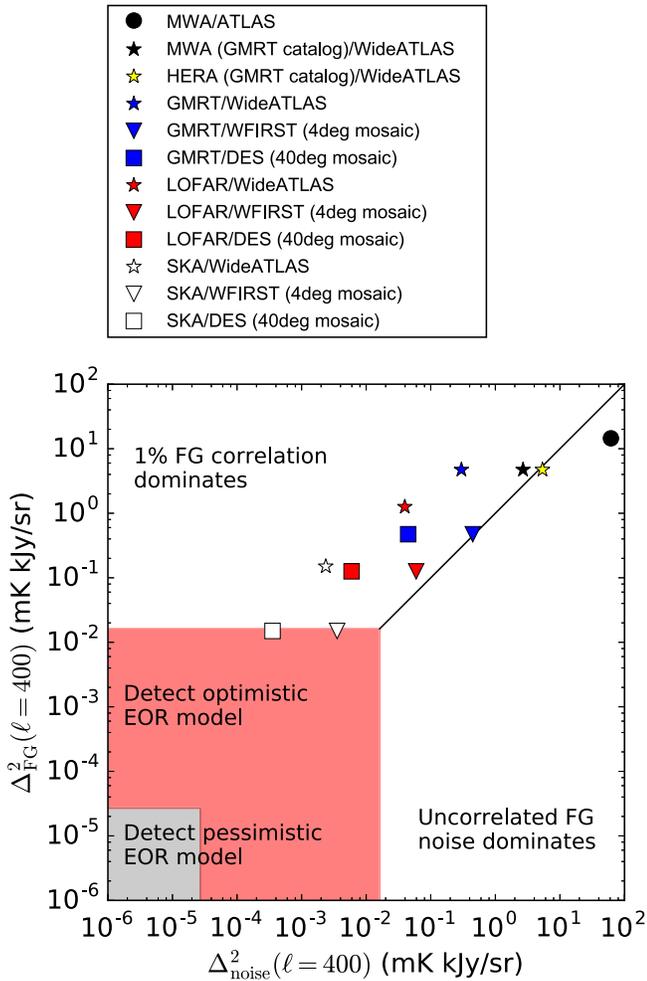
Figure 15. Foreground cross spectrum due to percent-level geometric correlations (y-axis) due to sample variance noise due to the uncorrelated component of foregrounds (x-axis) for a number of possible radio–IR intensity mapping survey pairs. This work (black circle) is the only one limited by sample variance noise; even a nominal improvement using the deeper GMRT catalog becomes foreground correlation limited. Deeper radio (e.g., LOFAR and SKA) and IR (e.g., *HST* and DES mosaics) are needed to give the foreground removal required to suppress the geometric foreground correlation below the optimistic expected EOR signal.

be limited by the geometric foreground correlations. Motivated by our results in Section 4.2, we have assumed here that the coherence remains at the percent level irrespective of mask depth, though more work is needed to model this effect more completely and self-consistently. In any case, this geometric foreground correlation appears to be strictly positive, while the predicted correlation of 21 cm–Lyα from the EOR is negative, giving us a test of whether foregrounds remain in the data.

6. Discussion

Radio and infrared surveys are on the cusp of direct observations of the two sides of the EOR. Deep infrared surveys are beginning to probe the bright end of the ionizing galaxy population itself, while low-frequency radio surveys are constraining the 21 cm brightness of the neutral IGM. New radio and infrared instruments are all but assured to yield the first direct constraints on the astrophysics of the EOR, and combining these measurements will be crucial to confirm these first detections and extend their science reach. In particular,

measurements of the predicted anticorrelation between redshifted radio and Lyα emission can shed light on properties and statistics of the ionizing sources that 21 cm measurements are not directly sensitive to, while simultaneously yielding redshift information on the near-infrared background for the first time.

3D correlation measurements are the ultimate goal, but making infrared maps with sufficient spectral resolution over large enough fields to match the comoving volumes probed by low-frequency 21 cm experiments is challenging and expensive. In contrast, 2D (i.e., broadband) infrared maps can be measured by many existing ground- and space-based observatories and are a natural first step.

We have shown that foregrounds pose two significant challenges for 2D intensity mapping correlation experiments: (1) simple geometric effects result in percent-level correlations between radio and infrared fluxes, even if their luminosities are uncorrelated; and (2) the largely uncorrelated radio and infrared foreground residuals contribute a sample variance noise to the cross spectrum. The first challenge demands better foreground masking and subtraction, while the second requires measurements over large fields of view with many independent samples to average down uncorrelated radio and infrared power.

In the first part of this paper we searched for correlations between radio and infrared catalog fluxes (i.e., point sources). We began with the 185 MHz MWA catalog and the 4.5 μm *WISE* catalog; any foreground correlation between these bands would limit 21 cm–Hα correlation analyses at $z \sim 7$. We detect a few correlations at $>6\sigma$ significance after masking 4.5 μm sources down to 18.25 mag, which corresponds to the flux below which extragalactic sources dominate the catalog. We reproduce this observed 185 MHz–4.5 μm correlation in simulations, confirming that it is sourced by SF galaxies rather than AGNs. The narrow radio and infrared luminosity functions of the former make distance a much stronger effect than for AGNs in deriving fluxes from luminosities.

We then performed 850 nm observations of the MWA field using the ATLAS telescope and searched for correlations between 185 MHz and 850 nm catalog fluxes in MWA and ATLAS catalogs. Any correlation between these bands would limit 21 cm–Lyα analyses. We observed a marginal correlation at the $\sim 3\sigma$ level that grew slightly stronger toward our deepest 850 nm mask depth of 0.1 mJy. This is consistent with our finding from SDSS data that at 850 nm stars still dominate source counts above this flux. Deeper near-infrared studies are needed to study how these correlations depend on mask depth to quantify what level of foreground removal is required to mitigate them. For now, we assume that radio–infrared flux correlations are at the percent level for the purpose of weighing foreground flux correlations versus foreground sample variance noise in our experimental design study.

In the second part of this paper we analyzed the residual 185 MHz and 850 nm foregrounds in MWA and ATLAS images after subtracting radio sources and masking infrared sources. We computed power spectra of the band-averaged MWA cubes and found that they agree with the $k_{\parallel} = 0$ bin of the 3D power spectra of Beardsley et al. (2016) after conversion from 3D to 2D units. Despite the fact that the MWA image is not thermal noise limited, we find that foreground subtraction improves significantly (by a factor of 2 in power) when the same number of observations are spread over 2 weeks instead of clustered within a single night, consistent with ionosphere-induced calibration errors. If the

ionosphere is indeed the cause, then the implication is that observing time should be rotated between multiple fields on a nightly cadence.

We then used the optimal quadratic estimator of Tegmark (1997) to compute the power spectrum of our 850 nm ATLAS images after masking sources at $2''$ resolution. These results agree well over $300 < \ell < 4000$ with the 1.1 μm power spectrum measured by Zemcov et al. (2014) using CIBER, a sounding rocket with $7''$ resolution and $\sim 1°$ field of view. Our larger field of view permits significantly improved precision at $\ell < 1000$, and while our slightly increased resolution *should* permit better masking, the lack of foreground reduction indicates that the power spectrum has hit a floor at these angular scales. Mitchell-Wynne et al. (2015) show that mild improvement is possible using the *HST*'s significantly improved $0''.1$ resolution, but argue that further improvement at these angular scales is stymied by Galactic dust. They demonstrate that significant improvement is possible at larger ℓ, though these scales are largely inaccessible to 21 cm observatories focusing on EOR power spectrum measurements (e.g., PAPER, the MWA, and HERA). Taking advantage of these infrared foreground reductions will require the higher-resolution images of LOFAR and SKA-LOW.

Turning to cross-spectrum measurements, we simulate the loss of signal in band averaging from 3D to 2D maps. We find that 2D experiments can recover much of the signal at $k < 0.2$ Mpc^{-1}, but that power on shorter length scales is suppressed by a factor of 10–10^2 in power due to averaging out of of $k_{\parallel} > 0$ modes.

Finally, we used our foreground-subtracted and masked MWA and ATLAS images to set the first limits on the cross spectrum of residual foregrounds of 21 cm and Lyα emission from $z \sim 7$. We adapted the optimal quadratic estimator to the cross-spectrum case to properly account for both nonuniform radio uv sampling and infrared image space masking. Our results are consistent with zero correlation, and our strictest upper limit on the residual foreground cross spectrum is $\Delta^2 < 181$ (kJy sr^{-1} mK) (95%) at $\ell \sim 800$.

A percent-level correlation between radio and infrared foreground fluxes, of the sort suggested by our catalog correlation analyses, remains below this upper limit, but will be crucial for future experiments. We weigh the impact of a percent-level radio–infrared foreground correlation against the sample variance noise due to their uncorrelated part for several possible future experiments. The simplest improvement we consider is increasing the ATLAS survey area by a factor of ~ 10 and subtracting radio fluxes below the MWA's confusion limit using the GMRT catalog. We predict that the sensitivity boost of these improvements should reveal the cross-spectrum floor due to the percent-level geometric flux correlations. Pushing down this floor requires foreground reduction using higher-resolution radio surveys such as LOFAR and SKA-LOW and infrared surveys such as *HST* and the Dark Energy Survey, though future work is needed to better understand these geometric flux correlations at very deep flux cuts.

We conclude that detection of the predicted EOR anticorrelation between redshifted 21 cm and Lyα emission in 2D (i.e., broadband) images is challenging, but within reach of future surveys.

In the near term, by probing correlation properties of radio/infrared foregrounds, these 2D experiments will be valuable complements to future 3D correlation analyses using infrared

cubes from the proposed SPHEREx and Cosmic Dawn Intensity Mapper telescopes. These 3D surveys exchange foreground challenges for sensitivity and cost challenges, but will eventually probe the short spatial scales inaccessible to 2D surveys. The decade of direct observation of the EOR is upon us, and correlation experiments will help us move from the era of detection to the era of astrophysics.

Appendix A
Power Spectrum of Photon Shot Noise

In Section 5.2 we measure the maximum airglow to be $I_{air} = 5 \times 10^3$ kJy sr^{-1}, and in this appendix we calculate the power spectrum of this photon shot noise. We must observe that the mean number of photons collected by a pixel during each observation is $\langle N_{ph} \rangle = I_{air} A t_{int} \Delta f d\theta^2 / hf$, where $A = (0.5 \text{ m})^2$ is the collecting area of ATLAS, $t_{int} = 30$ s, Δf and f are the frequency bandwidth and center frequency of I band, respectively, and $d\theta$ is the pixel size. The passband has $\Delta\lambda = 150$ nm and $\lambda = 800$ nm.

The shot noise contribution to the power spectrum is given by

$$C_{IR,shot}(\ell) = \left\langle \left| \sum_{m,n} I_{shot}(m,n) e^{-2\pi i(ma+nb)/N} \right|^2 \right\rangle \frac{d\theta^2}{N^2}, \quad (26)$$

where $I_{shot}(m,n) \equiv I(m,n) - \langle I(m,n) \rangle$ denotes the photon shot noise contribution to pixel (m,n), and N is the number of pixels on each side of the square image. Then using the fact that the shot noise is uncorrelated between different pixels, we find

$$C_{IR,shot}(\ell) = \sum_{m,n} \langle I_{shot}^2(m,n) \rangle \frac{d\theta^2}{N^2}. \quad (27)$$

Note that $I(m,n) = N_{ph}(m,n) hf / \Delta f A t_{int} d\theta^2$ and $\langle N_{ph}^2 \rangle = \langle N_{ph} \rangle$, so we have

$$C_{IR,shot}(\ell) = \langle N_{ph} \rangle \left(\frac{hf}{\Delta f A t_{int} d\theta^2} \right)^2 d\theta^2, \quad (28)$$

$$C_{IR,shot}(\ell) = \frac{I_{air} h\lambda}{\Delta\lambda A t_{int}}, \quad (29)$$

which gives $\Delta(\ell = 10^3) = 6 \times 10^{-2}$ kJy sr^{-1} for our deep fields with $t_{int} = 2.5$ minutes.

Appendix B
Relation between the Power Spectrum of Image Cubes and Broadband Images

We focus in this paper on the spherical power spectrum of broadband images, C_ℓ, instead of that of image cubes, $P(k)$, as 21 cm observations have focused on. Here we work out the approximate relation between the two over small fields of view (i.e., for large ℓ) to facilitate comparison with past 21 cm power-spectrum results. In particular, we calculate the scaling factor B relating the purely transverse modes of the power

spectrum $P(k_\perp, k_\parallel = 0)$ of an image cube $I(\theta_x, \theta_y, f)$ to the spherical power spectrum of a broadband image C_ℓ as

$$P(k_\perp, k_\parallel = 0) = B C_{\ell(k_\perp)}. \tag{30}$$

Using the Fourier transform convention discussed in Section 2, the left-hand side of the equation is given by

$$P(k_\perp, k_\parallel = 0) = \frac{1}{N_\perp^2 N_\parallel dV} \langle |\tilde{I}(k_x, k_y, k_\parallel = 0)|^2 \rangle, \tag{31}$$

where $N_\perp \equiv N_x = N_y$ is the number of pixels in each of the two transverse dimensions of the image cube and $N_\parallel$ is the number of pixels in the line-of-sight (i.e., frequency) dimension. The comoving pixel volume is $dV = (D_c d\theta)^2 (\Delta D_c / N_\parallel)$, where D_c is the line-of-sight comoving distance from the present day to the center of the cube and ΔD_c is the comoving line-of-sight thickness of the cube. Lastly, recall that $k_\perp$ is related to k_x and k_y as $k_\perp = \sqrt{k_x^2 + k_y^2}$.

Now substituting the definition of the Fourier transform, we find

$$P(k_\perp, k_\parallel = 0)$$
$$= \frac{1}{N_\perp^2 N_\parallel dV} \left\langle \left| dV \sum_{\theta_x, \theta_y, f} I(\theta_x, \theta_y, f) e^{i D_c (k_x \theta_x + k_y \theta_y)} \right|^2 \right\rangle. \tag{32}$$

Simplifying and writing this in terms of the broadband image $I_{\Delta f}(\theta_x, \theta_y) \equiv \frac{1}{N_\parallel} \sum_f I(\theta_x, \theta_y, f)$, we find

$$P(k_\perp, k_\parallel = 0)$$
$$= (D_c^2 \Delta D_c) \frac{d\theta^2}{N_\perp^2} \left\langle \left| \sum_{\theta_x, \theta_y} I_{\Delta f}(\theta_x, \theta_y) e^{i D_c (k_x \theta_x + k_y \theta_y)} \right|^2 \right\rangle. \tag{33}$$

Now denote $k_x = a \cdot dk$, $k_y = b \cdot dk$, $\theta_x = m \cdot d\theta$, and $\theta_y = n \cdot d\theta$, where $dk = 1 / N_\perp D_c d\theta$. Then

$$P(k_\perp(\ell(a, b)), k_\parallel = 0)$$
$$= (D_c^2 \Delta D_c) \frac{d\theta^2}{N_\perp^2} \left\langle \left| \sum_{m, n} I_{\Delta f}(m, n) e^{2\pi i (am + bn)/N_\perp} \right|^2 \right\rangle. \tag{34}$$

Comparing with Equations (5)–(7), we see that $B \equiv P(k_\perp, k_\parallel = 0) / C_{\ell(k_\perp)} = D_c^2 \Delta D_c$ and $\ell = D_c k_\perp$.

Appendix C
Extending the Optimal Quadratic Power Spectrum Estimator to the Cross-spectrum Case

The optimal quadratic estimator formalism presented in Section 5.2 was constructed to estimate the power spectrum of an image with arbitrary pixel sampling and noise properties. In this section we extend this formalism to achieve the same advantages in cross-spectrum measurements.

Consider measurements at two bands over the same set of pixels on the sky, x_1 and x_2, each a column vector with N elements. Let us combine these together into a single column vector containing all measurements as $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, whose covariance is given by

$$C \equiv \langle xx^\dagger \rangle - \langle x \rangle \langle x^\dagger \rangle = \begin{pmatrix} C_1 & C_{12} \\ C_{12}^\dagger & C_2 \end{pmatrix} \tag{35}$$

and

$$\frac{dC}{dp_{12}^\alpha} = \begin{pmatrix} 0 & C_{,\alpha} \\ C_{,\alpha}^\dagger & 0 \end{pmatrix}. \tag{36}$$

C_1 and C_2 depend only on the auto-power spectra of the different fields; only C_{12} depends on the cross spectrum. Said another way, $C_{,\alpha}$ is the same matrix used in Section 5.2, but used here in the off-diagonal parts of dC/dp_{12}^α so as to capture the cross products between the two fields.[15] And as before, the unnormalized estimator q_α of the power in band α is given by

$$q_\alpha = \frac{1}{2}(x - \langle x \rangle)^t C^{-1} \frac{dC}{dp_{12}^\alpha} C^{-1}(x - \langle x \rangle) \tag{37}$$

and the elements of the Fisher matrix are

$$F_{\alpha\beta} = \frac{1}{2} \mathrm{tr} \left(C^{-1} \frac{dC}{dp_{12}^\alpha} C^{-1} \frac{dC}{dp_{12}^\beta} \right). \tag{38}$$

In the case where the correlation between the two fields is expected to be weak and we are primarily interested in setting an upper limit, we can get a significant speedup by approximating $C_{12} \approx 0$ in our guess covariance. The expressions for q_α and $F_{\alpha\beta}$ then simplify to

$$q_\alpha \approx (x_1 - \langle x_1 \rangle)^t C_1^{-1} C_{,\alpha} C_2^{-1}(x_2 - \langle x_2 \rangle), \tag{39}$$

$$F_{\alpha\beta} \approx \mathrm{tr}(C_1^{-1} C_{,\alpha} C_2^{-1} C_{,\beta}). \tag{40}$$

ORCID iDs

Abraham R. Neben ● https://orcid.org/0000-0001-7776-7240
Brian Stalder ● https://orcid.org/0000-0003-0973-4900
John L. Tonry ● https://orcid.org/0000-0003-2858-9657

References

Ali, Z. S., Parsons, A. R., Zheng, H., et al. 2015, ApJ, 809, 61
Alvarez, M. A., Finlator, K., & Trenti, M. 2012, ApJL, 759, L38
Amorín, R., Fontana, A., Pérez-Montero, E., et al. 2017, NatAs, 1, 0052
Atek, H., Richard, J., Jauzac, M., et al. 2015, ApJ, 814, 69
Beardsley, A. P., Hazelton, B. J., Morales, M. F., et al. 2013, MNRAS, 429, L5
Beardsley, A. P., Hazelton, B. J., Sullivan, I. S., et al. 2016, ApJ, 833, 102
Beardsley, A. P., Morales, M. F., Lidz, A., Malloy, M., & Sutter, P. M. 2015, ApJ, 800, 128
Bernard, S. R., Carrasco, D., Trenti, M., et al. 2016, ApJ, 827, 76
Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393
Bertin, E., Mellier, Y., Radovich, M., et al. 2002, in ASP Conf. Proc. 281, Astronomical Data Analysis Software and Systems XI, ed. D. A. Bohlender, D. Durand, & T. H. Handley (San Francisco, CA: ASP), 228
Bouwens, R. 2016, in Understanding the Epoch of Cosmic Reionization: Challenges and Progress, ed. A. Mesinger (Basel: Springer), 111
Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2011, ApJ, 737, 90
Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2014, ApJ, 793, 115
Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2015, ApJ, 803, 34
Bouwens, R. J., Oesch, P. A., Illingworth, G. D., Ellis, R. S., & Stefanon, M. 2016a, arXiv:1610.00283

---

[15] One might object to our form of dC/dp_{12}^α, arguing that artificially limiting ourselves to cross products between the two fields is tantamount to throwing out a significant amount of the information contained in the data sets, and thus our estimator cannot be optimal. There are certainly situations in which this would be the case. If we had some theory of how each field was related to the matter density field of the universe, then both the auto-products and cross products contain similar information. However, we take an empirical approach where we assume we know nothing about either field and want only to know their correlation properties. Only the cross products contain that information.

Bouwens, R. J., Oesch, P. A., Labbé, I., et al. 2016b, ApJ, 830, 67
Bowler, R. A. A., Dunlop, J. S., McLure, R. J., et al. 2015, MNRAS, 452, 1817
Bowler, R. A. A., Dunlop, J. S., McLure, R. J., & McLeod, D. J. 2017, MNRAS, 466, 3612
Bowman, J. D., Cairns, I., Kaplan, D. L., et al. 2013, PASA, 30, e031
Bradley, L. D., Trenti, M., Oesch, P. A., et al. 2012, ApJ, 760, 108
Calvi, V., Trenti, M., Stiavelli, M., et al. 2016, ApJ, 817, 120
Carroll, P. A., Line, J., Morales, M. F., et al. 2016, MNRAS, 461, 4151
Chang, T.-C., Pen, U.-L., Bandura, K., & Peterson, J. B. 2010, Natur, 466, 463
Coe, D., Bradley, L., & Zitrin, A. 2015, ApJ, 800, 84
Comaschi, P., Yue, B., & Ferrara, A. 2016, MNRAS, 463, 3193
Cooray, A., Bock, J., Burgarella, D., et al. 2016, arXiv:1602.05178
Cooray, A., Smidt, J., de Bernardis, F., et al. 2012, Natur, 490, 514
Cutri, R. M., Wright, E. L., Conrow, T., et al. 2013, Explanatory Supplement to the AllWISE Data Release Products, Tech. Rep. (http://wise2.ipac.caltech.edu/docs/release/allsky/expsup/index.html)
Dark Energy Survey Collaboration, Abbott, T., Abdalla, F. B., et al. 2016, MNRAS, 460, 1270
de Jong, T., Klein, U., Wielebinski, R., & Wunderlich, E. 1985, A&A, 147, L6
DeBoer, D. R., Parsons, A. R., Aguirre, J. E., et al. 2017, PASP, 129, 045001
Dewdney, P., & Lazio, T. 2008, URSI Radio Science Bulletin, 81, 4
Dewdney, P. E., Hall, P. J., Schilizzi, R. T., & Lazio, T. J. L. 2009, IEEEP, 97, 1482
Di Matteo, T., Perna, R., Abel, T., & Rees, M. J. 2002, ApJ, 564, 576
Dillon, J. S., Liu, A., & Tegmark, M. 2013, PhRvD, 87, 043005
Dillon, J. S., Liu, A., Williams, C. L., et al. 2014, PhRvD, 89, 023002
Dillon, J. S., Neben, A. R., Hewitt, J. N., et al. 2015, PhRvD, 91, 123011
Dodelson, S., Narayanan, V. K., Tegmark, M., et al. 2002, ApJ, 572, 140
Doré, O., Bock, J., Ashby, M., et al. 2014, arXiv:1412.4872
Doré, O., Werner, M. W., Ashby, M., et al. 2016, arXiv:1606.07039
Dunlop, J. S., Rogers, A. B., McLure, R. J., et al. 2013, MNRAS, 432, 3520
Eisenstein, D. J., Weinberg, D. H., Agol, E., et al. 2011, AJ, 142, 72
Ellis, R. S., McLure, R. J., Dunlop, J. S., et al. 2013, ApJL, 763, L7
Ewall-Wice, A., Bradley, R., Deboer, D., et al. 2016, ApJ, 831, 196
Feng, C., Cooray, A., & Keating, B. 2017, arXiv:1701.07005
Fernandez, E. R., Zaroubi, S., Iliev, I. T., Mellema, G., & Jelić, V. 2014, MNRAS, 440, 298
Finkelstein, S. L. 2016, PASA, 33, e037
Fixsen, D. J., Moseley, S. H., & Arendt, R. G. 2000, ApJS, 128, 651
Fu, H., Yan, L., Scoville, N. Z., et al. 2010, ApJ, 722, 653
Furlanetto, S., Oh, S., & Briggs, F. 2006, PhR, 433, 181
Furlanetto, S. R., & Lidz, A. 2007, ApJ, 660, 1030
Gardner, J. P., Mather, J. C., Clampin, M., et al. 2006, SSRv, 123, 485
Gong, Y., Cooray, A., Silva, M. B., et al. 2017, ApJ, 835, 273
Gong, Y., Silva, M., Cooray, A., & Santos, M. G. 2014, ApJ, 785, 72
Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, ApJS, 197, 35
Hall, P. J. (ed.) 2005, The Square Kilometre Array: An Engineering Perspective (Dordrecht: Springer)
Helgason, K., Ricotti, M., Kashlinsky, A., & Bromm, V. 2016, MNRAS, 455, 282
Helou, G., Soifer, B. T., & Rowan-Robinson, M. 1985, ApJL, 298, L7
Henden, A. A., Templeton, M., Terrell, D., et al. 2016, yCat, 2336, 0
Heneka, C., Cooray, A., & Feng, C. 2016, arXiv:1611.09682
High, F. W., Stubbs, C. W., Stalder, B., Gilmore, D. K., & Tonry, J. L. 2010, PASP, 122, 722
Huang, K.-H., Bradač, M., Lemaux, B. C., et al. 2016, ApJ, 817, 11
Illingworth, G. D., Magee, D., Oesch, P. A., et al. 2013, ApJS, 209, 6
Intema, H. T., Jagannathan, P., Mooley, K. P., & Frail, D. A. 2017, A&A, 598, A78
Jacobs, D. C., Pober, J. C., Parsons, A. R., et al. 2015, ApJ, 801, 51
Kashlinsky, A., Arendt, R. G., Ashby, M. L. N., et al. 2012, ApJ, 753, 63
Kashlinsky, A., Arendt, R. G., Mather, J., & Moseley, S. H. 2005, Natur, 438, 45
Kashlinsky, A., Arendt, R. G., Mather, J., & Moseley, S. H. 2007, ApJL, 657, L131

Lidz, A., Zahn, O., Furlanetto, S. R., et al. 2009, ApJ, 690, 252
Liu, A., & Tegmark, M. 2011, PhRvD, 83, 103006
Livermore, R. C., Finkelstein, S. L., & Lotz, J. M. 2017, ApJ, 835, 113
Mao, X.-C. 2014, ApJ, 790, 148
Masui, K. W., Switzer, E. R., Banavar, N., et al. 2013, ApJL, 763, L20
Mauch, T., & Sadler, E. M. 2007, MNRAS, 375, 931
McLeod, D. J., McLure, R. J., Dunlop, J. S., et al. 2015, MNRAS, 450, 3032
McLeod, D. J., McLure, R. J., & Dunlop, J. S. 2016, MNRAS, 459, 3812
Mesinger, A. (ed.) 2016, Understanding the Epoch of Cosmic Reionization (Basel: Springer)
Mesinger, A., Furlanetto, S., & Cen, R. 2011, MNRAS, 411, 955
Mingo, B., Watson, M. G., Rosen, S. R., et al. 2016, MNRAS, 462, 2631
Mitchell-Wynne, K., Cooray, A., Gong, Y., et al. 2015, NatCo, 6, 7945
Morales, M. F., & Wyithe, J. S. B. 2010, ARA&A, 48, 127
Neben, A. R., Bradley, R. F., Hewitt, J. N., et al. 2016, ApJ, 826, 199
Nikutta, R., Hunt-Walker, N., Nenkova, M., Ivezić, Ž., & Elitzur, M. 2014, MNRAS, 442, 3361
Oesch, P. A., Bouwens, R. J., Illingworth, G. D., et al. 2013, ApJ, 773, 75
Offringa, A. R., Wayth, R. B., Hurley-Walker, N., et al. 2015, PASP, 32, e008
Park, J., Kim, H.-S., Wyithe, J. S. B., & Lacey, C. G. 2014, MNRAS, 438, 2474
Parsons, A. R., Liu, A., Aguirre, J. E., et al. 2014, ApJ, 788, 106
Patra, N., Parsons, A. R., DeBoer, D. R., et al. 2017, arXiv:1701.03209
Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, A&A, 594, A13
Pober, J. C., Ali, Z. S., Parsons, A. R., et al. 2015, ApJ, 809, 62
Pober, J. C., Liu, A., Dillon, J. S., et al. 2014, ApJ, 782, 66
Prandoni, I., & Seymour, N. 2015, in Advancing Astrophysics with the Square Kilometre Array, The Continuum Universe (Trieste: SISSA), 67
Prescott, M., Mauch, T., Jarvis, M. J., et al. 2016, MNRAS, 457, 730
Pritchard, J. R., & Loeb, A. 2012, RPPh, 75, 086901
Pullen, A. R., Doré, O., & Bock, J. 2014, ApJ, 786, 111
Repp, A., Ebeling, H., & Richard, J. 2016, MNRAS, 457, 1399
Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2014, Proc. SPIE, 9143, 914320
Roberts-Borsani, G. W., Bouwens, R. J., Oesch, P. A., et al. 2016, ApJ, 823, 143
Robertson, B. E., Furlanetto, S. R., Schneider, E., et al. 2013, ApJ, 768, 71
Schmidt, K. B., Treu, T., Trenti, M., et al. 2014, ApJ, 786, 57
SDSS Collaboration, Albareti, F. D., Allende Prieto, C., et al. 2016, arXiv:1608.02013
Seymour, N., Stern, D., De Breuck, C., et al. 2007, ApJS, 171, 353
Silva, M. B., Santos, M. G., Gong, Y., Cooray, A., & Bock, J. 2013, ApJ, 763, 132
Smit, R., Bouwens, R. J., Labbé, I., et al. 2014, ApJ, 784, 58
Spergel, D., Gehrels, N., Breckinridge, J., et al. 2013, arXiv:1305.5422
Sullivan, P. W., & Simcoe, R. A. 2012, PASP, 124, 1336
Tegmark, M. 1997, PhRvD, 55, 5895
Tegmark, M., Dodelson, S., Eisenstein, D. J., et al. 2002, ApJ, 571, 191
Tegmark, M., & Zaldarriaga, M. 2002, PhRvD, 66, 103508
Tingay, S. J., Goeke, R., Bowman, J. D., et al. 2013, PASA, 30, e007
Tonry, J. L. 2011, PASP, 123, 58
Tonry, J. L., Stubbs, C. W., Lykke, K. R., et al. 2012, ApJ, 750, 99
Trenti, M., Bradley, L. D., Stiavelli, M., et al. 2011, ApJL, 727, L39
Vaate, J., de Lera Acedo, E., Virone, G., et al. 2011, in 2011 XXXth URSI General Assembly and Scientific Symp. (Piscataway, NJ: IEEE), 3340
van Haarlem, M. P., Wise, M. W., Gunst, A. W., et al. 2013, A&A, 556, A2
Wilkins, S. M., Bouwens, R. J., Oesch, P. A., et al. 2016, MNRAS, 455, 659
Willott, C. J., Rawlings, S., Jarvis, M. J., & Blundell, K. M. 2003, MNRAS, 339, 173
Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, AJ, 140, 1868
Wyithe, J. S. B., & Morales, M. F. 2007, MNRAS, 379, 1647
Xu, C., Lisenfeld, U., & Voelk, H. J. 1994, A&A, 285, 19
Yatawatta, S., de Bruyn, A. G., Brentjens, M. A., et al. 2013, A&A, 550, A136
Yue, B., Ferrara, A., & Salvaterra, R. 2016, arXiv:1601.02514
Yun, M. S., Reddy, N. A., & Condon, J. J. 2001, ApJ, 554, 803
Zemcov, M., Smidt, J., Arai, T., et al. 2014, Sci, 346, 732