

IMPROVING FOREGROUND SUBTRACTION IN STATISTICAL OBSERVATIONS OF 21 cm EMISSION FROM THE EPOCH OF REIONIZATION

MIGUEL F. MORALES,¹ JUDD D. BOWMAN,² AND JACQUELINE N. HEWITT²

Received 2005 October 3; accepted 2006 May 22

ABSTRACT

Statistical observations of the epoch of reionization using the 21 cm line of neutral hydrogen have the potential to revolutionize our understanding of structure formation and the first luminous objects. However, these observations are complicated by a host of strong foreground sources. Several foreground-removal techniques have been proposed in the literature, and it has been assumed that these would be used in combination to reveal the epoch of reionization (EOR) signal. By studying the characteristic subtraction errors of the proposed foreground-removal techniques, we identify an additional subtraction stage that can further reduce the EOR foreground contamination, and study the interactions between the foreground-removal algorithms. This enables us to outline a comprehensive foreground-removal strategy that incorporates all previously proposed subtraction techniques. Using this foreground-removal framework and the characteristic subtraction errors, we discuss the complementarity of different foreground-removal techniques and the implications for array design and the analysis of EOR data.

Subject headings: cosmology: observations — radio lines: galaxies

1. INTRODUCTION

Highly redshifted 21 cm neutral hydrogen emission from the epoch of reionization (EOR) is a unique cosmological probe, and observations with planned low-frequency radio telescopes could revolutionize our understanding of galaxy and structure formation and the emergence of the first luminous objects.

The potential of 21 cm observations was first recognized by Sunyaev & Zeldovich (1972), and further developed by Scott & Rees (1990), Madau et al. (1997), Tozzi et al. (2000), and Iliev et al. (2002). There are five possible experimental signatures produced by neutral hydrogen during the EOR that can be targeted with low-frequency radio observations: the global-frequency step (Shaver et al. 1999), direct imaging, H I forest (Carilli et al. 2002, 2004b), Stromgren sphere mapping, and statistical observations (see Carilli et al. [2004a] for overview). Of the five experimental signatures, the statistical observations developed by Zaldarriaga et al. (2004), Morales & Hewitt (2004), and Bharadwaj & Ali (2005) offer the most promise and cosmological power, and are being targeted by the Mileura Widefield Array (MWA), the Low Frequency Array (LOFAR), and the Chinese Meter Array (21CMA, formerly PAST).

Unlike the cosmic microwave background (CMB) emission, which is inherently two-dimensional (2D; sky position), the EOR data is three-dimensional (3D) because the redshift of the observed neutral hydrogen emission maps to the line-of-sight distance. This allows us to extend the statistical techniques developed for the CMB to three dimensions, and use power-spectrum statistics to study the EOR. These statistical analysis techniques dramatically increase the sensitivity of first-generation EOR observations, and allow much more detailed studies of the cosmology (Bowman et al. 2006; Furlanetto et al. 2004; Morales 2005). The major remaining question of EOR observations is whether the foreground contamination—which is ~ 5 orders of magnitude brighter than the neutral hydrogen radio emission—can be

removed with the precision needed to reveal the underlying EOR signal.

Several foreground-subtraction techniques have been suggested in the literature. Bright sources can be identified and removed as in CMB and galaxy clustering analyses, but the faint emission of sources below the detection threshold will still overwhelm the weak EOR signal (Di Matteo et al. 2002). Additional foreground contamination can be removed by fitting a spectral model to each pixel in the sky to remove the many faint continuum sources in each line of sight (F. H. Briggs et al. 2005, private communication; Wang et al. 2004). A similar technique proposed by Zaldarriaga et al. (2004) moves to the visibility space (2D FT of image-frequency cube to obtain wavenumbers in sky coordinates and frequency along the third axis) and then fits smooth spectral models for each visibility (Santos et al. 2005). This method should be better at removing emission on larger angular scales, such as continuum emission from our own Milky Way. Morales & Hewitt (2004) introduced a subtraction technique that exploits the difference between the spherical symmetry of the EOR power spectrum and separable-axial symmetry of the foregrounds in the 3D Fourier space, and is particularly well suited for removing very faint contaminants.

Foreground removal has been envisioned as a multistaged process in which increasingly faint contaminants are subtracted in a stepwise fashion. By studying the errors made by proposed foreground-subtraction algorithms, we identify an additional subtraction stage where the average fitting errors of the proposed algorithms are subtracted from the 3D power spectrum. This residual error subtraction step can significantly reduce the residual foreground contamination of the EOR signal, and differs from CMB techniques by relying on the statistics of the errors and separating the residual contamination from the power spectrum instead of the image.

Because the residual error subtraction relies on the statistical characteristics of the subtraction errors, the foreground-removal steps become tightly linked, and we must move from focusing on individual subtraction algorithms to the context of a complete foreground-removal framework. This paper outlines a comprehensive foreground-removal strategy that incorporates all previously proposed subtraction techniques and introduces the new residual

¹ Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, MS-51 P-225, Cambridge, MA 02138-1516.

² MIT Kavli Institute for Astrophysics and Space Research, 77 Massachusetts Avenue, Cambridge, MA 02139.

error subtraction stage. Treating the foreground-removal process as a complete system also allows us to study the interactions (covariance) of the subtraction algorithms and identify the types of foreground contamination each algorithm is best suited for.

Section 2 reviews the properties of the data produced by the first-generation EOR observatories. We then introduce the foreground-removal framework in § 3, along with several detailed examples of the subtraction errors. Sections 4 and 5 then discuss the implications of the foreground-removal framework and how it can be used to improve the design of EOR observatories and foreground-removal algorithms.

2. EXPERIMENTAL DATA

While there are important differences in the data processing requirements of the MWA, LOFAR, and 21CMA (and their data analysis systems are rapidly evolving), all three experiments follow the same basic data-reduction strategy.

All three observatories are composed of thousands of simple detection elements (dual-polarization dipoles for MWA and LOFAR, and single-polarization YAGIs for 21CMA). The signals of the individual detecting elements are then combined with analog and digital systems into “antennas” of tens to hundreds of elements, which are then cross-correlated to produce the visibilities of radio astronomy. These visibilities are the basic observable, and are the spatial Fourier transform of the brightness distribution on the sky at each frequency. The visibilities from each experiment—up to 4 billion per second in the case of the MWA—must then be calibrated and integrated to form one long exposure. The final data product is a visibility cube representing a few hundred hours of observation, which can be Fourier transformed either along the angular dimensions to produce an image cube for mapping, or along the frequency axis to produce the Fourier representation for the power-spectrum analysis.

Going from the raw visibilities produced by the correlators to visibility cubes representing hundreds of hours of observation is a Herculean task, and we do not wish to minimize the effort involved in this stage of the processing. The ionospheric distortion must be corrected using radio adaptive optics, and the time-variable gain, phase, and polarization of each antenna must be precisely calibrated. Going from the raw visibilities to the long integration visibility cube tests and displays the art of experimental radio astronomy. However, in the end all three experiments will produce the same basic data product, and for our purposes we will concentrate on how to process this long integration visibility cube to remove the astrophysical foregrounds and reveal the cosmological EOR signal.

3. FOREGROUND SUBTRACTION FRAMEWORK

Fundamentally, all the proposed foreground subtraction techniques exploit symmetry differences between the foregrounds and the EOR signal, and are targeted at removing different types of foreground contamination. Because the EOR signal is created by a redshifted line source, the observed frequency can be mapped to the line-of-sight distance. This produces a cube of space where we can observe the H I intensity as a function of position. The EOR emission appears as bumps along both the frequency and angular directions, and since space is isotropic (rotationally invariant), leads to a spherical symmetry in the Fourier space (the 3D Fourier Transform of the image cube [Morales & Hewitt 2004]). This contrasts with most of the foreground sources, which either have continuum emission that is very smooth in the frequency direction, such as synchrotron radiation, or emission-line radiation that is not redshifted and thus at set frequencies, such as radio recombination lines from the Milky Way. The

foreground-removal techniques all use the difference between the clumpy-in-all-directions EOR signal and the foregrounds, which are smooth in at least one of the dimensions.

The foreground-subtraction framework can be divided into three stages: bright source removal, spectral fitting, and residual error subtraction, where each step subtracts increasingly faint foreground contamination. The first two steps utilize well-developed radio analysis techniques, and have been previously proposed. The residual error subtraction stage extends the symmetry ideas of Morales & Hewitt (2004) to identifying and removing the average fitting error of the first two stages. In §§ 3.1–3.3, we step through the three subtraction stages. Section 3.4 then shows several examples of how to calculate the characteristics of the subtraction errors, and § 3.5 discusses how the subtraction errors affect the EOR sensitivity.

One subtlety that arises is which sources should be considered foreground in an interferometric observation. Diffuse synchrotron emission from our own galaxy is the single brightest source of radio emission at EOR frequencies. The brightness temperature of the synchrotron emission toward the Galactic poles is several hundred degrees, and dominates the thermal noise of the telescope and receiver system. However, the diffuse Galactic synchrotron emission is so spatially and spectrally smooth that it is not customarily included in discussions of foreground subtraction—the majority of this nearly DC emission is resolved out by interferometric observations. Instead, the Galactic synchrotron contribution is included in sensitivity calculations as the dominant source of system noise (Bowman et al. 2006). Similarly, polarized emission can become a major foreground (Haverkorn et al. 2003) by leaking into the intensity maps. However, since the contamination of the intensity map is through errors in the polarization calibration, this is customarily considered a calibration issue and not foreground subtraction. Following these conventions here, we focus on removing the unpolarized contributions from resolved foreground sources.

3.1. Bright Source Removal

In the first stage, the bright contaminating sources, both astrophysical and man-made, are removed. Because the spatial and frequency response of an array is not a delta function, emission from a bright source will spill over into neighboring pixels and frequency channels. The goal of the first foreground-removal stage is to subtract the contributions from all sources that can contaminate distant locations in the image cube.

The worst of the radio frequency interference (RFI) will be cut out of the data prior to forming the long integration visibility cube. What will remain is a sea of faint transmissions. The easiest way to remove narrowband transmissions is to identify the affected channels (elevated rms) and excise them. Modern polyphase filters have very high dynamic range, so only a few channels will need to be removed for all but the very brightest transmitters. This leads to slices of missing frequency information and complicates the experimental window function, but is very effective if the dynamic range of the polyphase filters is high enough.

Removing astrophysical sources is conceptually similar, but is more difficult because of the lower spatial dynamic range of most radio arrays. Emission from a bright astrophysical source will leak into pixels far from the source position because of the imperfect point-spread function of an array. Thus, we need to use the traditional radio astronomy subtraction technique of removing the sources directly from the visibilities to subtract the array sidelobes along with the central emission.

The signal strength of the EOR is a few mK, so the astrophysical and RFI sources must be subtracted until the sidelobes

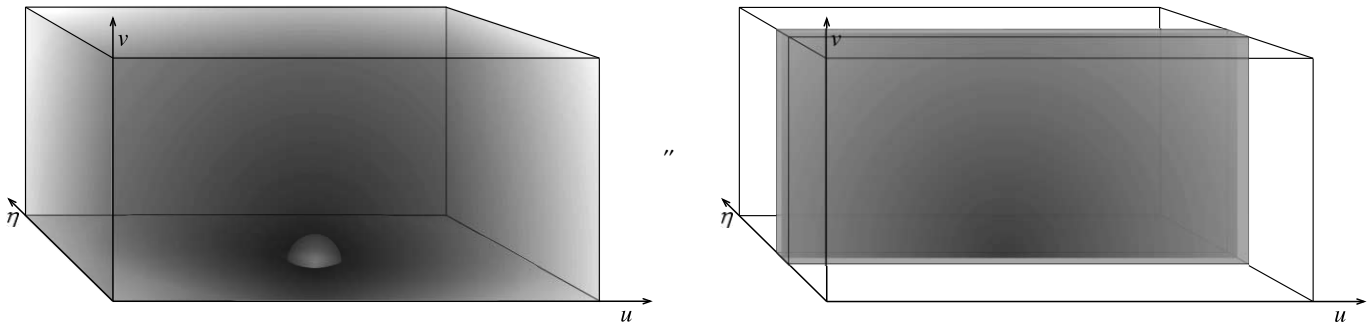


FIG. 1.—*Left*: Spherically symmetric power spectrum of the EOR signal (zero is in the center of the lower face); *right*: separable-axial power-spectrum template typical of the residual foregrounds. The power spectrum shapes are known, and the amplitudes can be fit in the residual error subtraction stage to separate the residual foreground-subtraction errors from the faint EOR signal.

are $\lesssim \text{mK}$. This places strong constraints on the spatial and frequency dynamic range of an array, as well as the RFI environment. Unfortunately, even the faint emission of galaxies below the detection threshold will overwhelm the weak EOR signal (Di Matteo et al. 2002), and we must resort to more powerful subtraction techniques to reveal the EOR signal.

3.2. Spectral Fitting

At the end of the bright source foreground-removal stage, all sources bright enough to corrupt distant areas of the image cube have been removed, and we are left with a cube in which all of the contamination is local. Here we can use foreground-subtraction techniques that target the frequency characteristics of the foreground emission.

In every pixel of the image cube, there will be contributions from many faint radio galaxies. The spectrum within one pixel is well approximated by a power law, and can be fit and removed. Since the EOR signal is bumpy, fitting smooth power-law models nicely removes the foreground contribution, while leaving most of the cosmological signal (F. H. Briggs, et al. 2005, private communication). There are a number of subtle effects that must be carefully monitored, such as changing pixel size, but this is an effective way of removing the contributions of the faint radio galaxy foreground.

Zaldarriaga et al. (2004) proposed a similar method, in which smooth spectral models are fit to individual spatial-frequency pixels in the visibility space. While there is a lot of overlap between this method and the image method, the visibility foreground-subtraction technique should be superior for more extended objects, such as fluctuations in the Milky Way synchrotron emission.

The last type of spectral fitting is to remove radio recombination lines from our own Galaxy. The intensity of these lines is uncertain, but since they occur at known frequencies, template spectra can be used to subtract them. Unlike the smooth power-law spectra, the structure in the recombination line spectrum has much more power on small scales (line-of-sight redshift distance). More work is needed to accurately determine the strength of this foreground and develop template spectra.

The errors made in the spectral-fitting stage can be classified into two types: *model errors* due to foreground spectra that cannot be fit exactly by the model parameters, and *statistical errors* due to slight misestimates of the model parameters in the presence of thermal noise. These errors are discussed at length in § 3.4.

3.3. Residual Error Subtraction

While the vast majority of the foreground contamination will be removed in the first two analysis stages, residual foreground contamination will remain because of errors in the subtraction

process. In the absence of foregrounds, the EOR power spectrum could be measured by dividing the individual power measurements in the Fourier space into spherical annuli, and averaging the values within each shell to produce a single power spectrum measurement at the given length scale (Morales & Hewitt 2004). This reduces the billions of individual power measurements down to of order 10 statistical measurements, and is behind the extraordinary sensitivity of cosmological power spectrum measurements (Bowman et al. 2006).

However, the first two stages in the foreground subtraction are not perfect. For example, in the bright source removal stage, the flux of each source will be slightly misestimated, leading to faint residual positive and negative sources at the locations of the subtracted sources. These faint residual sources inject spurious power to the 3D power spectrum. It is impossible to determine what these subtraction errors are individually (otherwise we would improve them), however, we can predict, measure, and remove the *average* effects of this residual foreground contamination from the power spectrum. Since the power spectrum is related to the square of the intensity, residual positive and negative sources have the same power spectrum signature, and the amplitude of the residual power spectrum signal is related to the standard deviation of the subtraction errors made in the first two stages. Different types of foreground subtraction errors produce distinct shapes in the 3D power spectrum, and are easily differentiated from the approximate spherical symmetry of the EOR signal (see Morales & Hewitt [2004] and § 3.4). Figure 1 shows the 3D power-spectrum shapes typical of the signal and residual foreground components.

So in the presence of foregrounds, our final stage of the analysis becomes a multiparameter fit, with each component of the residual foreground and the EOR signal being represented by a corresponding 3D power spectrum template and amplitude. The measurements are then decomposed into template amplitudes to separate the EOR signal from the residual contamination from foreground-subtraction errors in the first two stages. In effect, this final subtraction stage allows us to fit not only the local foreground parameters (position, spectra, etc.) as in the first two stages, but also the width of the subtraction errors.

The errors produced by the first two foreground-subtraction stages depend on the details of both the algorithm and the array. For example, errors made in the spectral-fitting stage depend on both the spectral model used (quadratic, power law, etc.), and how the pixel shape varies with frequency (array design). This precludes defining a set of residual error templates that is generally applicable, but calculating the templates for a specific analysis is straightforward, as demonstrated in the following section.

3.4. Example Subtraction Error Templates

To separate the residual foreground and cosmological signals in the residual error subtraction stage of the EOR analysis, we need to predict the shape of the residual foreground contamination as seen in the 3D power spectrum. The first two stages of foreground subtraction remove the majority of the contamination, so what we see in the residual error subtraction stage is not the original power spectrum shape of the foregrounds, but instead the shape of the errors characteristic of the first two foreground-removal stages. In the following subsections, we provide examples of how to calculate the residual error templates, and discuss the characteristic power spectrum shapes.

3.4.1. Statistical Spectral Fitting Errors

In the spectral-fitting foreground-subtraction stage, a smooth spectral model is fit to each pixel to remove the contributions of faint continuum sources. However, due to the presence of thermal noise, the fit spectrum is not exactly the same as the true foreground. These slight misestimates of the foreground spectra in each pixel produce a characteristic power spectrum component. The exact shape of this power spectrum template of course depends on the spectral model one chooses. Over the relatively modest bandwidths of proposed EOR measurements, the foreground emission is reasonably well modeled by a quadratic spectrum (F. H. Briggs et al. 2005, private communication), and as an illustrative example we demonstrate how to calculate the power-spectrum template in the case of a simple quadratic spectral model.

For a quadratic foreground-subtraction algorithm, the residual foreground contamination is given by

$$\Delta S(f) = \Delta a df^2 + \Delta b df + \Delta c, \quad (1)$$

where df is the difference between the observed frequency and the center of the band, and Δa , Δb , and Δc represent the difference between the true parameter value for the foreground and the fit value. Figure 2 depicts errors in fitting parameter b for one pixel.

Moving to the line-of-sight wavenumber η with a finite Fourier transform gives

$$\Delta S(\theta_x, \theta_y, \eta) = \frac{\Delta a B}{\pi \eta^2} - \frac{i \Delta b B}{\pi \eta} + \Delta c \delta^K(\eta), \quad (2)$$

where B is the bandwidth of the observation, and we have explicitly shown this for a particular line of sight θ_x, θ_y .³

To compare with the three-dimensional EOR power spectrum, we need to move from the angular coordinates θ_x, θ_y to the spatial frequencies u, v . Even though the foreground is spatially clustered, we expect the fitting errors to be distributed like white noise. This allows us to calculate the Fourier transform for the zero-frequency term and generalize to the other values of u and v :

$$\Delta S(u, v, \eta) = \sum_{\Theta} \Delta S(\theta_x, \theta_y, \eta) d\theta_x d\theta_y, \quad (3)$$

where Θ is the field of view and $d\theta_x d\theta_y = d\Omega$ is the angular resolution of the array (measured in steradians). The rms of the

³ Equation (2) assumes the function is sampled at the center of each bin, whereas in most measurements the value is integrated over the bin. The two definitions converge as $\Delta a \rightarrow 0$, so we use eq. (2) as a very good approximation.

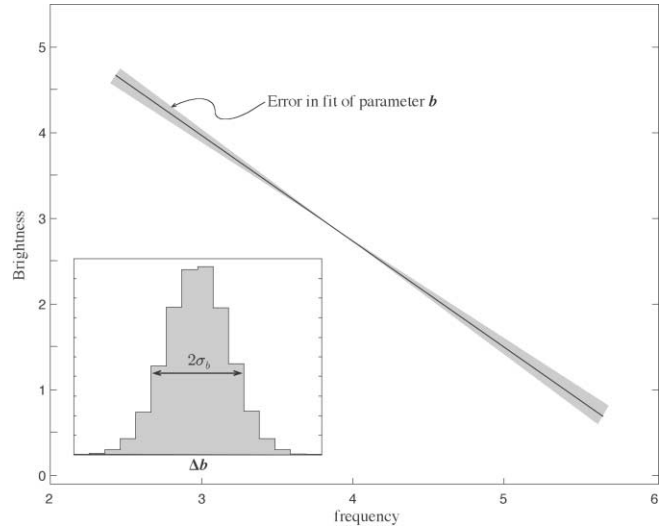


FIG. 2.—Cartoon showing the true foreground continuum spectrum observed in one pixel as a black line, and the error in fitting parameter b due to thermal noise. The inset shows the expected Gaussian profile of $\Delta b \equiv (b_T - b)$ and the width of the distribution σ_b .

sum is given by the square root of the number of independent lines of sight $[(\Theta/d\Omega)^{1/2}]$ times $d\Omega$:

$$\Delta S(u, v, \eta)_{\text{rms}} = \Delta S(\theta_x, \theta_y, \eta)_{\text{rms}} \sqrt{\Theta d\Omega}. \quad (4)$$

If the errors in the fitting parameters are Gaussian distributed, the average power spectrum will be

$$\langle P_s(\eta) \rangle = 2\Theta d\Omega B^2 \left[\frac{\sigma_a^2}{\pi^2 \eta^4} + \frac{\sigma_b^2}{\pi^2 \eta^2} + \sigma_c^2 \delta^K(\eta) \right], \quad (5)$$

where σ_a and σ_b are the standard deviations of Δa and Δb , respectively, and the term σ_c has been redefined to include all the contributions proportional to the Kronecker delta function $\delta^K(\eta)$. For each visibility u, v , the power spectrum will be exponentially distributed around the average, but will become Gaussian distributed when many lines of sight are averaged together.

Equation (5) gives the average power spectrum due to statistical fitting errors with a simple quadratic spectral model. A plot of this fitting-error power-spectrum template along the line of sight is shown in Figure 3. For pixel-based fitting algorithms, the thermal noise is nearly constant for all pixels, and thus the magnitude of the residual power spectrum is nearly equal for all visibilities. This provides a particularly simple 3D power-spectrum template that falls as a high power of η and is constant for all visibilities, and can be visualized as a plane of power near $\eta = 0$ that quickly falls away (see the right panel of Fig. 1 for the basic shape). For visibility-based subtraction algorithms tuned for larger scale structure (Zaldarriaga et al. 2004), the thermal noise at large angular scales is much lower because of the central condensation of realistic arrays, so the power will be concentrated near the $\eta = 0$ plane, but reduced in amplitude at small visibilities.

The σ values in equation (5) represent the standard deviation of the quadratic spectral-fitting algorithm and determine the amplitude of the residual subtraction errors to the signal. The σ values represent an ensemble statistic, as compared to the individual fits to each pixel made by the spectral-fitting algorithm.

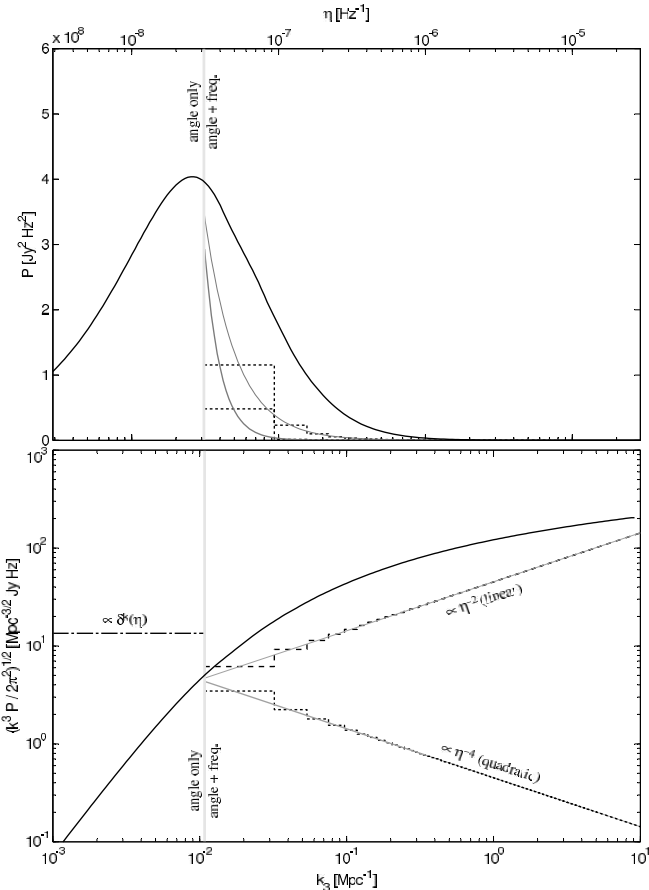


FIG. 3.—Figure showing the shape of the power spectrum contributions from the EOR signal and the statistical spectral fit residuals in observer’s units (*top panel*) and theorist’s units (*bottom panel*) along the line-of-sight direction (η or k_3). The observed EOR signal is shown as a solid black line, and the linear ($\propto \eta^{-2}$) and quadratic ($\propto \eta^{-4}$) residuals are shown as thick and thin gray lines, respectively. The vertical gray line shows the edge of the $\eta = 0$ bin for a 16 MHz bandwidth. All power spectrum measurements to the left of the vertical line are purely angular, while the measurements to the right of the line use both angular and frequency information. The dashed lines show the binning effects for the linear and quadratic components of the residual foreground, with the dash-dotted line showing the δ^k -function contribution from the offset term. The amplitudes of the residual foreground components depend only on the standard deviations of the fitting parameter errors ($\sigma_a, \sigma_b, \sigma_c$), and are fit using parameter estimation in the residual subtraction stage of the analysis.

In the full analysis, we fit local parameters for each pixel/visibility during the spectral-fitting stage, and then we fit the errors (σ) in the residual foreground-subtraction stage. This pattern of fitting local parameters in the first two stages and the statistical distribution of their errors in the third is repeated for each type of error.

The example power spectrum template presented here only applies to a simple quadratic spectral model, but similar results can be obtained if power-law or other foreground models are used instead. The key is to determine the shape of the power spectrum produced by the local spectral-fitting errors, which can then be fit globally in the residual error subtraction stage of the analysis.

3.4.2. Spectral Model Errors

In addition to the statistical fitting errors discussed in the previous subsection, there is a class of model errors that can be made in the spectral-fitting stage. Simple foreground spectral models may be unable to fit the underlying foreground spectrum. Even if all the source spectra were perfect power laws, there are

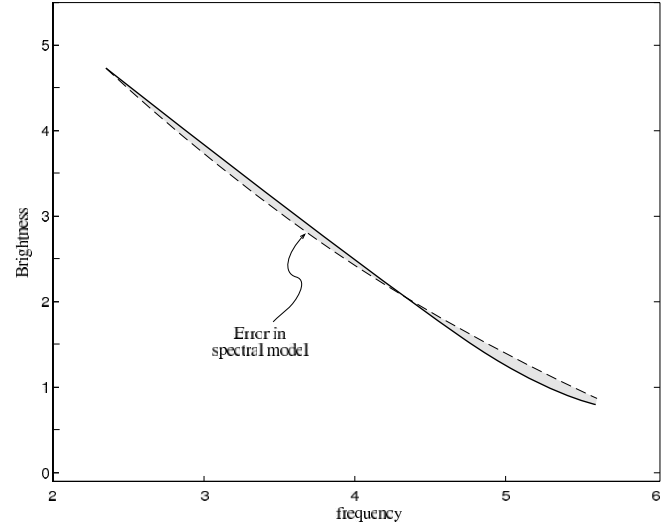


FIG. 4.—Cartoon showing the difference between the spectrum in a pixel (*solid line*) and the best-fit model in the absence of measurement noise (*dashed line*). Because the model cannot exactly follow the true spectrum, an additional component is added to the power spectrum that can be estimated and removed in the residual error subtraction stage of the analysis.

many sources per pixel, which can lead to a complex cumulative spectrum that cannot be exactly fit by a simple spectral model. Figure 4 shows the origin of the model errors.

Because the summation of sources and spectra that forms the cumulative spectrum in each pixel is a statistical process, we expect the model errors to be statistical. This allows one to follow the same process we used in the previous section and form an expected power spectrum shape due to model errors, which can be fit in the residual error subtraction stage of the analysis. This model error template will depend sensitively on both the chosen spectral model and the imaging characteristics of the array.

F. H. Briggs et al. (2005, private communication) showed that using realistic brightness counts dN/dS and spectral slopes, the cumulative spectrum is typically dominated by the brightest source in the pixel, and even for a simple quadratic spectral fit the model error is expected to be less than the cosmological EOR signal. Thus, we expect the power added by model errors to be quite small. But by including the model error power spectrum in the parameter fit, we can eliminate any bias that could be introduced into the EOR signal.

3.4.3. Bright Source Subtraction Errors

Astrophysical sources that are bright enough to contaminate distant areas of the image cube are removed in the bright-source subtraction stage of the analysis (§ 3.1). The errors in this foreground-removal stage are primarily due to misestimates of the source intensities. The bright-source subtraction errors can be envisioned as residual positive and negative sources at the locations of the subtracted sources, and will produce a distinct foreground signature in the 3D power spectrum.

The accuracy of determining the fluxes of foreground sources depends on their brightness—the strongest sources may be subtracted to a few parts per billion, while faint sources are only subtracted to a few tenths of a percent accuracy. Thus the residual sources mirror the positions of subtracted foreground sources and have a zero mean Gaussian distribution of flux, but with a greatly reduced dynamic range (the standard deviation of the subtraction error σ is *not* proportional to the brightness S). This produces a map of the expected residual sources, which can be convolved with the point-spread function of the array and Fourier

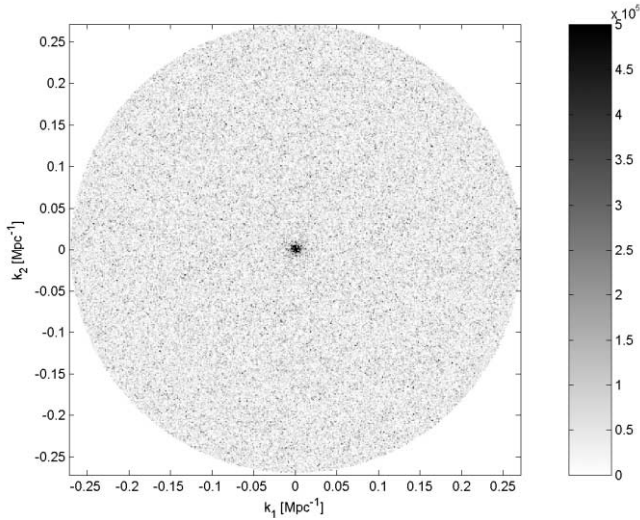


FIG. 5.—Example of the expected u, v power spectrum template due to bright source subtraction errors. In this example, 14,510 sources from the Westerbork Northern Sky Survey (WENSS) at 325 MHz (Rengelink et al. 1997) were chosen in a field centered at 90° R.A., 60° decl., with a 31° FOV (diameter). To model the compression of the dynamic range, the variance of the source subtraction error was chosen to be proportional to the log of the flux, and the color map is linear with an arbitrary scale. Note the increased amplitude toward the center due to the angular galaxy correlation. The exact power spectrum template will depend on the chosen field and the source identification algorithm, but will have this basic shape in the u, v plane and will be highly concentrated toward $\eta = 0$.

transformed and squared to create the expected 3D power-spectrum template for the errors introduced by the bright source subtraction stage of the analysis.

The source-subtraction error template will depend on the particular field observed and spectral model (see § 3.2), but will be strongly peaked at low η [primary error leads to $\delta^K(\eta)$], and concentrated at small visibilities because of angular clustering of the foreground sources (Di Matteo et al. 2002). The compression of the dynamic range must be either modeled or fit with an additional parameter in the residual error subtraction stage of the analysis. An example power-spectrum template is shown in Figure 5, using sources from the Westerbork Northern Sky Survey (WENSS). As can be seen from the example, the errors made by the bright source subtraction stage add power, particularly at small u, v and $\eta = 0$. A similar procedure can be used for RFI, which spills into neighboring frequency channels, and depends on the spectral dynamic range of the array and the RFI removal procedure.

3.5. Uncertainty Calculations

In the final residual foreground-subtraction stage, parameter estimation is used to separate the EOR signal from the residual foreground contaminants, using the characteristic power spectrum templates calculated in § 3.4. We are left with a measurement of the EOR signal strength (typically in ranges of k) and the residual foreground amplitudes. The uncertainty of the EOR signal determination depends on two factors: the uncertainty in the residual foreground amplitude determinations, and their covariance with the EOR signal.

Calculating the uncertainty of the foreground amplitudes ($A_x \propto \sigma_x^2$) is complicated by their additional statistical correlations, as compared to the EOR signal. The observed Fourier space has a fundamental correlation, imprinted by the field of view and bandwidth of the observation. This can be used to define a natural cell size for the data, where the observed EOR signal in each cell is largely independent (Morales & Hewitt 2004; Morales 2005). The observed data can then be represented as a vector of length $N_u \times N_v \times N_\eta$, where N is the number of cells in the u, v , and η dimensions, respectively. For the EOR signal vector \mathbf{s} , all of the cells are nearly independent and $\langle \mathbf{s}\mathbf{s} \rangle \approx \mathbf{s}\mathbf{I}$ where \mathbf{I} is the identity matrix.

This is *not* true for most of the residual foreground templates. The residual foreground contributions are often highly correlated between Fourier cells ($\langle \mathbf{f}\mathbf{f} \rangle \neq \mathbf{f}\mathbf{I}$), and so average differently than the EOR signal. For example, the linear and quadratic spectral fitting errors (σ_a and σ_b terms) each imprint a specific residual in all the η -channels for a given u, v pixel—the amplitude of the residual will vary from one u, v pixel to the next, but are perfectly correlated (deterministic) for the η -values within one pixel.

When averaging over many Fourier cells and lines of sight, the uncertainty in the amplitude of any component becomes approximately Gaussian distributed and equal to

$$\Delta A_x \equiv A_x^T - A_x \propto \sqrt{A_x/N_m}, \quad (6)$$

where A_x is the amplitude of the x component of the signal or residual foreground contribution, T is the true value, and N_m is the number of independent measurements (realizations) of this contribution. For the example of the σ_a and σ_b terms of the spectral fitting residual errors, there are only $N_m = N_u \times N_v$ independent measurements as compared to $N_u \times N_v \times N_\eta$ for the EOR signal and the thermal noise. This correlation along the η -axis comes from using the frequency channels to make the original spectral fit, and means that these spectral errors will only average down by adding more lines of sight, not increasing the

TABLE 1
FUNCTIONAL DEPENDENCE AND STATISTICAL CORRELATIONS OF SIGNAL AND FOREGROUND COMPONENTS IN FOURIER SPACE

Parameter	Functional Dependence	Statistical Correlations	N_m
EOR signal.....	\mathbf{k}	...	$N_u \times N_v \times N_\eta$
Thermal noise.....	$\sqrt{u^2 + v^2}$...	$N_u \times N_v \times N_\eta$
Spectral fitting (σ_a, σ_b).....	η	η	$N_u \times N_v$
Spectral fitting ($\sigma_{c'}$).....	$\delta^K(\eta)$...	$N_u \times N_v$
Spectral model.....	η	η	$N_u \times N_v$
Source subtraction.....	u, v, η	u, v, η	$N_{\text{sources}} \times N_{\text{spectral parameters}}$

NOTES.—Table shows the functional dependence and statistical correlations of the signal and foreground components in the Fourier space (\mathbf{k} or u, v, η), and the number of independent measurements of the component's amplitude N_m . The observed Fourier space has a fundamental correlation imprinted by the field of view and bandwidth of the observation, which can be used to define a natural cell size for the data set where the observed EOR signal and thermal noise in each cell is independent (Morales & Hewitt 2004; Morales 2005). However, the residual foreground components have additional statistical correlations between cells, as indicated in the table and described in the text.

number of cells along the η -axis. The correlations of the various foregrounds are listed in Table 1.

In addition to the uncertainties in the amplitude determinations, we must also calculate the covariance of the parameters. Since the power spectrum templates of the residual fitting errors and signal form the basis functions for the parameter estimation in the residual error subtraction stage, they define the covariance of the amplitude terms. Thus, for a given observatory and choice of foreground-subtraction algorithms, we can calculate the residual foreground power-spectrum templates and resulting uncertainty in the EOR measurement.

4. IMPLICATIONS FOR ARRAY AND ALGORITHM DESIGN

Since the power spectrum templates of the subtraction errors quantify the interactions between the three analysis stages, they enable us to study the effects of array design on our ability to isolate the statistical EOR signal. The difference in the shapes of the power spectrum templates determines how easy it is for the parameter estimation stage of the analysis to separate different contributions. If the power spectrum shapes of two contributions are similar, it will be difficult to accurately determine the amplitudes of the contributions. Mathematically, the shapes of the power spectrum templates determine the covariance of the parameter estimation matrix, with the covariance decreasing as the shapes become more orthogonal.

The power spectrum templates depend on the details of both the array design and analysis technique. For example, the model-fitting error template depends on both the angular resolution of the array (detailed pixel shape) and whether a quadratic or logarithmic power law is used in the fit. The performance advantages and tradeoffs of different arrays and analysis techniques are captured in the shapes and covariances of the power spectrum templates.

To date, the experimental community has been uncertain how to best design arrays and analysis systems to detect the statistical EOR signal. Much of this confusion is because no quantitative measure has been available for comparing design choices. We feel that the power spectrum templates can provide the necessary figure of merit. The shape of the EOR power spectrum signal is known (given a theoretical model), and so we are concerned with the amplitude and covariance of the foreground-subtraction errors of a given array and analysis system with the known EOR power spectrum. The power spectrum templates define the performance of the array and analysis, and allow design tradeoffs to be accurately compared. In many cases, making a plot analogous

to Figure 3 and comparing the amplitude and shape of the residual templates will be sufficient to guide the array design.

5. TOWARD PRECISION FOREGROUND CALCULATIONS

The residual foreground-contamination levels shown in Figure 3 are unrealistically small for the first generation of EOR observatories. However, the EOR signal can still be detected even if the amplitude of the residual foregrounds greatly exceeds the EOR signal in the figure. The key question does not concern the amplitudes of the residual foregrounds, but the uncertainties they create in measuring the EOR signal, as discussed in § 3.5. The foregrounds shown in Figure 3 are what would be needed to detect the EOR power spectrum in a single pixel, whereas all of the first-generation EOR observatories rely on combining information from many lines of sight.

Unfortunately, the uncertainty due to foreground contamination depends strongly on characteristics of the array and observing strategy, for example, a widefield observation will average over more lines of sight, and thus be able to tolerate higher standard deviation in the subtraction than a narrowfield observation. This precludes defining a set of amplitudes that must be obtained to observe the EOR signal for a generic observatory.

The dependence of the subtraction precision on the details of the array makes the task of foreground modelers much more difficult. The precision of the foreground removal is now array dependent, and most researchers are not familiar with the subtle array details needed to accurately calculate the sensitivity of a given observation.

However, the power spectrum templates and σ values do offer a way of translating the characteristics of a foreground-removal algorithm into the sensitivity of an array. Modelers can determine the shape of the residual foreground contamination (as in eq. [5]) and the scaling of the σ values for their foreground-removal algorithm. Experimentalists can then use the predicted power spectrum shapes and σ scalings to determine the effects of the foreground subtraction for a specific observation. This allows researchers studying the foreground removal to avoid doing detailed calculations for each array and observing strategy, while still providing robust results that can guide the experimental design of the next-generation EOR observations. In this way, we hope the foreground-removal framework presented in this paper will facilitate a conversation between foreground modelers and experimentalists and enable accurate array-specific predictions of the foreground subtraction effects on the upcoming EOR observations.

REFERENCES

- Bharadwaj, S., & Ali, S. K. S. 2005, *MNRAS*, 356, 1519
 Bowman, J. D., Morales, M. F., & Hewitt, J. N. 2006, *ApJ*, 638, 20
 Carilli, C. L., Furlanetto, S., Briggs, F., Jarvis, M., Rawlings, S., & Falcke, H. 2004a, *NewA Rev.*, 48, 1029
 Carilli, C. L., Gnedin, N., Furlanetto, S., & Owen, F. 2004b, *NewA Rev.*, 48, 1053
 Carilli, C. L., Gnedin, N. Y., & Owen, F. 2002, *ApJ*, 577, 22
 Di Matteo, T., Perna, R., Abel, T., & Rees, M. J. 2002, *ApJ*, 564, 576
 Furlanetto, S. R., Zaldarriaga, M., & Hernquist, L. 2004, *ApJ*, 613, 16
 Haverkorn, M., Katgert, P., & de Bruyn, A. G. 2003, *A&A*, 403, 1045
 Iliev, I. T., Shapiro, P. R., Ferrara, A., & Martel, H. 2002, *ApJ*, 572, L123
 Madau, P., Meiksin, A., & Rees, M. J. 1997, *ApJ*, 475, 429
 Morales, M. F. 2005, *ApJ*, 619, 678
 Morales, M. F., & Hewitt, J. 2004, *ApJ*, 615, 7
 Rengelink, R. B., Tang, Y., de Bruyn, A. G., Miley, G. K., Bremer, M. N., Roettgering, H. J. A., & Bremer, M. A. R. 1997, *A&AS*, 124, 259
 Santos, M. G., Cooray, A., & Knox, L. 2005, *ApJ*, 625, 575
 Scott, D., & Rees, M. J. 1990, *MNRAS*, 247, 510
 Shaver, P. A., Windhorst, R. A., Madau, P., & de Bruyn, A. G. 1999, *A&A*, 345, 380
 Sunyaev, R. A., & Zeldovich, Ya. B. 1972, *A&A*, 20, 189
 Tozzi, P., Madau, P., Meiksin, A., & Rees, M. J. 2000, *ApJ*, 528, 597
 Wang, X., Tegmark, M., Santos, M., & Knox, L. 2005, preprint (astro-ph/0501081)
 Zaldarriaga, M., Furlanetto, S. R., & Hernquist, L. 2004, *ApJ*, 608, 622