

FIRST RESULTS FROM COPSS: THE CO POWER SPECTRUM SURVEY

GARRETT K. KEATING¹, GEOFFREY C. BOWER², DANIEL P. MARRONE³, DAVID R. DEBOER¹, CARL HEILES¹, TZU-CHING CHANG⁴, JOHN E. CARLSTROM⁵, CHRISTOPHER H. GREER³, DAVID HAWKINS⁶, JAMES W. LAMB⁶, ERIK LEITCH^{5,6}, AMBER D. MILLER^{7,8}, STEPHEN MUCHOJEJ^{6,9}, AND DAVID P. WOODY⁶

¹ Radio Astronomy Laboratory, 501 Campbell Hall, University of California, Berkeley, CA 94720, USA; karto@astro.berkeley.edu

² Academia Sinica Institute of Astronomy and Astrophysics, 645 N. A'ohoku Place, Hilo, HI 96720, USA; gbower@asiaa.sinica.edu.tw

³ Steward Observatory, University of Arizona, 933 North Cherry Avenue, Tucson, AZ 85721, USA

⁴ Academia Sinica Institute of Astronomy and Astrophysics, P.O. Box 23-141, Taipei 10617, Taiwan

⁵ Department of Astronomy & Astrophysics, Kavli Institute for Cosmological Physics, University of Chicago, Chicago, IL 60637, USA

⁶ California Institute of Technology, Owens Valley Radio Observatory, Big Pine, CA 93513, USA

⁷ Columbia Astrophysics Laboratory, Columbia University, New York, NY 10027, USA

⁸ Department of Physics, Columbia University, New York, NY 10027, USA

⁹ California Institute of Technology, Department of Astronomy, Pasadena, CA 91125, USA

Received 2015 August 25; accepted 2015 October 22; published 2015 November 30

ABSTRACT

We present constraints on the abundance of carbon monoxide in the early universe from the CO Power Spectrum Survey. We utilize a data set collected between 2005 and 2008 using the Sunyaev–Zel'dovich Array (SZA), which was previously used to measure arcminute-scale fluctuations of the cosmic microwave background. This data set features observations of 44 fields, covering an effective area of 1.7 square degrees, over a frequency range of 27–35 GHz. Using the technique of intensity mapping, we are able to probe the CO(1–0) transition, with sensitivity to spatial modes between $k = 0.5\text{--}2\text{ h Mpc}^{-1}$ over a range in redshift of $z = 2.3\text{--}3.3$, spanning a comoving volume of $3.6 \times 10^6\text{ h}^{-3}\text{ Mpc}^3$. We demonstrate our ability to mitigate foregrounds, and present estimates of the impact of continuum sources on our measurement. We constrain the CO power spectrum to $P_{\text{CO}} < 2.6 \times 10^4\text{ }\mu\text{K}^2\text{ (h}^{-1}\text{ Mpc)}^3$, or $\Delta_{\text{CO}}^2(k = 1\text{ h Mpc}^{-1}) < 1.3 \times 10^3\text{ }\mu\text{K}^2$, at 95% confidence. This limit resides near optimistic predictions for the CO power spectrum. Under the assumption that CO emission is proportional to halo mass during bursts of active star formation, this corresponds to a limit on the ratio of CO(1–0) luminosity to host halo mass of $A_{\text{CO}} < 1.2 \times 10^{-5}\text{ }L_{\odot}\text{ }M_{\odot}^{-1}$. Further assuming a Milky Way-like conversion factor between CO luminosity and molecular gas mass ($\alpha_{\text{CO}} = 4.3\text{ }M_{\odot}\text{ (K km s}^{-1}\text{ pc}^{-2})^{-1}$), we constrain the global density of molecular gas to $\rho_{z\sim 3}(M_{\text{H}_2}) \leq 2.8 \times 10^8\text{ }M_{\odot}\text{ Mpc}^{-3}$.

Key words: galaxies: evolution – galaxies: high-redshift – ISM: molecules – methods: statistical

1. INTRODUCTION

Molecular gas serves a vital role in star formation as the natal material from which stars form. Though the main constituent of this gas is molecular hydrogen, the H_2 molecule lacks a permanent dipole moment, making it a poor radiator of energy and hence difficult to observe. Traditionally, the CO molecule—the next most abundant molecule after H_2 —has been used as a tracer of molecular hydrogen (e.g., Wilson et al. 1970; Young & Scoville 1982; Young et al. 1995; Regan et al. 2001; Bolatto et al. 2013). Unlike H_2 , the CO molecule possesses a permanent dipole moment, with an excitation temperature of $T_{\text{ex}} \approx 5.5\text{ K}$ for the $J = 1 \rightarrow 0$ transition, making it ideal for probing the cold, dense gas of molecular clouds.

Within the local universe, the CO luminosity (L_{CO}) of galaxies—and, by extension, their molecular gas mass—shows strong correlation with far-infrared luminosity (L_{FIR}), $\text{H}\alpha$ emission, and $\text{Ly}\alpha$ emission; all are strong indicators of star formation within these galaxies (Downes et al. 1993; Solomon et al. 1997; Kennicutt 1998). This relationship between molecular gas and star formation rates (SFRs) is typically referred to as the Kennicutt–Schmidt (KS) relationship (Schmidt 1959; Kennicutt 1998), and demonstrates a deep connection between the abundance of molecular gas and the formation of stars.

It is unknown how the connection between CO, bulk molecular gas, and star formation evolves over cosmic time. Observations of distant star-forming galaxies suggest that L_{CO} –

SFR correlation persists up to $z \lesssim 2$ (Tacconi et al. 2013), implying that the correlation between CO abundance and the amount of molecular gas available for star formation has remained relatively unchanged in the several billion years following the peak of cosmic star formation (Hopkins & Beacom 2006). Prior to this epoch, however, early galaxies (with their short star formation histories) may not contain enough metals to form an appreciable amount of CO, or may possess too little dust to shield the CO from dissociation by UV starlight (Genzel et al. 2012; Bolatto et al. 2013). Some models of the conversion factor between CO luminosity and molecular gas mass (α_{CO}) predict a steep power-law relationship between α_{CO} and the gas metallicity of galaxies (e.g., Israel 1997). Should their predictions hold true, many high-redshift galaxies may lack significant CO emission, despite the presence of molecular gas (Wolfire et al. 2010; Muñoz & Furlanetto 2014). Other theoretical work (e.g., Obreschkow et al. 2009b; Glover & Mac Low 2011) suggests that CO is not so strongly affected by the lower metallicity and dust masses of early galaxies, offering a more optimistic outlook for CO as a tool for exploring molecular gas at high redshift.

The makeup of the molecular gas content of star-forming galaxies in the early universe is currently an active area of observational research. Recently, Decarli et al. (2014) and Walter et al. (2014) used the Plateau de Bure Interferometer (PdBI) to make a very deep (100 hr integration time) observation of a portion of the GOODS-N field (Dickinson

et al. 2003), and probed the CO luminosity function at $z \sim 3$ to a limit of $L_{\text{CO}}^f \gtrsim 10^{10} \text{ K km s}^{-1} \text{ pc}^2$ for the $J = 3 \rightarrow 2$ rotational transition of CO. For $z \sim 3$, this line luminosity limit corresponds to galaxy SFRs greater than $\sim 1.2 \times 10^2 M_{\odot} \text{ yr}^{-1}$ (Tacconi et al. 2013). Clearly, such studies are limited to massive galaxies that are rapidly forming stars and miss the lower luminosity and/or less massive systems that are expected to make up a large fraction of star-forming galaxies (Obreschkow et al. 2009b; Lagos et al. 2011; Bouwens et al. 2012; Smit et al. 2012; Sargent et al. 2014).

Exploration of the properties of more typical galaxies may be done through “intensity mapping,” where the signals from hundreds or thousands of galaxies—both bright and dim—are detected in aggregate as larger-scale fluctuations in the mean line intensity. The method of CO intensity mapping has been investigated in numerous recent theoretical studies (Righi et al. 2008; Visbal & Loeb 2010; Carilli 2011; Gong et al. 2011; Lidz et al. 2011; Visbal et al. 2011; Pullen et al. 2013; Li et al. 2015). These analyses have predicted the mean brightness temperature of CO at $z \sim 3$ to be of the order of $\langle T_{\text{CO}} \rangle \sim 1 \mu\text{K}$ —within reach of existing instruments with very deep integrations, provided that observational systematics can be controlled.

The Sunyaev–Zel’dovich Array (SZA), an eight-element interferometer of 3.5 m dishes with 1 cm receivers, is well matched to performing a CO intensity mapping experiment. The SZA is capable of observing the $J = 1 \rightarrow 0$ rotational transition of CO, which we will herein refer to as CO(1–0), at a redshift range of $z = 2.3\text{--}3.3$, with greatest sensitivity to comoving size scales of $0.5\text{--}2 h^{-1} \text{ Mpc}$. The data set for our analysis was previously used by Sharp et al. (2010, hereafter S10) as a measurement of the power spectrum of the cosmic microwave background (CMB) on arcminute angular scales.

In this paper, we discuss our search for measurable anisotropy in the three-dimensional distribution of molecular gas by characterizing the variance spectrum of the S10 data. This analysis resembles the two-dimensional power spectrum measurement of that work, extended to a third dimension using the frequency channels recorded in the data. Sensitive measurements may eventually image the intensity variations due to the large-scale structure of CO-emitting galaxies, but at the depth of these data we expect, at best, only a statistical detection of CO fluctuations via variance that exceeds that expected from the thermal noise of the data set.

This paper is structured as follows: Section 2 discusses the SZA instrument and the observational data used for this experiment. Section 3 discusses the software pipeline built for the analysis of the data. Section 4 presents the results of our analysis, and Section 5 discusses these results in the context of theoretical expectations. Conclusions are given in Section 6.

2. OBSERVATIONS

In this section, we present a brief description of the instrument and data set used in our analysis. A more thorough description of each can be found in Muchovej et al. (2007) and Sharp et al. (2010), respectively.

2.1. Instrument Description

The SZA is an eight-element interferometer designed for measurements of the Sunyaev–Zel’dovich (SZ) effect signature

of galaxy clusters on arcminute angular scales (Muchovej et al. 2007). At the time of data collection, the SZA was located at the Owens Valley Radio Observatory (OVRO); it was later incorporated into the nearby Combined Array for Research in Millimeter-wave Astronomy.

Each SZA antenna has a single-polarization (left-hand circular) 1 cm receiver, capable of observing between 27 and 35 GHz, which corresponds to $z \approx 2.3\text{--}3.3$ for CO(1–0). At 30 GHz, an SZA antenna has a typical system temperature $T_{\text{sys}} \approx 40 \text{ K}$, and a typical aperture efficiency of $\eta_{\text{eff}} \approx 0.6$. Each antenna is 3.5 m in diameter, which corresponds to a FWHM $\theta_{\text{B}} \approx 11'$ at 30 GHz and covers an effective solid angle of $\Omega_{\text{B}} = 0.03 \text{ deg}^2$. The SZA antennas were arranged in a compact group of six antennas (4.5–11.5 m spacings), with two outrigger antennas $\sim 50 \text{ m}$ away to separate the extended SZ signal from compact radio continuum sources. Figure 1 shows both the layout of the SZA as well as the uv coverage of a single field during a typical observation.

The SZA operates using a 2 bit digital XF correlator—with an efficiency of $\eta_{\text{corr}} = 0.87$ —providing a total of 28 cross-correlations and 8 autocorrelations. The correlator has a total bandwidth of 8 GHz split across 16 different windows of 500 MHz bandwidth. Within each window are 16 spectral channels, of width 31.25 MHz. One channel is discarded within each window due to edge effects, leaving 468.75 MHz of usable bandwidth per window.

2.2. Data Description

The S10 data were obtained between 2005 and 2008. The data consist of 44 telescope pointings, arranged as 11 groups of four fields. Each group is composed of four pointings at constant declination separated by 4 minutes in R.A., such that each field is observed over the same hour angle over a series of sequential 4 minute observations. The duration of the loop through the four fields, consisting of twenty 20 second integrations per field, plus several minutes on a gain calibrator is approximately 20 minutes. A bandpass calibrator was typically observed for 5 minutes at the beginning or end of the track. We refer to a contiguous track of data, typically 6 hr in length, as an “observing block.” Each group of four fields was observed for approximately 45 days, providing an hour a day of integration time, with an average total of 20 hr of integration time per field after taking into account data flagging. A listing of the position of the lead field for each group is provided in Table 1.

3. ANALYSIS

3.1. Pipeline Overview

This section provides a broad overview of the data processing and calibration software used in our analysis. The power spectrum analysis and null tests are described in Sections 3.2 and 3.3, respectively. All of the routines used here were developed within MATLAB.¹⁰

The calibration of the raw SZA data follows similar procedures to those described in Muchovej et al. (2007) and S10. The raw data are recorded as complex correlation amplitudes with associated time-tagged status information. They are converted to a physical power scale using system

¹⁰ Mathworks, Version 2013b, <http://www.mathworks.com/products/matlab/>

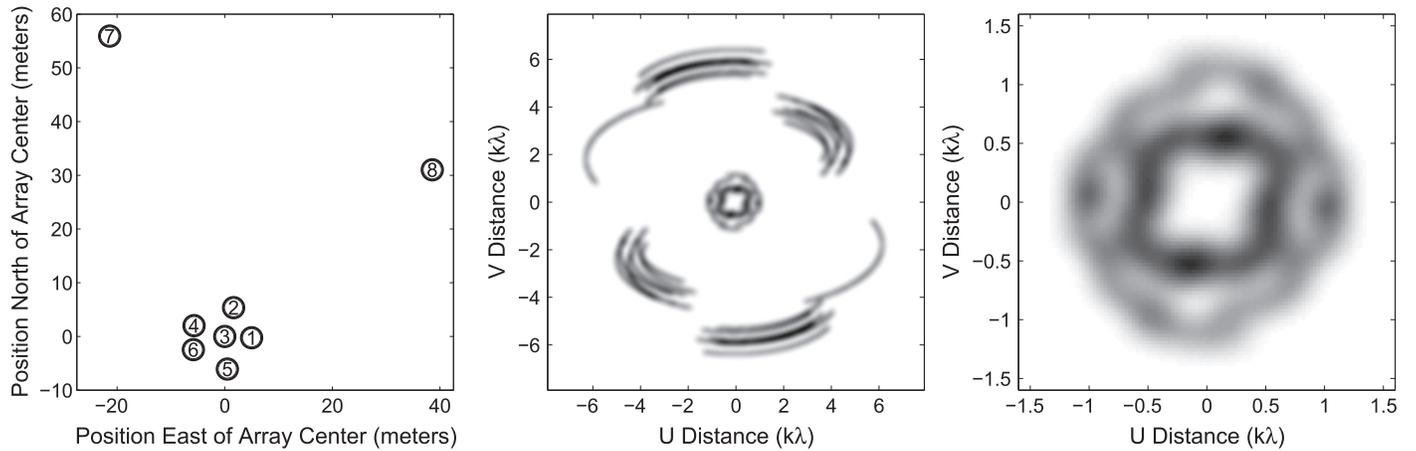


Figure 1. Left: SZA antenna positions. The antennas are drawn as 3.5 m diameter circles. Center: the uv coverage of the full array for one spectral window (30.2 GHz) for one of the fields in the S10 survey, with darker shades corresponding to areas with greater sensitivity. The baselines to the outrigger antennas, 7 and 8, are well separated from the baselines within the compact portion of the array. Right: the inner region of the uv plane.

Table 1

The Lead Field for Each of the 11 Groups Observed in the S10 Data Set

Field Name	R.A.	decl.	Gain Cal.
cmbA1	02 ^h 12 ^m 00 ^s .0	+33°00′00″	J0237+288
cmbAA1	21 ^h 24 ^m 38 ^s .7	+25°29′37″	J2139+143
cmbBB1	21 ^h 24 ^m 38 ^s .1	+25°59′24″	J2025+337
cmbCC1	02 ^h 11 ^m 31 ^s .3	+33°27′43″	J0237+288
cmbDD1	13 ^h 18 ^m 40 ^s .1	+35°01′42″	J1131+305
cmbEE1	14 ^h 18 ^m 39 ^s .2	+35°31′52″	J1331+305
cmbI1	02 ^h 12 ^m 00 ^s .0	+32°37′08″	J0237+288
cmbR1	02 ^h 12 ^m 15 ^s .6	+32°11′24″	J0237+288
cmbY1	02 ^h 12 ^m 00 ^s .0	+31°51′24″	J0237+288
cmbXX1	21 ^h 24 ^m 38 ^s .7	+24°59′37″	J2139+143
cmb07	02 ^h 07 ^m 37 ^s .0	+34°00′00″	J0237+288

temperature measurements that are made during every source–calibrator cycle. Absolute telescope and system efficiencies, derived from Mars via the brightness temperature model of Rudy et al. (1987), are applied. As these factors are identical to those used in S10, we expect they are accurate to 10%. The data are flagged to remove bad data, bandpass calibration is determined from a strong point source in each track, and relative gain calibration is determined from the gain calibrator observed on a 20 minute cycle. Some features of these steps are outlined below.

Flagging of data is done using three principal methods. First, data affected by known hardware issues (e.g., an antenna fails to point correctly) are marked as bad. Second, data are passed through various statistical tests and checked against theoretical estimates to see if they behave in a Gaussian fashion (data that exceed estimated noise thresholds are flagged as bad). Finally, data adjacent (i.e., belonging to preceding or subsequent frequency channels or integrations) to bad data are removed as well. Flagging of data is typically done with an iterative approach, identifying outliers in small groups of data (and removing them when appropriate) before reevaluating the data over larger groupings. This approach helps limit the impact that a few bad data points may have on an otherwise good data set, at the expense of additional processing time. We remove 28.7% of the data for known hardware problems (including shadowing of antennas), 4.8% of data for exceeding noise thresholds, and 6.5% of data for bad neighboring data. We also discard all data

from our shortest baseline (amounting to 3.6% of our data additionally removed) due to known systematic noise issues. In total, 43.6% of all data are flagged.

The SZA system produces very stable calibration across frequency and time. Gain amplitude and phase typically vary by 2% and $<10^\circ$ across a track, respectively. When large gain shifts are observed ($>10\%$ in amplitude or $>30^\circ$ in phase) between gain calibrator observations, data are marked as bad and excluded from later analysis. Phase solutions are linearly interpolated, while the very stable gain amplitudes are averaged over the track. Bandpass solutions are also typically stable to 1% between observing blocks.

One additional calibration step is performed to account for discrepancies between the expected and actual noise within our measurements. System temperatures are measured for each window, but RF/IF features and quantization effects can introduce variability in the system temperature on a channel-by-channel basis. To account for this effect, the variance within each channel of each baseline is calculated for all data within a single observing block (excluding calibrator data, and after subtraction of known sources in each field), and the “system equivalent flux density (SEFD) correction” is determined as the difference between the theoretical and measured noise. These differences are believed to be due to antenna-based effects (primarily due to standing waves in the receiver bandpass that introduce a spectrally varying signal level at the digitizers, resulting in varying quantization noise), hence an antenna-based correction factor is determined (using a χ^2 fit) for each frequency channel. The correction factors are determined once per track and are seen to be consistent to 1% between observing blocks. An example SEFD correction spectrum is shown in Figure 2.

Figure 3 presents images of all 44 fields, deconvolved via the CLEAN algorithm (Högbom 1974). No primary beam correction is applied to these images. The resolution of the images is ~ 0.5 and the median theoretical noise is 0.15 mJy. We find that the measured RMS residual noise in our images following deconvolution is consistent with theoretical estimates based on thermal noise, with the exception of cmbI17 (due to the presence of a bright point source at the field edge). We detect 24 sources above 10σ across the 44 fields, with a median flux of 2.5 mJy and a maximum (primary beam corrected) flux of 100 mJy. These sources are subsequently removed from our

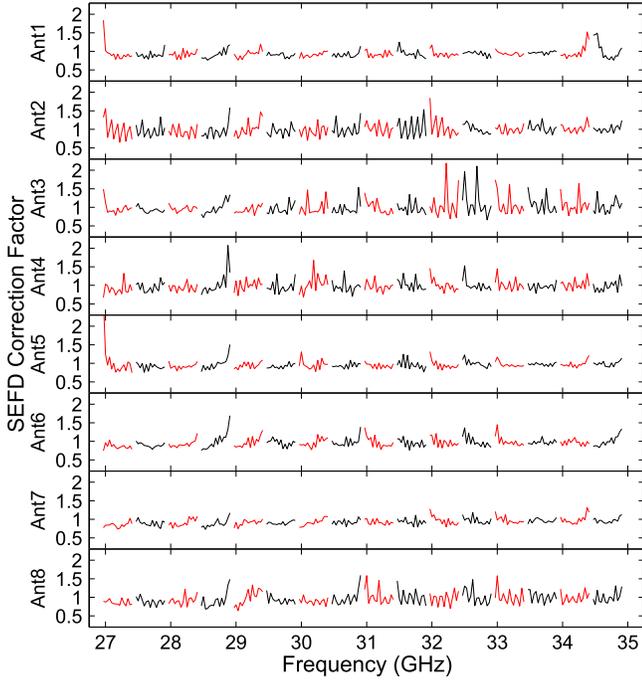


Figure 2. Example of the SEFD correction solution generated for BB20070906—a single observing block within our data set. SEFD corrections show good day-to-day agreement, similar to the bandpass calibration solutions.

data prior to our power spectrum analysis. Sources below the 10σ threshold are ignored because of their limited impact on our power spectrum measurement (see Section 3.2.2). These results are consistent with those reported by Sharp et al. (2010).

3.2. Power Spectrum Analysis

The goal of our power spectrum analysis is to transform the frequency-resolved interferometric visibilities into a measurement of the CO power spectrum. We provide an overview of the method here, reserving the full details for Appendix A.

The power spectrum measurement is performed in a three-dimensional spatial frequency domain that is the Fourier transform of the real-space universe. The interferometer naturally provides measurements of spatial frequencies in the two dimensions of the sky, while the third is constructed by Fourier transforming the frequency dimension of the data, which maps to the line-of-sight distance via the redshift of the CO line. The data are Fourier transformed one 500 MHz window at a time, providing redshift-segmented measurements of the Fourier-transformed intensity field, $\tilde{I}(u, v, \eta, z)$, with z representing the median redshift of the window. We refer to these as “delay-visibilitys.” The delay-visibilitys are then converted to temperature units, weighted by their variance estimates and gridded in the (u, v, η, z) space. Here u and v are the standard spatial frequency variables for interferometry, and η represents the Fourier transform of the frequency axis. The (u, v, η) domain is closely related to the \mathbf{k} space, differing only by the conversion factors X and Y , which convert between comoving physical size and angular distance or frequency with units Mpc rad^{-1} and Mpc Hz^{-1} , respectively (e.g., Parsons et al. 2012). With the application of the X and Y conversion factors, we find the Fourier dual of the specific intensity in the three-dimensional \mathbf{k} -vector space, $\tilde{I}(\mathbf{k})$. Within an individual window our coverage of \mathbf{k} space is similar in the η direction

and the u and v directions, considering only the compact portion of the array.

Formally, the intensity power spectrum is defined as

$$\begin{aligned} \Delta^2(k) &\equiv \frac{k^3}{2\pi^2} P(k), \\ &= \frac{k^3}{2\pi^2} \frac{X^4 Y^2}{V_z} \left(\frac{c^2}{2k_B V^2} \right)^2 \langle |\tilde{I}^2| \rangle_{\mathbf{k}=\mathbf{k}^2}. \end{aligned} \quad (1)$$

In Equation (1), $\Delta^2(k)$ (in units of μK^2 , often called the dimensionless power spectrum; e.g., Furlanetto et al. 2006) is the variance in brightness temperature at comoving spatial frequency k (in units of $h \text{ Mpc}^{-1}$, where $h = H_0/100 \text{ km s}^{-1} \text{ Mpc}^{-1}$ and H_0 is the current Hubble parameter) per $\ln(k)$. The power spectrum, $P(k)$, is the Fourier transform of the autocorrelation function of the CO intensity field, and has units $\text{Mpc}^3 \mu\text{K}^2$. The Boltzmann factor is represented by k_B , and the speed of light by c . The volume probed by the measurement at redshift z is $V_z = X^2 Y B_z \Omega_B / 2$, for a solid angle of the telescope primary beam, Ω_B , and a bandwidth, B_z , considered around this redshift. For a single pointing, the volume surveyed by a single correlator window (in this case, the central one) is a cylinder of diameter $12 \text{ Mpc}/h$ and length $45 \text{ Mpc}/h$.

As a dimensionless quantity, $\Delta^2(k)$ is sometimes favored over $P(k)$ (Dodelson 2003). However, $P(k)$ has the advantage that it maintains the same value for different values of k when in the cosmological shot-noise limit; for both the CO measurement and some foreground contaminants, our experiment resides firmly in the shot-noise regime. As such, values in both conventions are presented throughout this paper.

Values for the power spectrum are calculated using the following equation:

$$\begin{aligned} \mathcal{P}(\mathbf{k}, z) &= \frac{\sum_{\mathbf{k}'} \sigma_{\mathbf{k}}^{-2} \sigma_{\mathbf{k}'}^{-2} C(\mathbf{k} - \mathbf{k}') \left(|\tilde{I}^*(\mathbf{k}, z) \tilde{I}(\mathbf{k}', z)| \right)}{\sum_{\mathbf{k}'} \sigma_{\mathbf{k}}^{-2} \sigma_{\mathbf{k}'}^{-2} C^2(\mathbf{k} - \mathbf{k}')} \\ &\quad - \mathcal{A}_{\mathbf{k}} \\ P(k, z) &= \langle \mathcal{P}(\mathbf{k}, z) \rangle_{\mathbf{k}=\mathbf{k}^2}. \end{aligned} \quad (2)$$

In Equation (2), gridded data within a single redshift window are cross-multiplied against one another, weighted by their estimated thermal noise variance, $\sigma_{\mathbf{k}}^2$, and their normalized covariance, C , for the cross-multiplied data. The covariance is the analytically calculated correlation between adjacent points in the (u, v, η) space. For the uv plane, this amounts to a consideration of the overlap of the visibilities being multiplied, each of which samples an area in the uv plane described by the autocorrelation of the telescope illumination pattern. The correlation is similarly calculated between η channels. To eliminate noise bias from our measurement, the autocorrelations of individual delay-visibilitys are averaged over each grid cell ($\mathcal{A}_{\mathbf{k}}$) and removed from the power spectrum measurement. As a result, the autocorrelation measured in our power spectrum analysis may produce negative values when the result is noise-dominated. The power spectrum measurement is collapsed down to a single dimension, averaging over spheres in the \mathbf{k} space. A power spectrum is created for each redshift window, within each source field.

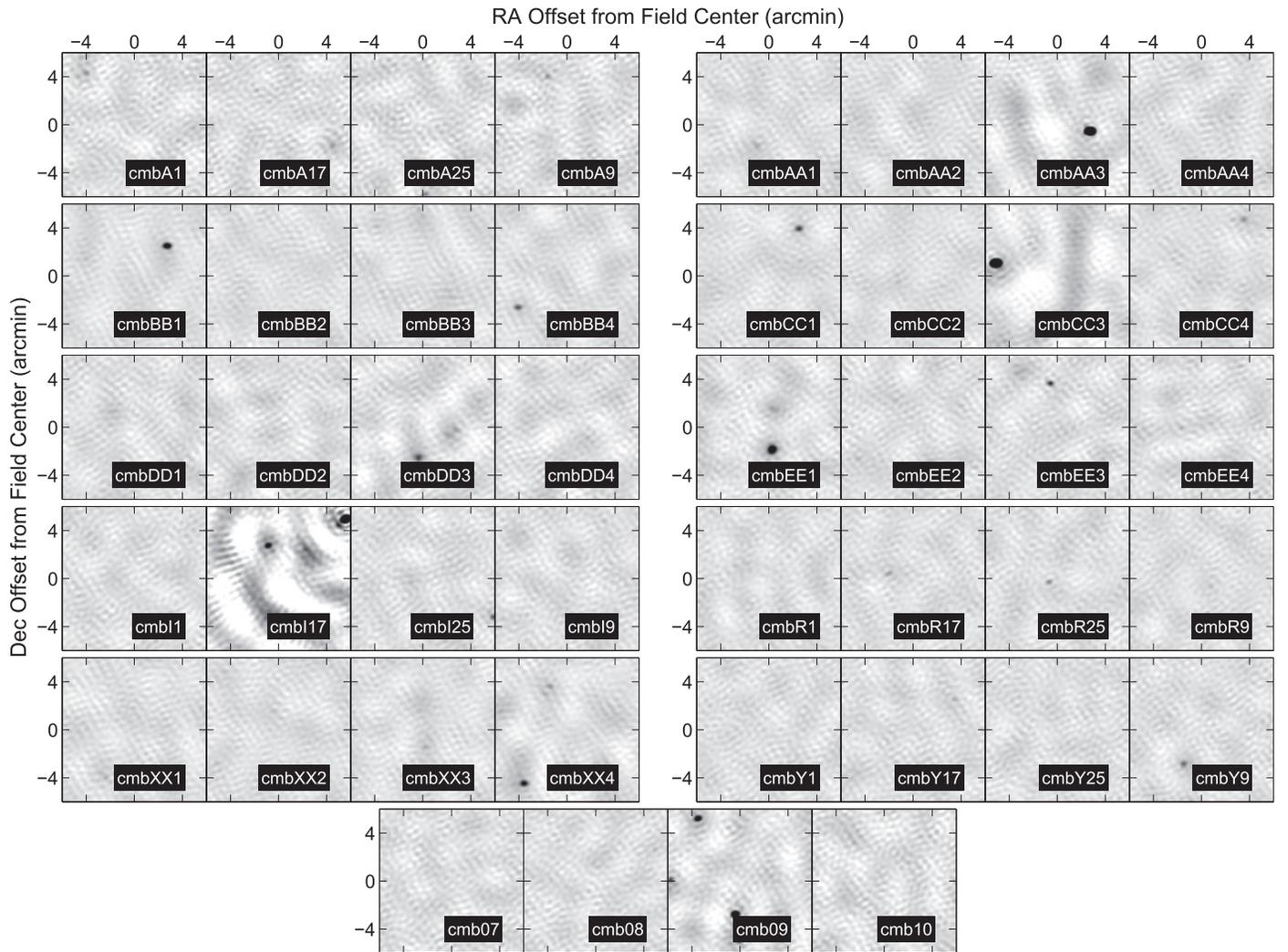


Figure 3. Thumbnail images of the deconvolved maps for all 44 fields included in the S10 survey. On average, we find there to be a rough average of one bright point source ($S \gtrsim 1$ mJy) per pointing. Most fields after deconvolution show noise-like residuals, though cmbI17 shows some artifacts due to the presence of a 100 mJy point source at the edge of the primary beam—arguably the worst-case scenario for contamination of our measurements (see Section 3.3).

To test the fidelity of the power spectrum pipeline, we generate a series of simulated observations of mock fields. Visibilities are generated based on a randomized set of spectral line sources. Gaussian noise is added to each visibility using noise estimates based on system temperature, instrument parameters, and SEFD corrections. Effects such as primary beam attenuation and mode-mixing (i.e., bandwidth smearing) are taken into account in these simulations, to better understand the limits of our analysis. To account for shot noise, we simulate data for 10^4 fields, which are then averaged together to provide a final power spectrum result. We find that the input power is recovered across our full k range to 1%, limited by small spectral and gridding corrections.

3.2.1. Ground Subtraction

The lead-trail design of the S10 observations allows the removal of any contributions to the power spectrum that are correlated with antenna position, such as ground contamination or antenna cross-talk. It was found in S10 that without ground subtraction there was weak evidence for a ground-correlated contamination of the CMB power spectrum measurement. Similar tests, described in the next section, suggest that

removal of such contamination is important to cleaning our data.

We create a model for the contamination using a variance-weighted average of the four fields in each group. The average is generated visibility by visibility, so that the first 20 s integrations on each field in the group are averaged together, preserving the individual baselines and frequency channels, as are the second integrations, and so on. This model is subtracted from the individual visibilities in each of the four fields, reducing the number of independent measurements in our experiment by 25%, which degrades the sensitivity by approximately 12%.

3.2.2. Point Source Contamination

Continuum point sources have the potential to significantly contaminate our measurements. Point sources provide of the order of 10^3 more power (in units of μK^2) than CMB anisotropies or Galactic foregrounds over the range of multipoles ($\ell \sim 3000$) that the SZA is sensitive to (S10; Bennett et al. 2013; Planck Collaboration et al. 2015). As such, we expect point sources to be the dominant astrophysical contaminant in our measurements. Fortunately, most sources

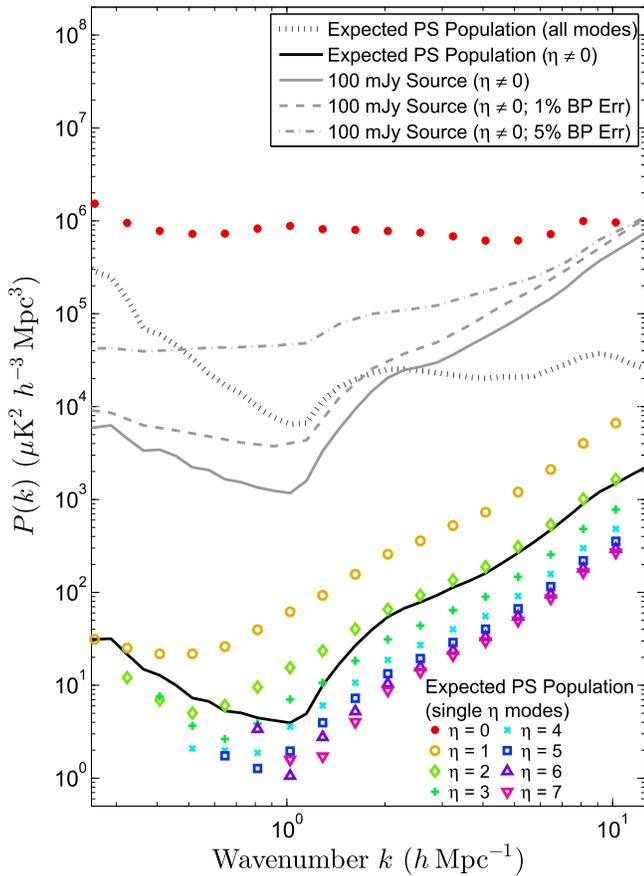


Figure 4. Results of simulations to probe the impact of point sources (PS) on our analysis. Separate simulations were used to probe the impact of a bright 100 mJy source (gray solid) and of the expected population of point sources. The measured power, averaged over three-dimensional shells in k , is shown before (black dotted) and after (black solid) removal of the $\eta = 0$ modes. For the 100 mJy source simulations, we additionally show the impact that bandpass (BP) errors have on the measured residual power, for RMS errors of 1% (gray dashed) and 5% (gray dotted-dashed). For the simulations of the expected point source population, we additionally show the power measured when using data from a single η channel.

at 30 GHz will be close to spectrally flat over the frequency range of one window (Muchovej et al. 2010), and therefore will contribute power primarily to Fourier modes with $\eta = 0$. To reject this contamination, we drop the $\eta = 0$ channel. Dropping this channel does not significantly affect the k range in the experiment, as the length of the smallest three-dimensional k vector is determined primarily by the shortest baselines, rather than the smallest η mode.

We test the effectiveness of our method of rejecting continuum source power with three sets of simulations. First, we simulated observations of mock fields with a 100 mJy point source of varying spectral index, inserted at a random position within twice the FWHM of the primary beam. Such a source is much brighter than average, and is somewhat brighter than the brightest single source found in our data. Second, we simulated observations of mock fields following the 30 GHz point source counts and spectral index distributions measured by the SZA in Muchovej et al. (2010; which we refer to as the “expected point source population”). We simulated sources of flux density up to 15 mJy, above which observations show a steep drop in point source counts. Third, we introduced calibration errors (i.e.,

baseline length, gains, and bandpass) into both our expected point source population and 100 mJy source simulations.

The results of these simulations are shown in Figure 4. For simulations probing the impact of the expected point source population, we additionally present a breakdown of the power measured in Fourier modes for individual η channels. Power from continuum sources is predominantly found in the $\eta = 0$ channel; as the one-dimensional power spectrum is an average over all remaining η , the effect of contamination is an average of the η channels shown.

In the absence of calibration errors, we find that rejection of the $\eta = 0$ mode reduces the contribution of continuum sources by factors of 10^1 – 10^4 , depending on k . We find that the residuals rise roughly as k^2 for $k \gtrsim 2 h \text{ Mpc}^{-1}$. This is due to the appreciable change of the baseline length over the bandwidth of a window for the long baselines of the array—an effect sometimes referred to as “mode-mixing” (Datta et al. 2010; Morales et al. 2012). Our experiment does not achieve interesting sensitivity to these modes, so we do not expect that this effect requires more thorough mitigation. We find that rejection of the $\eta = 0$ channel reduces the power spectrum contributions of a 100 mJy point source to levels comparable to the sensitivity of our experiment. In the real data, the suppression of power from bright point sources should be greater, as we remove such sources from the visibilities before calculating power spectra. We also find that rejection of the $\eta = 0$ channel reduces the power spectrum contributions of the expected source population to levels well below our sensitivity threshold. Further testing reveals that the point source residual power decreases by an additional factor of ~ 10 after removing sources of flux density greater than 1.5 mJy (roughly corresponding to the flux limit for source subtraction as discussed in Section 3.1).

In simulations that include calibration errors, we find that bandpass errors can significantly increase the residual power from continuum point sources. We find that for RMS fractional bandpass errors, σ_{BP} , between 1% and 10%, the residual power scales as σ_{BP}^2 . We expect that our bandpass errors are only 1% (discussed in Section 3.1); our simulations indicate that this will increase point source residuals by 30%–300%, depending on k . This effect is negligible for the expected point source population, though it may allow residuals from bright point sources to exceed our detection threshold. We note these simulations represent a worst-case scenario, as only a single set of bandpass errors are incorporated into each mock field. In real data, a new bandpass solution is derived for each observing block; hence, data for a typical field implement 45 different bandpass solutions. Assuming that the errors in these bandpass solutions are dominated by thermal noise, we expect the increase in residual power to be only a few percent of what we have found in these simulations.

We further find that gain and baseline errors do not significantly increase the residual point source power, though they do impact the fidelity with which we are able to remove bright sources from our data set. We also note that calibration errors of all varieties will reduce the sensitivity of our measurement (as these errors will “wash out” the signal of interest), by roughly a factor equal to the fractional error in the calibration. Given that our calibration errors are of the order of one part in a hundred, we expect this loss of sensitivity is likely to be minimal.

3.3. Jackknife Tests

To determine the impact of systematic errors on our measurements, we conduct a set of null tests in which we remove the astronomical signal from our data via linear combination or randomization and search for residual power. These tests, commonly referred to as jackknife tests, are split into four different categories: intraday, interday, “cross-window,” and noise tests. Results from our jackknife analysis can be found in Section 4.1.

Intraday tests take temporally adjacent visibilities (i.e., subsequent integrations of the same channel within the same baseline) and modulate them in such a way that the sky signals are canceled out. Two tests are performed in this manner: “couplet” tests (visibilities adjacent in time in a given baseline are simply subtracted from one another) and “triplet” tests (three adjacent visibilities are multiplied by a phase offset of 0° , 120° , and 240° respectively before being summed together). Visibilities are gain-corrected and flagged prior to this operation. In the event that one visibility of the couplet/triplet group is flagged as bad, the entire group is thrown out and not considered in subsequent jackknife analysis. Both couplet and triplet tests are sensitive to high-cadence systematics within our data set, with the couplet and triplet tests evaluating the data set over respectively shorter and longer periods.

Interday tests utilize gridded data rather than visibilities. Data from different observing blocks (i.e., days) are summed into two separate stacks and subtracted from one another. Two different interday tests are performed. The “even-odd” test sorts observations from alternating days into two stacks, and then subtracts them. The “first-last” test sums together data from all days belonging to the first half of the set of days and subtracts them from the sum of the second half.

The cross-window test operates by correlating different redshift windows with one another. The CO signal is not expected to correlate between different redshift windows, hence no subtraction-like step is necessary. However, continuum contamination that is not rejected by the removal of the $\eta = 0$ channel will correlate between windows, so this jackknife is an empirical check on our continuum rejection.

To further test the analysis software and pipeline, we perform a pair of tests to measure and confirm our noise estimates. In both tests, values are multiplied by a random complex number with absolute value of 1 and random phase. These “randomized phase” tests give us our best verification of noise estimates, as any potentially coherent emission detected by the interferometer should be scattered and thus not affect the results of this test. The “random raw” test randomizes the phases of ungridded data, while the “random grid” randomizes the phases of gridded data. While the former is much more thorough, the latter requires much less processing time to complete (while still being a useful tool for the verification of noise estimates).

We show the results of our jackknife analysis in Figure 5 and in Table 2. Listed in the table are the measured values (and associated errors) for each of the tests. Additionally listed for each test is the probability to exceed (PTE)—the likelihood for a noise-like event to produce a measurement of equal or greater statistical significance. Our jackknife results appear to be consistent with noise, suggesting that our analysis is not dominated by systematics. We note that our cross-window test correlates windows that are two steps apart (e.g., window 1 with window 3, window 2 with window 4). When correlating

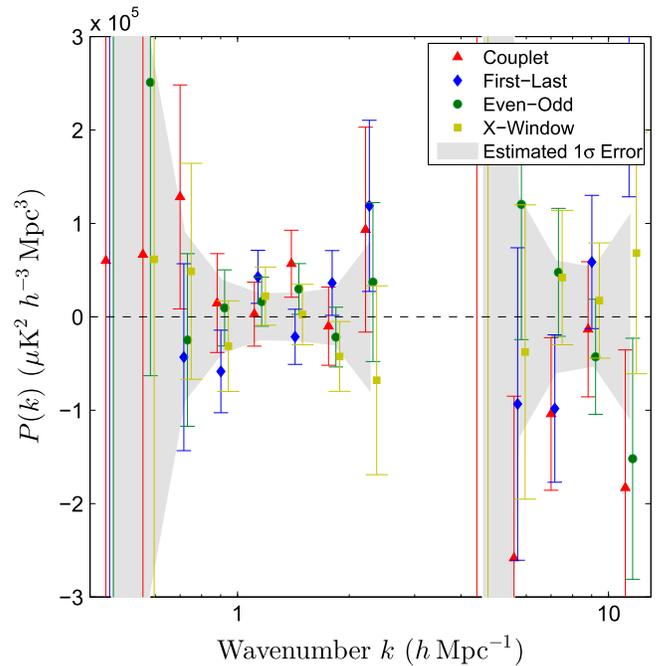


Figure 5. Jackknife results for our analysis of the S10 data set, alongside the 1σ noise threshold for our measurement (solid gray). For each bin, we find the results to be noise-like in their distribution, consistent with our assumption that the data are predominantly free of systematics that may affect our result. Our largest outlier has a significance of 2.0σ , consistent with what one would expect for a set of ~ 60 values that have been normally distributed.

Table 2
Table of Jackknife Test Results

Result Type	Measured P and 1σ errors ($10^4 \mu\text{K}^2 h^{-3} \text{Mpc}^3$)	PTE
Couplet	2.00 ± 1.74	0.25
Triplet	1.21 ± 2.46	0.62
Even-Odd	1.16 ± 1.34	0.39
First-Last	0.22 ± 1.47	0.88
Cross-Window	-0.16 ± 1.60	0.92

adjacent windows with one another (e.g., window 1 with window 2, window 3 with window 4), the cross-window test finds power at a level of 2.3σ significance. We attribute this power to signal leakage between windows (brought on by minor imperfections in analog bandpass filtering), and do not expect it to impact our analyses.

Additionally, we find good agreement between estimates for the RMS power induced by thermal noise and the random grid test. The RMS power is estimated by utilizing the measured system temperatures, aperture efficiency, and SEFD corrections to estimate the amount of noise power in a given delay-visibility (and propagating that noise estimate forward using Equation (1) to estimate the noise power in the power spectrum). The calculated uncertainty and the uncertainty derived from the random grid test agree to 10%, as we would expect for the 100 trials used in the test. We do find some minor differences between the two estimates in bins with lowest sensitivity; these bins have a small number of independent modes used to measure the power in the bin (hence their errors will not be normally distributed), therefore some discrepancy is to be expected.

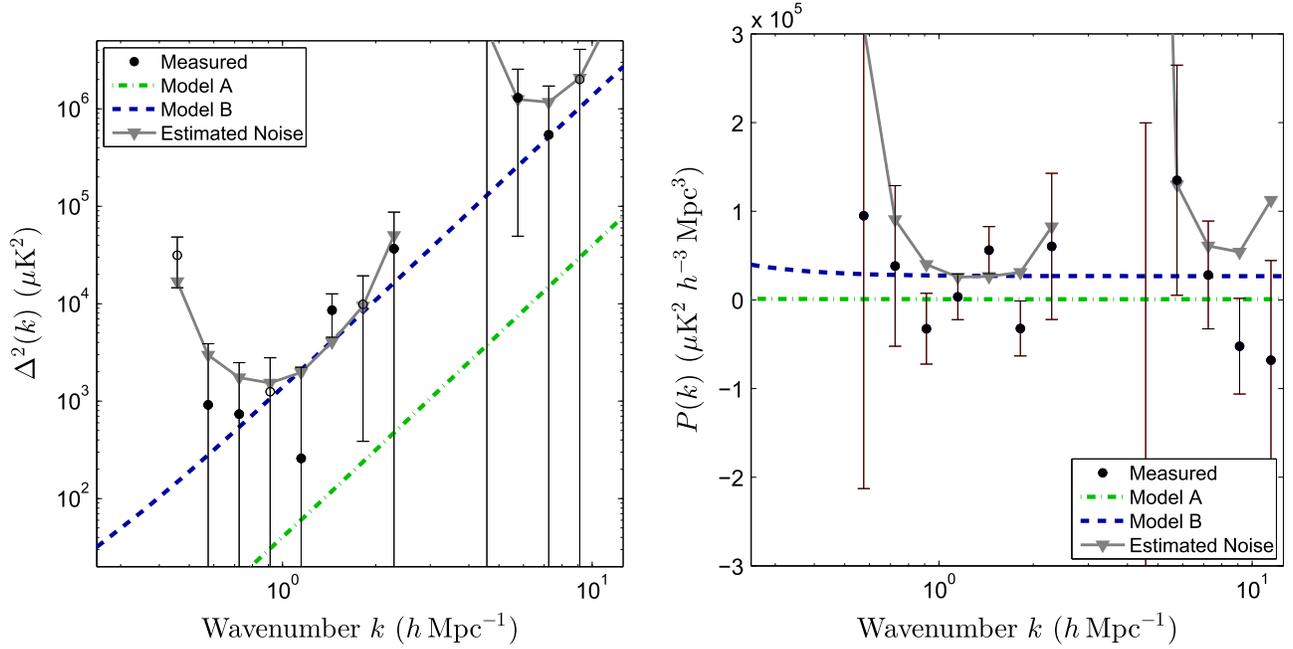


Figure 6. Left: the power spectrum result from our analysis of the S10 data, in the form $\Delta^2(k)$. Filled circles correspond to positive values for $\Delta^2(k)$, while open circles correspond to negative values, and the error bars correspond to the 1σ errors on our measured values. There exists a gap in sensitivity around $k \sim 4 h \text{ Mpc}^{-1}$, owing to the separation between baselines to the outrigger antennas and baselines within the compact portion of the array (see Figure 1). For reference, model A (dotted-dashed green) and model B (dashed blue) from Pullen et al. (2013) are shown (discussed further in Section 5.2), along with the estimated RMS noise power (gray triangle) in the absence of any astrophysical signal. Right: the power spectrum result, in the form of $P(k)$.

4. RESULTS

4.1. CO Power Spectra

Presented in Figure 6 are the final results of our analysis of the S10 data set. Our measurement has peak sensitivity at $k = 1.3 h \text{ Mpc}^{-1}$, with best sensitivity between $k = 0.5\text{--}2 h \text{ Mpc}^{-1}$. Integrating over all redshift windows and wavenumbers, we see no evidence for excess power from CO at $z \sim 3$ above a 2σ noise threshold of $P_N = 2.6 \times 10^4 \mu\text{K}^2 (h^{-1} \text{ Mpc})^3$. Placing this measurement into Δ_N^2 units, where Poisson power grows like k^3 , requires choosing a k value. At $k = 1 h \text{ Mpc}^{-1}$, $\Delta_N^2 = 1.3 \times 10^3 \mu\text{K}^2$.

Theoretical models (e.g., Pullen et al. 2013) suggest that there may be significant evolution between the redshift range sampled by these data ($z = 2.3\text{--}3.3$), thus we show in Figure 7 the results of our analysis for each redshift bin. With a mean 2σ noise limit of $P_N = 10^5 \mu\text{K}^2 (h^{-1} \text{ Mpc})^3$, $\Delta^2(k = 1 h \text{ Mpc}^{-1}) = 5 \times 10^3 \mu\text{K}^2$ within each redshift bin (of characteristic width $\Delta z = 0.06$), we see no evidence for excess power from CO within any of our individual redshift intervals.

5. DISCUSSION

5.1. Constraints on the CO Power Spectrum

The power spectrum for CO as a function of wavenumber and redshift is given by

$$P(k, z) = \langle T_{\text{CO}} \rangle^2 b^2(z) P_{\text{lin}}(k, z) + P_{\text{shot}}(z), \quad (3)$$

where $b(z)$ is the halo bias, P_{lin} is the linear matter power spectrum, and P_{shot} is the shot-noise contribution to the power spectrum. One can further define P_{shot} as

$$P_{\text{shot}}(z) = \int \left(\frac{c^3 (1+z)^2}{8\pi\nu_0^3 k_B H(z)} L \right)^2 \frac{dn(z)}{dL} dL, \quad (4)$$

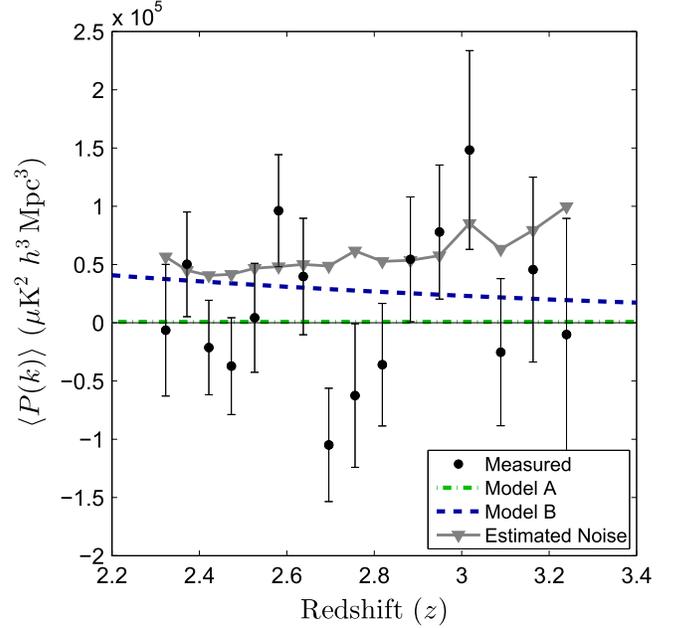


Figure 7. Measured variation of $P(k)$, averaged over all k , as a function of redshift, with the 1σ error bars shown for each point. We find that the results are consistent with noise, with a maximum significance 2.0σ at $z = 2.58$. The 2σ confidence upper limit resides just above the Model B prediction, with greatest sensitivity between $z = 2.3\text{--}2.8$.

where ν_0 is the rest frequency of the line (i.e., 115.271 GHz), $H(z)$ is the Hubble parameter, and $dn(z)/dL$ is the number of emitters per unit volume with luminosity L at a given redshift z . If there exists a linear relationship between CO luminosity and halo mass (as many models assume; see Table 3 of Li

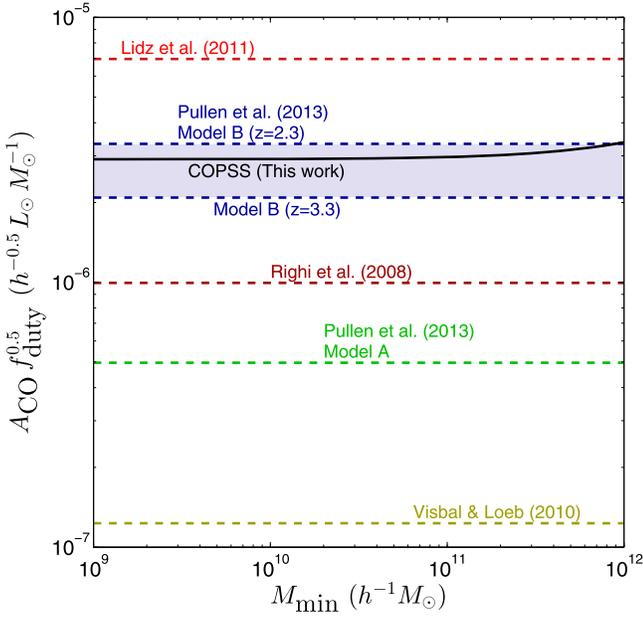


Figure 8. Constraints on A_{CO} as a function of M_{min} . The 2σ upper limit from our analysis (black solid) is shown vs. several theoretical expectations for A_{CO} (Breysse et al. 2014), multiplied by the square root of their adopted values for f_{duty} ($f_{\text{duty}} = 0.1$ for Visbal et al. 2011, $f_{\text{duty}} = t_s/t_H$ for all others). Righi et al. (2008) do not explicitly supply a value for A_{CO} or f_{duty} ; we have therefore adopted the value of A_{CO} calculated for this model by Breysse et al. (2014), and have adopted the value for f_{duty} from P13. The shaded region for model B reflects the model’s variation with redshift over the range of our measurement.

et al. 2015), then Equation (4) can be written as

$$P_{\text{shot}}(z) = \left(A_{\text{CO}} \frac{c^3(1+z)^2 L_{\odot}}{8\pi\nu_o^3 k_B H(z) M_{\odot}} \right)^2 \times f_{\text{duty}} \int_{M_{\text{min}}}^{\infty} M^2 \frac{dn(z)}{dM} dM, \quad (5)$$

where A_{CO} is the ratio of CO(1–0) luminosity to host halo mass for CO-luminous halos, expressed in units of $L_{\odot} M_{\odot}^{-1}$ (Lidz et al. 2011; Breysse et al. 2014), f_{duty} is the duty cycle of CO emitters, and $dn(z)/dM$ is the number of halos per unit mass as a function of redshift. Halos with masses below the low-mass limit, M_{min} , are assumed to be deficient in CO (either due to the suppression of star formation or because of a breakdown in the simple linear $M-L_{\text{CO}}$ relationship assumed), such that they do not appreciably contribute to P_{shot} (e.g., Pullen et al. 2013; Li et al. 2015). Similarly, the mean brightness temperature can be expressed as

$$\langle T_{\text{CO}} \rangle = A_{\text{CO}} \frac{c^3(1+z)^2 L_{\odot}}{8\pi\nu_o^3 k_B H(z) M_{\odot}} \times f_{\text{duty}} \int_{M_{\text{min}}}^{\infty} M \frac{dn(z)}{dM} dM. \quad (6)$$

Using Equation (5) and provided with an appropriate halo mass function (Tinker et al. 2008), one can use the upper limit presented in Section 4.1 to constrain the product $A_{\text{CO}} f_{\text{duty}}^{1/2}$. This constraint is shown in Figure 8 as a function of M_{min} , compared to several estimates from theoretical models.

5.2. Constraints on Theoretical Models

As shown in Figure 8, our constraint on A_{CO} falls below estimates from Lidz et al. (2011), though those estimates are tailored for $z \gtrsim 6$. Our measurement is not sensitive enough to place constraints on model A from Pullen et al. (2013, hereafter P13), or on estimates from Visbal & Loeb (2010) or Li et al. (2015; not shown in the figure, though comparable to model A). Our sensitivity limit is not far above model estimates by Righi et al. (2008); our measurement places constraints on optimistic versions of the Righi model, ruling out variants with an estimated power spectrum an order of magnitude stronger than the baseline prediction.

We consider the models of P13 in a bit more detail, as their model B is closest to our upper limit. P13 present two models (A and B) that predict the CO power spectrum based on locally observed correlations between SFR, far-infrared luminosity, and CO luminosity (Kennicutt 1998; Wang et al. 2011). Model A utilizes a predicted relationship between halo mass and SFR for power spectrum estimates, while model B uses SFR functions based on UV and IR observations (Smit et al. 2012).

Adopting the values for f_{duty} and M_{min} found in P13 ($M_{\text{min}} = 10^9 M_{\odot}$; $f_{\text{duty}} = t_s/t_{\text{age}}(z)$, where t_s is the star formation timescale, of the order of 10^8 yr, and $t_{\text{age}}(z)$ is the Hubble time at a given redshift), we constrain $A_{\text{CO}} < 1.2 \times 10^{-5} L_{\odot} M_{\odot}^{-1}$. Using Equation (6), we translate this to a constraint on the mean brightness temperature of the CO(1–0) transition at $z \sim 3$, $\langle T_{\text{CO}} \rangle < 4.8 \mu\text{K}$. As shown in Figure 6, our result is inconsistent (to 2σ significance) with the baseline expectation of model B from P13 at our average redshift of $z = 2.8$.

The CO abundance may change significantly over the redshift range of our measurement, owing in part to increasing metallicity and dust masses within galaxies (Valiante et al. 2009), rising feedback and quenching processes (Kereš et al. 2009; Wheeler et al. 2014), and the rapid depletion of neutral gas for star formation (Bauermeister et al. 2010) over cosmic time. Model B of P13 effectively makes such a prediction, with A_{CO} rising with decreasing redshift as one approaches the peak of cosmic star formation (as reflected by the shaded area in Figure 8). As previously discussed in Section 4.1, we see no evidence for a change in CO abundance as a function of redshift (evaluated over Δz of 0.12, 0.24, and 0.49). Evaluating the P13 models over shorter redshift ranges, we are able to rule out the baseline version of Model B to 2.1σ significance for $z = 2.3$ – 2.8 , and to 1.2σ significance for $z = 2.8$ – 3.3 .

We note that values of f_{duty} and M_{min} are not well constrained; in particular, the value of f_{duty} differs by more than an order of magnitude between different models. P13 argues $f_{\text{duty}} = t_s/t_{\text{age}}(z)$, which is ~ 0.05 for the redshift range of our measurement for their assumed $t_s = 10^8$ yr. This choice of t_s is based on arguments that the dominant source of CO emission at high redshift ($z \gtrsim 6$) will arise from galaxies undergoing extreme starburst events (Righi et al. 2008; Lidz et al. 2011). However, there are multiple observations indicating higher star formation duty cycles at $z \sim 1$ – 4 , with f_{duty} approaching 100% (Noeske et al. 2007; Lee et al. 2009; Tacconi et al. 2013). The Li et al. (2015) model implicitly assumes a 100% duty cycle, though they introduce intrinsic scatter model parameters for the SFR to L_{CO} and halo mass to SFR relations to allow for observed variations in halo activity. Following Equation (5), our constraint on A_{CO} depends on

$f_{\text{duty}}^{-0.5}$, so increasing it to unity would drop our limit on A_{CO} by a factor of 4, though the unaccounted scatter in halo properties noted by Li et al. (2015) would mimic a lower f_{duty} . We use the low value of f_{duty} from P13 for our conservative upper limit on A_{CO} , and show the combination of A_{CO} and f_{duty} in Figure 8. Variations in values of M_{min} have smaller effects on the model constraints, as can be seen in Figure 8. Uncertainties in the halo mass function (i.e., dn/dM) contribute insignificantly to the uncertainty. Tinker et al. (2008) found 1% accuracy for their fitting function over the range of masses of interest for this problem, which translates into a 1% uncertainty in the quantity of interest for determining A_{CO} , the second moment of dn/dM .

5.3. Constraints on the Cosmic CO Luminosity and Density of H_2

Following the arguments leading up to Equation (6), one may write a simple expression for the volume emissivity of CO (1–0), $\varepsilon_{\text{CO}(1-0)}$:

$$\varepsilon_{\text{CO}(1-0)} = f_{\text{duty}} A_{\text{CO}} \frac{L_{\odot}}{M_{\odot}} \int_{M_{\text{min}}}^{\infty} M \frac{dn}{dM} dM. \quad (7)$$

Adopting fiducial values from P13 ($M_{\text{min}} = 10^9 M_{\odot}$, $f_{\text{duty}} = t_s/t_{\text{age}}(z)$), from our constraint of $A_{\text{CO}} \leq 1.2 \times 10^{-5} L_{\odot} M_{\odot}^{-1}$ we calculate an upper limit for the CO(1–0) volume emissivity of $\varepsilon_{\text{CO}(1-0)} \leq 9 \times 10^3 L_{\odot} h^3 \text{Mpc}^{-3}$. Our measurement offers an improved constraint on the maximum cosmic CO(1–0) luminosity at $z \sim 3$ presented by Walter et al. (2014, who determined $\varepsilon_{\text{CO}(1-0)} \lesssim 2 \times 10^4 L_{\odot} h^3 \text{Mpc}^{-3}$), as it is an integrated measurement across the entire population range of galaxies, whereas Walter et al. (2014) exclude consideration of galaxies below their detection threshold.

Assuming a linear relationship between L_{CO} and M_{H_2} , one can use the conversion factor from CO luminosity to molecular gas mass, α_{CO} , to determine an upper limit for the cosmic density of molecular gas. Assuming a Milky Way-like $\alpha_{\text{CO}} = 4.3 M_{\odot} (\text{K km s}^{-1} \text{pc}^{-2})^{-1}$ (Frerking et al. 1982; Dame et al. 2001)—equivalent to $8.7 \times 10^4 M_{\odot} L_{\odot}^{-1}$ (Solomon et al. 1992)—we find an upper limit for the global density of molecular gas of $\rho_{z \sim 3}(M_{H_2}) \leq 9 \times 10^8 M_{\odot} h^3 \text{Mpc}^{-3}$, which lies only a factor of ~ 2 – 3 above theoretical predictions (Obreschkow et al. 2009a; Lagos et al. 2011; Sargent et al. 2014). Assuming $h = 0.7$, our limit is equal to $\rho_{z \sim 3}(M_{H_2}) \leq 2.8 \times 10^8 M_{\odot} \text{Mpc}^{-3}$, in agreement with measurements made by Walter et al. (2014), who calculate an upper limit of approximately $4 \times 10^8 M_{\odot} \text{Mpc}^{-3}$.

5.4. Cosmic Variance and Limits of Significance

We now consider the impact of cosmic variance on our measurement. In the shot-noise regime, the power measured is roughly proportional to the number density of emitters within the volume measured, n_e . The variance in the measured number density depends inversely on the total number of emitters detected over a given volume, $N_e = n_e V_z$. Therefore, provided with a halo mass function and a scaling relationship between L_{CO} and halo mass, one can calculate an estimate for cosmic variance. With a total survey volume of $3.6 \times 10^6 h^{-3} \text{Mpc}^3$, and assuming linear scaling between halo mass and L_{CO} with $f_{\text{duty}} = t_s/t_{\text{age}}(z)$, we find that cosmic variance effectively induces an error of $\Delta P/P \approx 0.1$ in our measurement. However, we note that f_{duty} has a strong effect on the impact

of cosmic variance, and that higher values of f_{duty} (suggested by some models, as discussed in Section 5.2) will decrease the impact of cosmic variance. In the case of an upper limit for the power spectrum, such as the one established in this experiment, we find that the impact of cosmic variance is negligible and may safely be ignored.

6. CONCLUSION

In this paper, we have constrained the power spectrum for CO at $z \sim 3$ to $P_{\text{CO}} < 2.6 \times 10^4 \mu\text{K}^2 (h^{-1} \text{Mpc})^3$, or $\Delta_{\text{CO}}^2(k=1 h \text{Mpc}^{-1}) < 1.3 \times 10^3 \mu\text{K}^2$, to 2σ confidence. We have used this constraint to place limits on the mean brightness temperature for CO and the global density of molecular gas. We have ruled out the model of Lidz et al. (2011) for $z \sim 3$, and have placed constraints on model B from Pullen et al. (2013).

In recent years, additional observations—focused on intensity mapping of CO(1–0) at $z \sim 3$ —have been performed with the SZA. Data from these observations will yield a significantly more sensitive measurement, and will offer improved constraints on the abundance of CO and molecular gas in the early universe. Other compact low-resolution centimeter-wave instruments, such as the Yuan-Tseh Lee Array (Ho et al. 2009), will offer increased sensitivity and will be capable of deeply probing the CO power spectrum at $z \sim 3$.

The authors would like to thank F. M. Fornasini and R. L. Plambeck for their valuable feedback during the preparation of this manuscript. This work was supported in part by the National Science Foundation University Radio Observatories Program, AST-1140031. We gratefully acknowledge the James S. McDonnell Foundation, the National Science Foundation, and the University of Chicago for funding to construct the SZA. The operation of the SZA was supported by NSF through award AST-0604982. Partial support was provided by NSF Physics Frontier Center grant PHY-0114422 to the Kavli Institute of Cosmological Physics at the University of Chicago, and by NSF grants AST-0507545 and AST-0507161 to Columbia University.

APPENDIX A THE INTENSITY MAPPING TECHNIQUE

The goal of intensity mapping is to use the two-point autocorrelation function, $\xi(\mathbf{x})$, to probe the underlying population of galaxies bearing molecular gas (through the tracer CO molecule) within some representative volume, V_z , of the universe at redshift z . The final analysis product is a power spectrum, $P(k)$, which along with its dimensionless counterpart, $\Delta^2(k)$, is related to the two-point autocorrelation function by

$$\begin{aligned} \Delta^2(k) &\equiv \frac{k^3}{2\pi^2} P(k), \\ &= \frac{k^3}{2\pi^2} \langle \mathcal{F}_{x \rightarrow k}(\xi(\mathbf{x})) \rangle_{k:k=k^2}, \end{aligned} \quad (8)$$

where $\mathcal{F}_{x \rightarrow k}$ is the Fourier transform from configuration to Fourier space, and \mathbf{k} is the vector wavenumber (of magnitude k) in Fourier space. The two-point autocorrelation function is

further defined as

$$\xi(\mathbf{x}) = \frac{1}{W(\mathbf{x}) \cdot W(\mathbf{x})} \left(\frac{c^2}{2k_B \nu^2} \right)^2 \times \left((W(\mathbf{x})I(\mathbf{x}')) \star (W(\mathbf{x})I(\mathbf{x})) \right), \quad (9)$$

where $I(\mathbf{x})$ is the specific intensity observed at a given frequency, ν . $W(\mathbf{x})$ is the spatial windowing function applied over the volume in question. $W(\mathbf{x}) \cdot W(\mathbf{x})$ is the normalization term of the autocorrelation function, which is equal to the effective volume probed, V_z . The Fourier dual of the autocorrelation function is $\tilde{\xi}(\mathbf{k})$, such that $P = \langle \tilde{\xi}(\mathbf{k}) \rangle$. The function $\tilde{\xi}(\mathbf{k})$ can then be expressed as

$$\tilde{\xi}(\mathbf{k}) = \frac{1}{V_z} \left(\frac{c^2}{2k_B \nu^2} \right)^2 \times \left((\tilde{W}^*(\mathbf{k}) * \tilde{I}^*(\mathbf{k})) (\tilde{W}(\mathbf{k}) * \tilde{I}(\mathbf{k})) \right), \quad (10)$$

where $\tilde{I}(\mathbf{k})$ is the Fourier dual of $\tilde{I}(\mathbf{x})$. We map between comoving physical size, $\mathbf{r} = (r_x, r_y, r_z)$, and the native observing units of $(l, m, \Delta\nu)$, where l and m describe the angular position (in units of radians) and $\Delta\nu$ is the change in line frequency due to expansion of the universe, with the following expressions:

$$l = \frac{r_x}{D_M(z)}, \quad m = \frac{r_y}{D_M(z)}, \quad \Delta\nu = \frac{H_0 E(z) \nu_{\text{rest}}}{c(1+z)^2} r_z. \quad (11)$$

In Equation (11), $D_M(z)$ is the comoving radial distance for redshift z , ν_o is the rest frequency of the line transition, H_0 is the current Hubble parameter, and $E(z)$ is the dimensionless Hubble parameter. To evaluate $\xi(\mathbf{k})$, one converts (u, v, η) , the Fourier dual of $(l, m, \Delta\nu)$, to (k_x, k_y, k_z) :

$$u = \frac{k_x X}{2\pi}, \quad v = \frac{k_y X}{2\pi}, \quad \eta = \frac{k_z Y}{2\pi},$$

using

$$X = D_M(z), \quad Y = \frac{c(1+z)^2}{H_0 E(z) \nu_{\text{rest}}}. \quad (12)$$

In Equation (12), we have included the variables X and Y (defined in Parsons et al. 2012) for brevity only, with units of Mpc rad^{-1} and Mpc Hz^{-1} respectively. Rewriting Equation (10) and making the proper variable substitutions, we derive the following expression:

$$\tilde{\xi}(\mathbf{k}) = \frac{X^4 Y^2}{V_z} \left(\frac{c^2}{2k_B \nu^2} \right)^2 \times \left((\tilde{W}^*(u, v, \eta) * \tilde{I}^*(u, v, \eta)) \right) \times \left(\tilde{W}(u, v, \eta) * \tilde{I}(u, v, \eta) \right). \quad (13)$$

With our simplified expression, we now turn our attention to the interferometer response to determine its relation to $\tilde{I}(u, v, \eta)$, and thus the measurement of $\tilde{\xi}(\mathbf{k})$. Under the flat-

sky approximation, the interferometer response is given by

$$\mathcal{V}(u, v, \nu) = \mathcal{F}_{(l,m) \rightarrow (u,v)}(I(l, m, \nu) \cdot A_\nu(l, m)) = \tilde{I}(u, v, \nu) * \tilde{A}_\nu(u, v), \quad (14)$$

where measurements for the interferometer response, $\mathcal{V}(u, v, \nu)$, are commonly referred to as ‘‘visibilities’’ (Thompson et al. 1986). \mathcal{F} in Equation (14) is the Fourier transform operator, which in this case transforms between (l, m) and (u, v) domains, and $A_\nu(l, m)$ is the primary beam response of the telescope at a given frequency ν and position in the sky (l, m) . We wish, however, to move from (u, v, ν) to (u, v, η) and create what we will refer to as ‘‘delay-visibility.’’ This requires performing one last Fourier transform over some window in frequency $\mathcal{B}(\nu)$:

$$\tilde{\mathcal{V}}(u, v, \eta) = \mathcal{F}_{\nu \rightarrow \eta}(V(u, v, \nu) \cdot \mathcal{B}(\nu)), = (\tilde{I}(u, v, \eta) * A_\nu(u, v)) * \tilde{\mathcal{B}}(\eta) = \tilde{I}(u, v, \eta) * \tilde{W}(u, v, \eta). \quad (15)$$

The delay-visibility in Equation (15) is the convolution of \tilde{I} and our windowing function, which we have defined as $\tilde{W}(u, v, \eta) = \tilde{A}(u, v) * \tilde{\mathcal{B}}(\eta)$, based on the windowing functions created by the primary beam pattern and the bandwidth window. We define $\Omega_{B,\nu} = \int A_\nu d\Omega$ to be the solid angle of the primary beam of the telescope, where A_ν is the primary beam pattern (and the Fourier dual of the aperture function, \tilde{A}), for a given frequency ν . Nominally both A and Ω are frequency-dependent, but the frequency range of each window is small enough that both $\Omega_{B,\nu}$ and A_ν can be treated as constant over the volume of interest. We also define $B_z = \int \mathcal{B}(\nu) d\nu$ as the effective bandwidth for our observation of the volume of interest, where $\mathcal{B}(\nu)$ is the frequency windowing function. In practice, this windowing function is generally equal to unity over the correlator window, hence for our measurement B_z is equal to the correlator window bandwidth. We will assume that our range in redshift over which our volume V_z subsides is small enough that X and Y remain constant. With this assumption we can further simplify our expression by substituting V_z with defined instrument parameters. The effective volume can be defined as

$$V = \int W \cdot W = X^2 Y \int A B \cdot A B d\Omega d\nu = X^2 Y B_z \Omega_{B,\nu} / 2. \quad (16)$$

In Equation (16), we have made use of the fact that for a Gaussian beam, $\int A^2 = \Omega_B / 2$, and have assumed a uniform weighting in frequency (i.e., $\mathcal{B} = 1$ over the frequency range of the window, otherwise $\mathcal{B} = 0$), as it is the weighting scheme used in our analysis. We can now simplify Equation (13), and come up with an expression for Equation (8) that contains only observables:

$$P(k) = \frac{2X^2 Y}{\Omega_B B_z} \left(\frac{c^2}{2k_B \nu^2} \right)^2 \left\langle \tilde{\mathcal{V}}^*(u, v, \eta) \tilde{\mathcal{V}}(u, v, \eta) \right\rangle_{\mathbf{k} \cdot \mathbf{k} = k^2}, \quad (17)$$

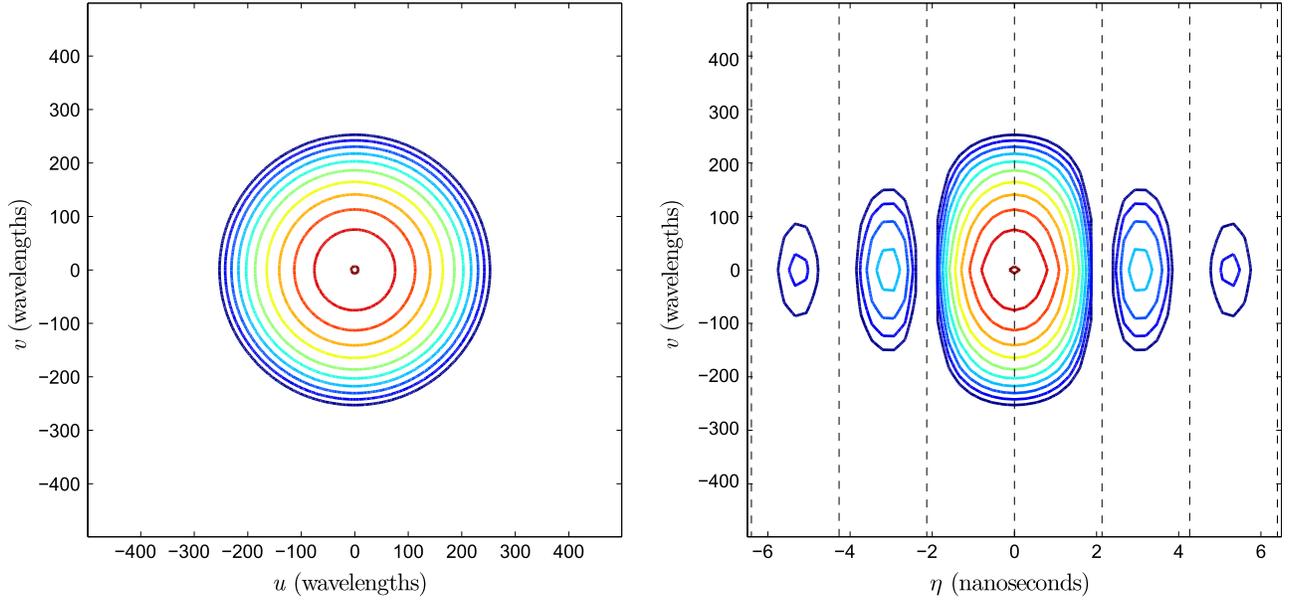


Figure 9. Left: a single cut through the function \tilde{W}^2 , where $\eta = 0$. Different contours correspond to 0.2 dex differences in sensitivity, from unity (dark red) to 0.01 (dark blue). Right: another slice through \tilde{W}^2 , where $u = 0$. Dashed lines mark the discrete values of η sampled by our analysis.

$$\Delta^2(k) = \frac{k^3}{\pi^2} \frac{X^2 Y}{\Omega_B B_z} \left(\frac{c^2}{2k_B \nu^2} \right)^2 \left\langle \tilde{\mathcal{V}}^*(u, v, \eta) \tilde{\mathcal{V}}(u, v, \eta) \right\rangle_{k \cdot k = k^2}. \quad (18)$$

A subtle feature of Equations (17) and (18) is that each delay-visibility is a weighted mean of several values of $\tilde{I}^2(u, v, \eta)$, smeared together by the windowing function \tilde{W}^2 (shown in Figure 9). This smearing effect in (u, v, η) has three significant consequences worth noting.

The first consequence is that the number of independent k -modes of the power spectrum, N_k , will be limited by the size of the windowing function. The size of the “footprint” of a measurement is $\Delta\tilde{V} = \int W^2$. While this value does not directly affect the estimators \hat{P} and $\hat{\Delta}^2$, it will affect the statistical significance of our measurement by limiting the number of independent modes one can measure.

The second consequence is that $|\hat{I}^2|$ may change appreciably over the range in (u, v, η) that our delay-visibility is spread over, particularly in cases where the delay-visibility probes a large range in spatial scales (a scenario not applicable to the S10 data set, where $\Delta k < k$). If unaccounted for, this may bias estimates of the power spectrum and lead to misinterpretation of results.

The final consequence is that delay-visibilitys centered at different positions in (u, v, η) may measure partially overlapping regions of $\tilde{I}(u, v, \eta)$, and can produce a measurement of the power spectrum. Invoking the assumption that the universe is homogeneous and isotropic on sufficiently large scales, orthogonal spatial modes for a randomized brightness temperature fluctuation field will be incoherent with one another, such that $\langle \tilde{I}^*(k) \tilde{I}(k') \rangle_{k \neq k'} = 0$. The product of two delay-visibilitys offset by some distance in (u, v, η) produces a coherent measurement of the specific intensity variance, I^2 , weighted by the inner product of the two offset windowing

functions:

$$C(\Delta u, \Delta v, \Delta \eta) = \frac{\int W(u, v, \eta) \cdot W(u + \Delta u, v + \Delta v, \eta + \Delta \eta) du dv d\eta}{\int W(u, v, \eta) \cdot W(u, v, \eta) du dv d\eta}. \quad (19)$$

$C(\Delta u, \Delta v, \Delta \eta)$ is the normalized covariance (with respect to the astronomical signal being measured) of two delay-visibilitys separated by some distance $(\Delta u, \Delta v, \Delta \eta)$, normalized such that two delay-visibilitys with zero separation have a normalized covariance of $C(0, 0, 0) = 1$. With this we can now define our estimator, $|\hat{I}^2|$, for the true specific intensity variance, $|I^2|$, as

$$|\hat{I}^2(u, v, \eta)| = \frac{\sum_n^{n \neq m} \tilde{\mathcal{V}}_m^*(u, v, \eta) \tilde{\mathcal{V}}_n(u', v', \eta') C(u - u', v - v', \eta - \eta') (\sigma_m^2 \sigma_n^2)^{-2}}{\sum_n^{n \neq m} (C(u - u', v - v', \eta - \eta') / \sigma_m^2 \sigma_n^2)^2} \quad (20)$$

where σ_m^2 is the noise estimate in a given visibility, \mathcal{V}_m . The sum is carried out for each visibility $\mathcal{V}_m(u, v, \eta)$ over all other independent visibilitys $\mathcal{V}_n(u', v', \eta')$. We discard any terms that arise from the product of a delay-visibility with its own complex conjugate in order to prevent our estimator from being positively biased by instrumental noise. Our sum in Equation (20) has been naturally weighted so as to maximize the sensitivity of our measurement (though other weighting schemes can be used). With this, we define our estimators for

the power spectrum, \hat{P} and $\hat{\Delta}^2$, as

$$\hat{P}(k) = \frac{2X^2Y}{\Omega_B B_z} \left(\frac{c^2}{2k_B \nu^2} \right)^2 \left\langle \left| \hat{I}^2 \right| \right\rangle_{k \cdot k = k^2}, \quad (21)$$

$$\hat{\Delta}^2(k) = \frac{k^3}{\pi^2} \frac{X^2Y}{\Omega_B B_z} \left(\frac{c^2}{2k_B \nu^2} \right)^2 \left\langle \left| \hat{I}^2 \right| \right\rangle_{k \cdot k = k^2}. \quad (22)$$

APPENDIX B SENSITIVITY ESTIMATES FOR INTENSITY MAPPING EXPERIMENTS

Utilizing Equations (17) and (18), one may construct a simple estimate for the sensitivity that one may achieve for a given choice of instruments and observations. The estimated noise for the power spectrum may be defined as

$$P_N = \frac{2X^2Y}{\Omega_B B_z} \left(\frac{c^2}{2k_B \nu^2} \right)^2 \frac{\tilde{V}_N^2}{N_V^{1/2} N_k^{1/2} N_{rd}}, \quad (23)$$

where \tilde{V}_N is the noise within a single delay-visibility, N_V is the number of independent volumes measured, N_k is the number of independent measurements made in the k -space volume, and N_{rd} is the number of redundant measurements made at a given position of (u, v, η) . Note that $N_{meas} = N_V N_k$. \tilde{V}_N can be further defined as

$$\tilde{V}_N = \frac{2k_B \nu^2 \Omega_b T_{sys}}{c^2} \frac{\eta_{eff}}{\tau_{int}} \sqrt{\frac{B}{\tau_{int}}}, \quad (24)$$

where T_{sys} is the system temperature, η_{eff} is the aperture efficiency, B is the bandwidth over which the measurement is being made, and τ is the integration time (Thompson et al. 1986; Parsons et al. 2012).

The number of independent volumes probed is simply the product of the number of fields, N_f , and the number of redshift windows, N_z . The number of redundant measurements, N_{rd} , and number of independent measurements in the k -space volume, N_k , are heavily intertwined and influenced by both the array and observing configuration. Consider an array with a number of baselines N_{base} . In a minimally redundant array, each of these baselines produces an independent measurement. However, in a redundant array (where elements are placed on an equally spaced grid) or in an extremely compact array, some baselines will overlap in the uv plane. If we define N_{ob} as the typical number of baselines that instantaneously share the same position in the uv plane (where $N_{ob} = 1$ corresponds to a minimally redundant array), then the number of independent measurements will decrease as N_{ob}^{-1} , but the number of redundant measurements will rise as N_{ob} . While compact, the SZA is laid out in such a way that the instantaneous baseline coverage produces only a few, partially overlapping baselines; hence, $N_{ob,SZA} \approx 1$.

When the array is not coplanar with the sky, the projected baselines will change position over time—an effect commonly referred to as Earth rotation aperture synthesis (ERAS; Thompson et al. 1986). ERAS reduces the amount of time that a given baseline is found at a particular position in the uv plane. Defining f_r as the fraction of the integration time in which baselines are found at already sampled positions in the uv plane (either because a baseline remains in a single position

for that time, or the baseline moves into a new position in the uv plane previously sampled by another baseline), then the number of redundant measurements scales linearly with f_r , though the number of independent measurements will scale as f_r^{-1} . The calculation of f_r will be heavily affected by the position of the source(s) of interest, as well as the range in hour angle over which observations are conducted, but a simple estimate for redundant fraction is $f_r = 2d_a^2/d_{B,med}^2$, where d_a is the diameter of the antenna and $d_{B,med}$ is the median baseline length. The justification for this estimate is that the interferometer will sweep out an area roughly equal to $\pi d_{B,med}^2$ in the uv plane, whereas the area instantaneously sampled by a single baseline is $2\pi d_a^2$ (the factor of 2 arises from the Hermitian nature of the uv plane). Therefore, the interferometer measures approximately $d_{B,med}^2/2d_a^2$ independent positions. Excluding baselines with the outrigger antennas, the median SZA baseline is approximately 6.5 m, leading to $f_{r,SZA} \approx 0.6$.

Finally, if one has N_{ch} frequency channels across a particular redshift window, then there are an equal number of delay-channels, each of which provides an independent measurement. One caveat to this statement is that if the frequency resolution is too fine, then the flux from an individual emitter may become spread across several channels, leaving some delay-visibilitys to resolve out the emission and not meaningfully contribute to the measurement (at a velocity resolution of ~ 300 km s $^{-1}$, this should not be an issue for the SZA). Combining all of the above pieces of information, the estimates for the number independent and redundant measurements may be written as

$$N_{rd} \approx N_{ob} f_r, \quad (25)$$

$$N_k \approx N_{base} N_{ob}^{-1} f_r^{-1} N_{ch}. \quad (26)$$

Combining Equations (25) and (26) with Equation (23),

$$P_N \sim 2X^2Y\Omega_B \times \frac{(T_{sys}/\eta_{eff})^2}{\tau_{int} (2N_{fields} N_z N_{base} N_{ob} N_{ch})^{1/2}} \frac{d_{B,med}}{d_a}. \quad (27)$$

Plugging in values from the S10 data set, we obtain an estimated sensitivity of $P_N \sim 1.3 \times 10^4 \mu\text{K}^2 (h^{-1} \text{Mpc})^3$, very close to the actual sensitivity achieved in Section 4.

REFERENCES

- Bauermeister, A., Blitz, L., & Ma, C.-P. 2010, *ApJ*, 717, 323
 Bennett, C. L., Larson, D., Weiland, J. L., et al. 2013, *ApJS*, 208, 20
 Bolatto, A. D., Wolfire, M., & Leroy, A. K. 2013, *ARA&A*, 51, 207
 Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2012, *ApJL*, 752, L5
 Breysse, P. C., Kovetz, E. D., & Kamionkowski, M. 2014, *MNRAS*, 443, 3506
 Carilli, C. L. 2011, *ApJL*, 730, L30
 Dame, T. M., Hartmann, D., & Thaddeus, P. 2001, *ApJ*, 547, 792
 Datta, A., Bowman, J. D., & Carilli, C. L. 2010, *ApJ*, 724, 526
 Decarli, R., Walter, F., Carilli, C., et al. 2014, *ApJ*, 782, 78
 Dickinson, M., Giavalisco, M., & GOODS Team 2003, in *The Mass of Galaxies at Low and High Redshift*, ed. R. Bender & A. Renzini (Berlin: Springer), 324
 Dodelson, S. 2003, *Modern Cosmology* (New York: Academic Press)
 Downes, D., Solomon, P. M., & Radford, S. J. E. 1993, *ApJL*, 414, L13
 Frerking, M. A., Langer, W. D., & Wilson, R. W. 1982, *ApJ*, 262, 590
 Furlanetto, S. R., Oh, S. P., & Briggs, F. H. 2006, *PhR*, 433, 181
 Genzel, R., Tacconi, L. J., Combes, F., et al. 2012, *ApJ*, 746, 69
 Glover, S. C. O., & Mac Low, M.-M. 2011, *MNRAS*, 412, 337
 Gong, Y., Cooray, A., Silva, M. B., Santos, M. G., & Lubin, P. 2011, *ApJL*, 728, L46
 Ho, P. T. P., Altamirano, P., Chang, C.-H., et al. 2009, *ApJ*, 694, 1610
 Högbom, J. A. 1974, *A&AS*, 15, 417
 Hopkins, A. M., & Beacom, J. F. 2006, *ApJ*, 651, 142

- Israel, F. P. 1997, *A&A*, **328**, 471
- Kennicutt, R. C., Jr. 1998, *ApJ*, **498**, 541
- Kereš, D., Katz, N., Fardal, M., Davé, R., & Weinberg, D. H. 2009, *MNRAS*, **395**, 160
- Lagos, C. D. P., Baugh, C. M., Lacey, C. G., et al. 2011, *MNRAS*, **418**, 1649
- Lee, K.-S., Giavalisco, M., Conroy, C., et al. 2009, *ApJ*, **695**, 368
- Li, T. Y., Wechsler, R. H., Devaraj, K., & Church, S. E. 2015, arXiv:1503.08833
- Lidz, A., Furlanetto, S. R., Oh, S. P., et al. 2011, *ApJ*, **741**, 70
- Morales, M. F., Hazelton, B., Sullivan, I., & Beardsley, A. 2012, *ApJ*, **752**, 137
- Muchovej, S., Leitch, E., Carlstrom, J. E., et al. 2010, *ApJ*, **716**, 521
- Muchovej, S., Mroczkowski, T., Carlstrom, J. E., et al. 2007, *ApJ*, **663**, 708
- Muñoz, J. A., & Furlanetto, S. R. 2014, *MNRAS*, **438**, 2483
- Noeske, K. G., Weiner, B. J., Faber, S. M., et al. 2007, *ApJL*, **660**, L43
- Obreschkow, D., Croton, D., De Lucia, G., Khochfar, S., & Rawlings, S. 2009a, *ApJ*, **698**, 1467
- Obreschkow, D., Heywood, I., Klöckner, H.-R., & Rawlings, S. 2009b, *ApJ*, **702**, 1321
- Parsons, A., Pober, J., McQuinn, M., Jacobs, D., & Aguirre, J. 2012, *ApJ*, **753**, 81
- Planck Collaboration, Aghanim, N., Arnaud, M., et al. 2015, arXiv:1507.02704
- Pullen, A. R., Chang, T.-C., Doré, O., & Lidz, A. 2013, *ApJ*, **768**, 15
- Regan, M. W., Thornley, M. D., Helfer, T. T., et al. 2001, *ApJ*, **561**, 218
- Righi, M., Hernández-Monteagudo, C., & Sunyaev, R. A. 2008, *A&A*, **489**, 489
- Rudy, D. J., Muhleman, D. O., Berge, G. L., Jakosky, B. M., & Christensen, P. R. 1987, *Icar*, **71**, 159
- Sargent, M. T., Daddi, E., Béthermin, M., et al. 2014, *ApJ*, **793**, 19
- Schmidt, M. 1959, *ApJ*, **129**, 243
- Sharp, M. K., Marrone, D. P., Carlstrom, J. E., et al. 2010, *ApJ*, **713**, 82
- Smit, R., Bouwens, R. J., Franx, M., et al. 2012, *ApJ*, **756**, 14
- Solomon, P. M., Downes, D., & Radford, S. J. E. 1992, *ApJL*, **398**, L29
- Solomon, P. M., Downes, D., Radford, S. J. E., & Barrett, J. W. 1997, *ApJ*, **478**, 144
- Tacconi, L. J., Neri, R., Genzel, R., et al. 2013, *ApJ*, **768**, 74
- Thompson, A. R., Moran, J. M., & Swenson, G. W. 1986, *Interferometry and Synthesis in Radio Astronomy* (New York: Wiley)
- Tinker, J., Kravtsov, A. V., Klypin, A., et al. 2008, *ApJ*, **688**, 709
- Valiante, R., Schneider, R., Bianchi, S., & Andersen, A. C. 2009, *MNRAS*, **397**, 1661
- Visbal, E., & Loeb, A. 2010, *JCAP*, **11**, 16
- Visbal, E., Trac, H., & Loeb, A. 2011, *JCAP*, **8**, 10
- Walter, F., Decarli, R., Sargent, M., et al. 2014, *ApJ*, **782**, 79
- Wang, R., Wagg, J., Carilli, C. L., et al. 2011, *AJ*, **142**, 101
- Wheeler, C., Phillips, J. I., Cooper, M. C., Boylan-Kolchin, M., & Bullock, J. S. 2014, *MNRAS*, **442**, 1396
- Wilson, R. W., Jefferts, K. B., & Penzias, A. A. 1970, *ApJL*, **161**, L43
- Wolfire, M. G., Hollenbach, D., & McKee, C. F. 2010, *ApJ*, **716**, 1191
- Young, J. S., & Scoville, N. 1982, *ApJ*, **258**, 467
- Young, J. S., Xie, S., Tacconi, L., et al. 1995, *ApJS*, **98**, 219