

FAST HOLOGRAPHIC DECONVOLUTION: A NEW TECHNIQUE FOR PRECISION RADIO INTERFEROMETRY

I. S. SULLIVAN¹, M. F. MORALES¹, B. J. HAZELTON¹, W. ARCUS², D. BARNES³, G. BERNARDI⁴, F. H. BRIGGS⁵, J. D. BOWMAN⁶, J. D. BUNTON⁷, R. J. CAPPALLO⁸, B. E. COREY⁸, A. DESHPANDE⁹, L. DESOUSA^{7,10}, D. EMRICH², B. M. GAENSLER^{10,11}, R. GOEKE¹², L. J. GREENHILL⁴, D. HERNE², J. N. HEWITT¹², M. JOHNSTON-HOLLITT¹³, D. L. KAPLAN¹⁴, J. C. KASPER⁴, B. B. KINCAID⁸, R. KOENIG⁷, E. KRATZENBERG⁸, C. J. LONSDALE⁸, M. J. LYNCH², S. R. MCWHIRTER⁸, D. A. MITCHELL¹¹, E. MORGAN¹², D. OBEROI⁸, S. M. ORD⁴, J. PATHIKULANGARA⁷, T. PRABU⁹, R. A. REMILLARD¹², A. E. E. ROGERS⁸, A. ROSHI⁹, J. E. SALAH⁸, R. J. SAULT¹⁵, N. UDAYA SHANKAR⁹, K. S. SRIVANI⁹, J. STEVENS^{7,16}, R. SUBRAHMANYAN⁹, S. J. TINGAY^{2,11}, R. B. WAYTH², M. WATERSON², R. L. WEBSTER^{11,15}, A. R. WHITNEY⁸, A. WILLIAMS¹⁷, C. L. WILLIAMS¹², AND J. S. B. WYITHE^{11,15}

¹ Department of Physics, University of Washington, Seattle, USA

² International Centre for Radio Astronomy Research, Curtin University, Australia

³ Centre for Astrophysics and Supercomputing, Swinburne University of Technology, Melbourne, Australia

⁴ Harvard-Smithsonian Center for Astrophysics, Cambridge, USA

⁵ Research School of Astronomy and Astrophysics, The Australian National University, Canberra, Australia

⁶ School of Earth and Science Exploration, Arizona State University, Tempe, USA

⁷ CSIRO Astronomy and Space Science, Australia

⁸ MIT Haystack Observatory, Westford, USA

⁹ Raman Research Institute, Bangalore, India

¹⁰ School of Physics, University of Sydney, Sydney, Australia

¹¹ ARC Centre of Excellence for All-sky Astrophysics (CAASTRO), Australia

¹² MIT Kavli Institute for Astrophysics and Space Research, Cambridge, USA

¹³ School of Chemical and Physical Sciences, Victoria University of Wellington, New Zealand

¹⁴ Department of Physics, University of Wisconsin–Milwaukee, Milwaukee, USA

¹⁵ School of Physics, The University of Melbourne, Melbourne, Australia

¹⁶ School of Mathematics and Physics, University of Tasmania, Hobart, Australia

¹⁷ Perth Observatory, Perth, Australia, and the University of Western Australia

Received 2012 April 9; accepted 2012 September 9; published 2012 October 11

ABSTRACT

We introduce the Fast Holographic Deconvolution method for analyzing interferometric radio data. Our new method is an extension of A-projection/software-holography/forward modeling analysis techniques and shares their precision deconvolution and wide-field polarimetry, while being significantly faster than current implementations that use full direction-dependent antenna gains. Using data from the MWA 32 antenna prototype, we demonstrate the effectiveness and precision of our new algorithm. Fast Holographic Deconvolution may be particularly important for upcoming 21 cm cosmology observations of the Epoch of Reionization and Dark Energy where foreground subtraction is intimately related to the precision of the data reduction.

Key words: dark ages, reionization, first stars – techniques: image processing – techniques: interferometric

Online-only material: color figure

1. INTRODUCTION

Interferometric radio arrays are some of the most scientifically productive tools of astrophysics. In particular, by choosing the appropriate antenna configuration, the angular sensitivity of an interferometer can be tuned to the scientific question at hand. However, this flexibility comes at the cost of increased demands on the precision of the data analysis.

Radio interferometry uses an array of antennas to sample the electric field from cosmic sources. Each antenna collects the electric field across its surface, which is then cross-correlated against all of the other antenna measurements to generate pairwise visibilities. Each visibility is a measurement of an angular Fourier mode of the sky (for an introduction to interferometry, see Rohlfs & Wilson 2000). The sparse sampling of angular Fourier modes allows a unique combination of resolution and angular sensitivity tailored to the science case, at the expense of a complicated point-spread function (PSF). Additionally, each antenna integrates the electric field across the antenna surface—fundamentally erasing information below the antenna scale and imprinting a direction-dependent gain and polarization response to the signal (see chapter 3 of Morales & Wyithe 2010 for a conceptual introduction to direction-dependent an-

tenna effects). Recent advances by Bhatnagar et al. (2008), Morales & Matejek (2009), and Myers et al. (2003)—variously referred to as “A-projection” or “software holography”—have enabled these direction-dependent gain and polarization effects to be included in the deconvolution process. While the resulting analyses are quite precise, they are slow compared to traditional CLEAN methods (Högbom 1974), particularly if all the antennas do not share the same gain pattern.

In this article we introduce the Fast Holographic Deconvolution (FHD) technique which increases the speed of A-projection/software-holography deconvolution. In Section 2, we introduce the FHD algorithm and compare it to commonly used radio deconvolution algorithms. We then describe our particular implementation of FHD in Section 3, and apply the algorithm to data from the 32 antenna MWA prototype in Section 4. We conclude with a discussion of future work in Section 5.

2. ALGORITHM

There are many variations of radio deconvolution in the literature, including CLEAN (Högbom 1974), Cotton–Schwab faceted CLEAN (Schwab 1984; Cotton et al. 2004; Cotton 2005), multi-scale CLEAN, maximum entropy methods,

A-projection-based CLEAN (Bhatnagar et al. 2008; Carozzi & Woan 2009), peeling (Mitchell et al. 2008), forward-modeling (Bernardi et al. 2011; Pindor et al. 2011), measurement equation (Smirnov 2011a), and compressive sampling (Li et al. 2011), to name just a few. Very crudely, the most popular styles of deconvolution can be divided into two families: basic CLEAN implementations that subtract model visibilities from the data, and peeling, A-projection, and forward modeling approaches that subtract improved visibility estimates from the data with direction-dependent antenna calibration. While this overly broad classification does not do justice to the many important advances in deconvolution, this simplification helps place the advances of FHD in context.

2.1. CLEAN

The basic CLEAN algorithm loop—omitting for clarity wide-field w-projection (Cornwell et al. 2008), faceting, major–minor cycles and other innovations—can be approximated with the following equation to determine the next component of the model:

$$\gamma \text{Max} \left\{ \mathbf{F}^T(\theta, \mathbf{u}) \mathbf{S}(\mathbf{u}, \mathbf{v}) \left(\mathbf{G}(\mathbf{v}, \mathbf{v}) \mathbf{v}_d - \sum_m \mathbf{v}_m \right) \right\}, \quad (1)$$

where we use the linear algebra notation $\mathbf{A}(\mathbf{a}, \mathbf{b}; x)$ to signify a transformation from coordinate vector \mathbf{b} to coordinate vector \mathbf{a} that depends on parameters x , and vector \mathbf{u} is the uv plane and \mathbf{v} represents a list of visibilities. Describing Equation (1) in words, the data visibilities \mathbf{v}_d are calibrated (\mathbf{G}) and the model visibilities \mathbf{v}_m for all the previously found components are subtracted. The terms in parentheses are then the residual visibilities which are gridded to the uv plane (\mathbf{S}) and Fourier Transformed (\mathbf{F}^T) to form the residual dirty map. The peak of the map is selected and the loop gain γ is applied to determine the next component of the model and the loop is repeated.

The model visibilities \mathbf{v}_m are calculated by Fourier transforming the model source components $I(\theta_m)$ to the uv plane and taking the δ -function of the correlation function at the baseline separations, as shown in this annotated and expanded version

$$\gamma \text{Max} \left\{ \mathbf{F}^T(\theta, \mathbf{u}) \overbrace{\mathbf{S}(\mathbf{u}, \mathbf{v})}^{\text{gridding}} \left(\overbrace{\mathbf{G}(\mathbf{v}, \mathbf{v}) \mathbf{v}_d}^{\text{calibrated data}} - \underbrace{\sum_m \delta(\mathbf{v}, \mathbf{u}) \mathbf{F}(\mathbf{u}, \theta) I(\theta_m)}_{\text{model visibilities}} \right) \right\}. \quad (2)$$

This is of course a gross simplification of the algorithms in common use, but it captures the three key features of most CLEAN implementations: the model visibilities are the component correlations at the baseline separation, source components are at pixel centers with a magnitude determined by a CLEAN gain γ that is less than one, and the model is subtracted from the data as visibilities.

2.2. A-projection

A-projection makes two significant changes to the basic CLEAN algorithm (Bhatnagar et al. 2008). A-projection uses the holographic beam pattern ($\tilde{\mathbf{B}}$) to grid residual visibilities to the uv plane, effectively applying direction-dependent calibration,

and it uses the beam pattern (\mathbf{B}) to calculate more accurate model visibilities:

$$\gamma \text{Max} \left\{ \mathbf{F}^T(\theta, \mathbf{u}) \tilde{\mathbf{B}}^T(\mathbf{u}, \mathbf{v}) \left(\mathbf{v}_d - \sum_m \mathbf{B}(\mathbf{v}, \mathbf{u}) \mathbf{F}(\mathbf{u}, \theta) I(\theta_m) \right) \right\}. \quad (3)$$

In words, the model components $I(\theta_m)$ are transformed to the uv plane which is integrated by the known antenna response $\mathbf{B}(\mathbf{v}, \mathbf{u})$ to form an improved estimate of the visibilities. Many array simulators such as MAPS (the MIT Array Performance Simulator¹⁸) operate by performing the uv integration described by $\mathbf{B}(\mathbf{v}, \mathbf{u})$. This improved model visibility is subtracted from the measured data, and the residual visibilities are gridded using the A-projection/software-holography kernels $\tilde{\mathbf{B}}$ to form a holographic dirty map. The peak of the image is used to produce the next model component, and the process is repeated.

In some situations it is faster to apply direction dependent gains ($\mathbf{G}(\mathbf{v}, \mathbf{v}; a, \theta_m)$) to each model visibility rather than integrate the uv plane with $\mathbf{B}(\mathbf{v}, \mathbf{u})$, giving

$$\gamma \text{MaxCentroid} \left\{ \mathbf{F}^T(\theta, \mathbf{u}) \tilde{\mathbf{B}}^T(\mathbf{u}, \mathbf{v}) \left(\mathbf{v}_d - \sum_m \mathbf{G}^T(\mathbf{v}, \mathbf{v}; a, \theta_m) \delta(\mathbf{v}, \mathbf{u}) \mathbf{F}(\mathbf{u}, \theta) I(\theta_m) \right) \right\}. \quad (4)$$

Mathematically these are identical as $\mathbf{B}(\mathbf{v}, \mathbf{u}) = \mathbf{G}(\mathbf{v}, \mathbf{v}; a, \theta_m) \delta(\mathbf{v}, \mathbf{u})$, but they can have very different speeds depending on context. Using the direction dependent gain is sometimes referred to as peeling (Mitchell et al. 2008). This is also similar to the measurement equation approach to direction-dependent calibration developed by Smirnov (2011b, 2011c).

2.3. Forward Modeling

Forward modeling (Bernardi et al. 2011; Pindor et al. 2011) is mathematically similar to A-projection and peeling, but subtracts holographic model and data images instead of the model and data visibilities:

$$\gamma \text{MaxCentroid} \left\{ \left[\mathbf{F}^T(\theta, \mathbf{u}) \tilde{\mathbf{B}}^T(\mathbf{u}, \mathbf{v}) \mathbf{v}_d \right]_1 - \mathbf{F}^T(\theta, \mathbf{u}) \tilde{\mathbf{B}}^T(\mathbf{u}, \mathbf{v}) \mathbf{B}(\mathbf{v}, \mathbf{u}) \sum_m \mathbf{F}(\mathbf{u}, \theta) I(\theta_m) \right\}. \quad (5)$$

The data are gridded and imaged (shown in square brackets) to form a holographic image (first line), and the model is converted to visibilities then gridded and imaged to form a holographic model image (second line), which are then subtracted. For traditional dirty images image subtraction would cause a significant degradation in deconvolution fidelity. However, because the antenna beam pattern ($\tilde{\mathbf{B}}$) is used to grid the data, holographic dirty maps contain all the information that was in the visibilities and subtracting images or visibilities works equally well (assuming calibration $\tilde{\mathbf{B}}$ is correct, Morales & Matejek 2009). This lossless information property was proved by Tegmark (1997) and is widely used for reducing cosmic microwave background data. One advantage of subtracting images is that the data only need to be gridded and imaged once (square brackets). Forward modeling also usually centroids the position of the source components as opposed to using the pixel centers, though component centroiding can be added to any of the deconvolution approaches.

¹⁸ <http://www.haystack.mit.edu/ast/arrays/maps/index.html>

2.4. Fast Holographic Deconvolution

The major change behind FHD is to introduce the Holographic Mapping function (HMF) $\mathbf{H}(\mathbf{u}, \mathbf{u}) \equiv \tilde{\mathbf{B}}^T(\mathbf{u}, \mathbf{v})\mathbf{B}(\mathbf{v}, \mathbf{u})$. In all of the holographic deconvolution algorithms (A-projection (Equation (3)), peeling with holographic gridding (Equation (4)), and forward modeling (Equation (5))), the majority of the time is spent integrating regions of the model uv plane to form the model visibilities ($\mathbf{B}(\mathbf{v}, \mathbf{u})$), and then gridding the visibilities with the holographic antenna beams ($\tilde{\mathbf{B}}^T(\mathbf{u}, \mathbf{v})$). Instead of repeating these operations for every model component and visibility, we combine them into an HMF. Conceptually the HMF $\mathbf{H}(\mathbf{u}, \mathbf{u})$ records *how* a uv location in the model is mapped to the holographic uv map, independent of the sky model used and including all of the baseline sampling and direction-dependent beam effects. Effectively we pre-compute the process of forming and gridding visibilities and use this repeatedly as we refine the sky model. The HMF allows us to quickly convert a uv model into an accurate estimate of the holographic uv map as seen by the instrument, including all of the direction-dependent array beam, antenna beam, and polarization effects. The resulting FHD algorithm is

$$\gamma \text{MaxCentroid} \left\{ \left[\mathbf{F}^T(\theta, \mathbf{u}) \tilde{\mathbf{B}}^T(\mathbf{u}, \mathbf{v}) \mathbf{v}_d \right]_1 - \mathbf{F}^T(\theta, \mathbf{u}) \mathbf{H}(\mathbf{u}, \mathbf{u}) \sum_m \mathbf{F}(\mathbf{u}, \theta) I(\theta_m) \right\}. \quad (6)$$

In words, the source components are used to build an uv model ($\sum_m \mathbf{F}(\mathbf{u}, \theta) I(\theta_m)$), the HMF converts this to a model holographic uv map including all of the instrumental and gridding effects, a Fourier transform creates a model holographic dirty image (lower line) which is subtracted from the dirty holographic map of the data (square brackets). Mathematically, FHD is equivalent to the A-projection, peeling, and forward modeling deconvolution algorithms,¹⁹ but is more efficient since the effect of individual direction-dependent antenna gains is pre-computed outside of the deconvolution loop.

The speed of FHD is due to the sparseness of the HMF. Formally $\mathbf{H}(\mathbf{u}, \mathbf{u})$ is a very large matrix as it can map any model uv point to any point in the holographic uv map. However, every visibility is the integral of a very limited region of the uv plane—about twice the diameter of an antenna. The gridding function is similarly compact. So in practice a location in the model uv plane is mapped only to nearby locations in the holographic uv plane. This combines with the proximity of nearby frequency and time samples to make the Holographic Mapping matrix $\mathbf{H}(\mathbf{u}, \mathbf{u})$ very sparse and amenable to sparse matrix computational methods. Determining the sparse HMF takes about ten times longer than one integration/gridding cycle (Table 1 and the associated text). But after $\mathbf{H}(\mathbf{u}, \mathbf{u})$ has been calculated, we can take any model uv contribution and very quickly map it into a holographic dirty map with high precision.

The power of the FHD technique comes from combining image-based subtraction and the HMF. This enables us to produce a precision deconvolution method, complete with direction and polarization effects, that is faster than current implementations.

¹⁹ Implicit gridding of the model uv plane can introduce small errors far from the field center.

3. IMPLEMENTATION

In this section, we detail our implementation of FHD. The FHD algorithm has three key steps: creating a holographic dirty image of the data, calculating the HMF $\mathbf{H}(\mathbf{u}, \mathbf{u})$, and iterating through the deconvolution loop (Equation (6)).

3.1. The Holographic Dirty Image

We begin FHD by constructing a dirty holographic image of the calibrated data for each of the four instrumental polarizations. We assume a dual linear reception element (labeled X and Y), and create separate holographic dirty maps from the XX, YY, XY, and YX visibilities. The maps are created by gridding each visibility with a kernel based on the individual holographic antenna maps for each polarization, and Fourier transforming the result. It is the holographic gridding kernel that enables the correction of the direction and polarization dependent response of each antenna. Our implementation supports arbitrary antenna kernels and any desired bandwidth synthesis. Since we use image-based subtraction (Equation (6)), we only need to grid the visibilities to the uv plane and fast Fourier transform (FFT) once for each instrumental polarization.

3.2. The Holographic Mapping Function

The HMF $\mathbf{H}(\mathbf{u}, \mathbf{u})$ must be pre-computed for every unique observation and for each instrumental polarization. Once computed, it is stored in the row-indexed sparse storage mode from *Numerical Recipes in C* and written to disk. Despite being very sparse, this is still typically a large file—a few gigabytes for our example in Section 4. However, the size of the matrix scales with uv coverage rather than the number of baselines or the length of integration, so we believe this approach will remain practical even for future large arrays. Finally, we note that the choice of pixelization of the uv plane is quite flexible. For example, the dimensions and resolution of the input and output uv planes do not have to match, or one could be a standard grid and the other non-rectilinear such as an uv analog of Healpix.

3.3. Deconvolution

Once the HMF has been built we begin the deconvolution loop described mathematically by Equation (6). In words, the steps within each iteration are:

1. We begin by applying the HMF to the model sky (in uv space) for each instrumental polarization, and take the FFT to create four model dirty holographic images.
2. These are subtracted from the corresponding dirty holographic images of the data.
3. The residual images are converted to Stokes I , Q , U , and V using the Jones matrices and an average beam model for each instrumental polarization.
4. The Stokes images are median filtered with a small box size to highlight point sources and masked to exclude regions of sky with low instrumental response. (Median filtering is equivalent to down weighting short baselines, but is faster for holographic algorithms.)
5. The positive peak in the filtered Stokes I image is identified. A centroid is fit around the peak and the gain factor γ is used to generate a new source component. (Q , U , and V amplitudes of the component may optionally also be fit.)
6. The new component is used to update the uv plane model sky in all four instrumental polarizations.

Table 1
Computational Costs

Single iteration		Full deconvolution loop	
τ_{model}	123 CPU-s	N	20,000 iterations
τ_{grid}	638 CPU-s	S	9.59
τ_{fit}	1.66 CPU-s	τ_A	4232 CPU-hours
τ_{HMF}	3.21 CPU-s	τ_{FHD}	29.1 CPU-hours

When deconvolution is finished, we condense all components within a threshold radius of half a pixel to form a list of discrete point source candidates. A robust source list may be constructed by applying standard statistical tests to the list of candidates, such as imposing a minimum signal to noise threshold. Extended sources which have more than a negligible number of close components outside of this threshold may be either left in component form or fit with a model profile.

3.4. Comparison of Computational Efficiency

Comparing the computational efficiency of very different algorithms is always difficult due to the details of implementation and computer hardware. Instead of counting the formal number of operations, which tends to ignore important effects such as memory bandwidth, we have created matched implementations of both FHD and A-projection. The IDL programming language was used for both, and no parallelization or optimizations were used (many optimizations work equally well for both). Once we finish algorithm development we intend to translate our FHD implementation to C or CUDA and parallelize as necessary.

The computational cost of deconvolving N components with A-projection can be represented by

$$\tau_A = N(\tau_{\text{model}} + \tau_{\text{grid}} + \tau_{\text{fit}}), \quad (7)$$

where τ_{model} is the time spent constructing model visibilities ($\tilde{\mathbf{B}}^T(\mathbf{u}, \mathbf{v})$), τ_{grid} is the time spent gridding the visibilities using the individual direction-dependent antenna gains to the uv plane ($\mathbf{B}(\mathbf{v}, \mathbf{u})$), and τ_{fit} is the time spent determining the next model source component. Forward modeling (Equation (5)) is computationally similar in cost to A-projection because it requires constructing model visibilities and gridding them back to the uv plane for each model component.

In FHD there is an up-front cost for calculating the HMF, followed by more efficient determination of each model component. The computational cost of FHD is represented by

$$\tau_{\text{FHD}} = S(\tau_{\text{model}} + \tau_{\text{grid}}) + N(\tau_{\text{fit}} + \tau_{\text{HMF}}). \quad (8)$$

The first term is the cost of pre-calculating the HMF, with one iteration of constructing model visibilities and gridding to the uv plane times a computation overhead S associated with building a sparse matrix. The second term is the time required to perform the deconvolution, consisting of the time to apply the HMF τ_{HMF} and the time to identify the next model component τ_{fit} times the number of model components N . The speed advantage of FHD comes from moving the costs involved in modeling and gridding visibilities outside of the deconvolution loop.

Table 1 lists the computational cost of each term in Equations (7) and (8) for our example MWA 32 data (Section 4). The total time for deconvolving 20,000 components with the A-projection algorithm (τ_A) is more than two orders of magnitude longer than with FHD (τ_{FHD}).

These costs do depend on the details of the instrument. In general τ_{model} and τ_{grid} will scale linearly with the number of

visibilities and the size of the gridding kernel, while τ_{HMF} scales linearly with uv coverage and with the square of the size of the gridding kernel. In practice there are many optimizations that can be applied to both algorithms, many of which will help both approaches. For our IDL implementation of FHD we have been able to reduce the deconvolution time (second term of Equation (8)) to 23 minutes per polarization. Written in a more efficient programming language and parallelized, real-time deconvolution of the example in Section 4 should be achievable on a high-end desktop computer.

4. RESULTS

In this section, we analyze a 5 minute snapshot observation by the 32 antenna MWA prototype (Ord et al. 2010) to demonstrate the effectiveness of FHD. Detailed descriptions of the full as-built MWA facility can be found in (Tingay et al. 2012) and (Lonsdale et al. 2009). This observation is from 139 MHz–169 MHz with Pictor A at the center of the primary beam field of view (FOV; approx. $41^\circ \times 41^\circ$ out to the first null), and was taken on 2010 September 25 at 5:19 am local time (21:19 UT) on our 14th field expedition. The Orion and Crab Nebulae are 40.5° and 67.9° to the north of field center in the first and second antenna side lobes, and the Galactic plane extends through the southern side lobes. The results of our FHD algorithm are shown in Figures 1 and 2. No attempt to deconvolve sources in the side lobes of the array beam was made, so artifacts from the bright Crab Nebula are visible in the northern corners of panels (b)–(d).

Panel (a) of Figure 1 shows the holographic dirty image prior to deconvolution, which is dominated by Pictor A and the array beam (PSF). We perform 20k iterations of deconvolution with a clean gain of 15% using only positive, centroided source components and removing all peaks above 600 mJy in the central region of the residual image. This peak detection threshold rises as sensitivity decreases toward the edge of the field to maintain a constant minimum signal to noise. We then collect all components within 6 arcmin to identify 2342 source candidates, shown in panel (b) (because each component is independently centroided, individual components of a real point source will differ slightly in location by small fractions of a pixel). The residual image is shown in panel (c) on an expanded linear scale to highlight the fluctuations (no smoothing or filtering has been applied to any of these images). There are very few features in the residual image, the residuals are highly Gaussian with noise of $\sigma = 230$ mJy in the center, and there is no evidence of the Pictor A array beam. As expected, the noise increases toward the beam edge where the sensitivity is down by a factor of 20. The flux density of Pictor A has been scaled to match the catalog value of 402 Jy to set an approximate overall scale.

The source counts are shown in panel (a) of Figure 2. The black solid line shows the counts from the 6th Cambridge survey at 151 MHz in the northern hemisphere (Hales et al. 1988), overlaid by our source counts in red. Because we deconvolve sources out to near the first null of the tile beam, our sensitivity naturally falls toward the field edge. At the faint end of the distribution (below 3 Jy) we divide our source counts into rings of nearly equal sensitivity (dashed lines). At the bright end (above 3 Jy) we use the full field as the uncertainty becomes dominated by the low source counts. There appears to be good correspondence between the source counts from FHD and the expected counts, with the expected roll-offs in sensitivity in the outer edges of the field. We observe excellent agreement of all bright source candidates with the Molonglo Reference

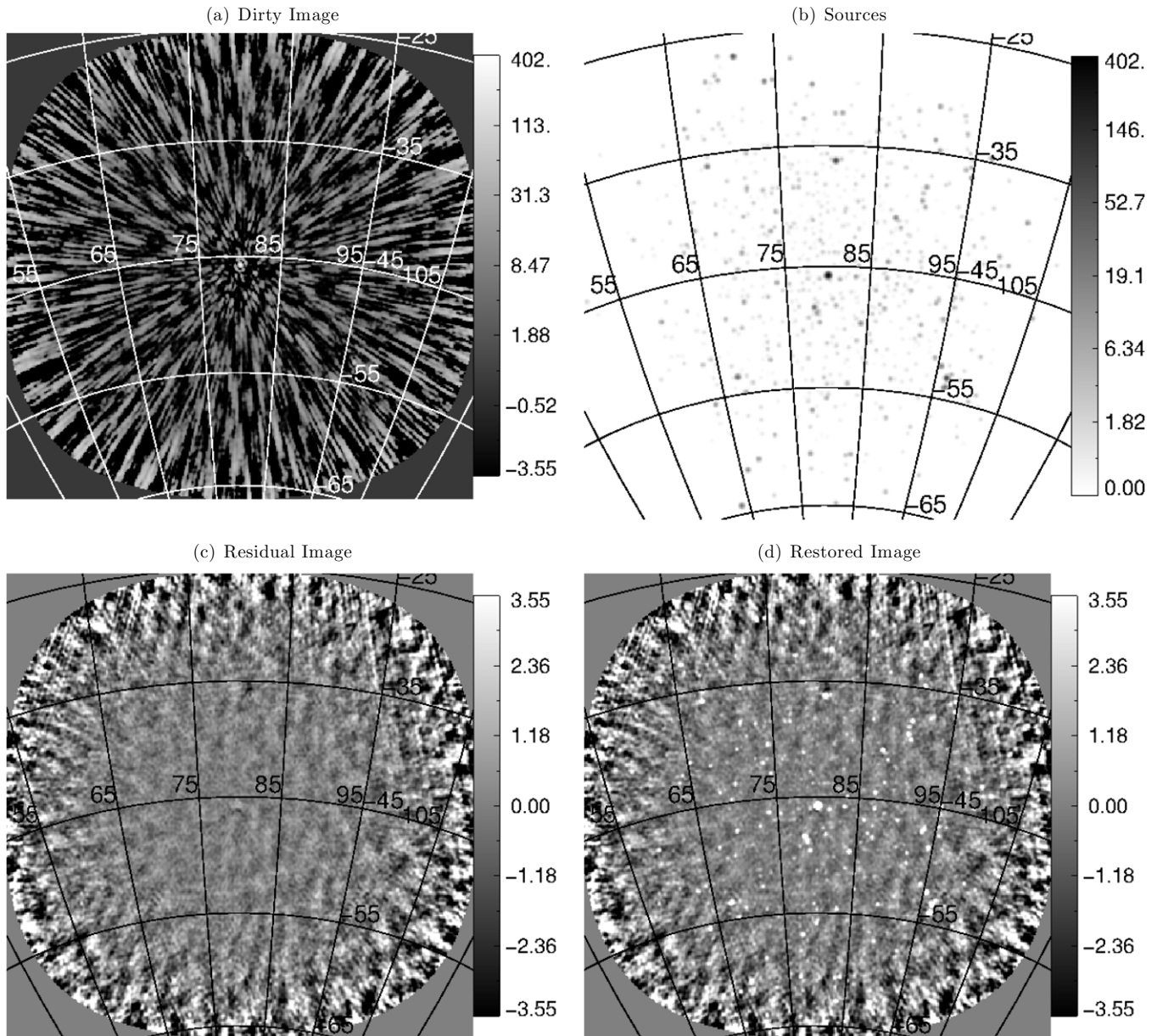


Figure 1. Data from the MWA 32 antenna prototype analyzed with Fast Holographic Deconvolution. The data shown are purposefully a torture test for deconvolution algorithms—a 5 minute snapshot with Pictor A at the center of the $\sim 41^\circ \times 41^\circ$ primary beam FOV, the Orion and Crab Nebulae $40:5$ and $67:9$ to the north of field center (well off the image), and the Galactic plane extending through the antenna side lobes. Panel (a) shows the Stokes I dirty image, dominated by Pictor A and the snapshot PSF (array beam) shown on a log stretch to highlight low level detail. Deconvolution reveals 2342 source candidates in this 5 minute snapshot in panel (b) at upper right (log stretch). Panel (c) shows the residual image after deconvolution on an expanded linear scale to highlight the residual emission, and panel (d) shows the reconstructed image (panels b + c) on the same scale. There are false sources in the upper corners of the image due to bright sources outside of the clean region which are not deconvolved. However, we see no residuals from Pictor A or other sources inside the deconvolved region. See the text for a detailed description of the features and Figure 2 for the $\log N / \log S$ plot and the Stokes Q image.

Catalog (MRC) of radio sources (Large et al. 1981), though we defer a thorough catalog comparison to future work with deeper integrations.

HERA²⁰ instruments such as the 32 antenna MWA prototype quickly reach the confusion limit (expected to be ~ 300 mJy). This makes measuring the dynamic range of a deconvolution algorithm problematic using common measures such as the rms in blank areas of the field. Within 7° of the field center, the source counts appear to follow the 6 C catalog to 600 mJy.

This implies a minimum snapshot dynamic range for FHD of 800:1 in this data set. This dynamic range will naturally improve with synthesis rotation. The polarization performance is represented by the residual Stokes Q image in Figure 2(b), with a fractional polarization of a few percent in the central region. With improved beam models and calibration we expect to achieve percent polarization residuals across the field.

5. CONCLUSION

A-projection and forward modeling deconvolution algorithms that use direction-dependent gain and polarization are very

²⁰ Hydrogen Epoch of Reionization Arrays <http://reionization.org/> (reionization.org)

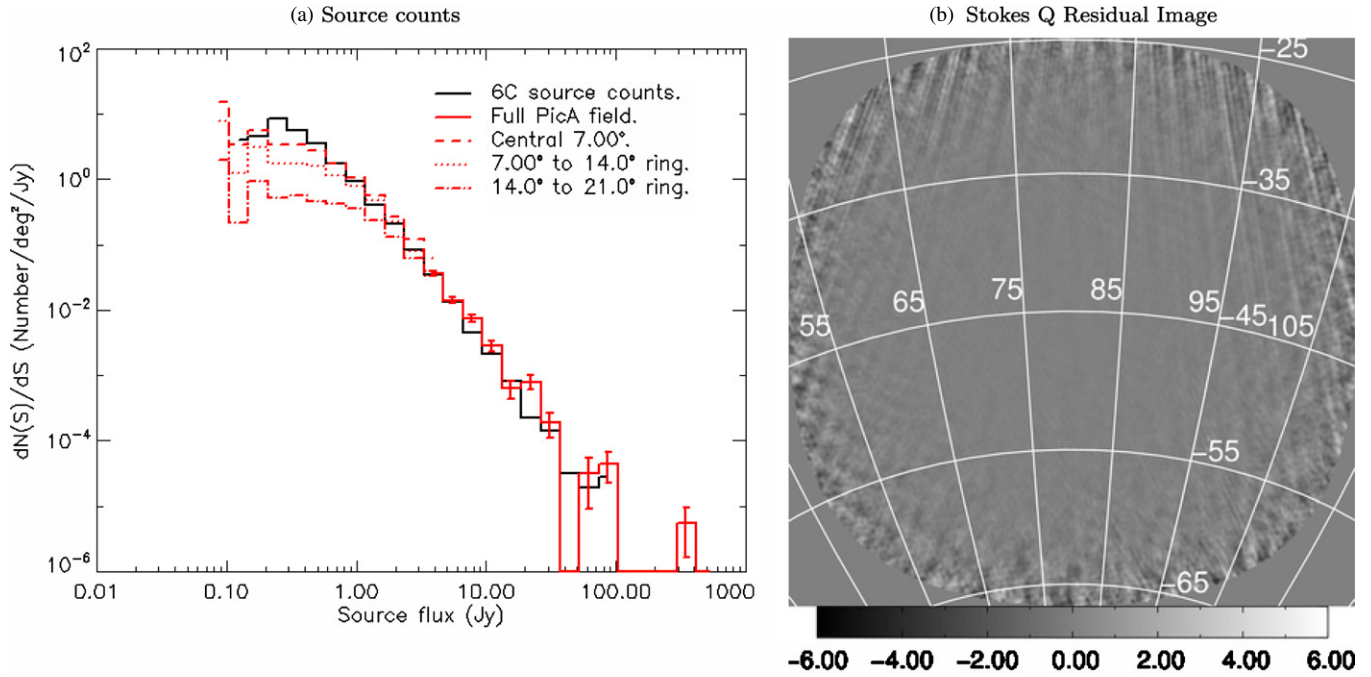


Figure 2. Panel (a) shows the $\log N / \log S$ plot for the sources in Figure 1, with error bars to indicate sample variance at high fluxes. We have divided the source statistics below 3 Jy into rings of increasing radius (dashed, dotted, dash-dotted). Careful investigation of each curve shows the turnover in the completeness increasing with radius as expected. In the central region of the image the sources follow the expected source counts until 600 mJy, which gives us a conservative snapshot dynamic range of $>800 : 1$. This dynamic range will naturally improve with synthesis rotation (dynamic range becomes hard to measure in longer MWA 32 observations due to the confusion limit). Panel (b) shows the residual Stokes Q image, with fractional polarization of a few percent in the central region (21° from antenna bore sight). With upgraded beam models and calibration we expect to achieve percent polarization across the field.

(A color version of this figure is available in the online journal.)

precise, but are much slower than the traditional CLEAN approach. In this paper we have introduced FHD, which uses the HMF to improve the speed of holographic deconvolution without sacrificing precision. We have demonstrated the power of the FHD algorithm using a 5 minute snapshot of data from the 32 tile MWA prototype.

In future work we plan to enhance the capabilities of our FHD algorithm; including parallelizing and optimizing our implementation, deconvolving diffuse sources, and implementing w-projection to enable long wide-field integrations (Cornwell et al. 2008; using holographic formulation in Morales & Matejek 2009). Because w-projection decreases the sparseness of the HMF (larger gridding kernels) it may be more efficient to do a joint deconvolution of a few intermediate length observations than one long track. We will also explore ways of using sparse matrix parallelization in addition to the more traditional frequency or time parallelizations. Finally, FHD lends itself naturally to modeling diffuse sources such as the galaxy, and we plan to explore multi-scale or other principle component deconvolution methods.

Support for this work came from the U.S. National Science Foundation (grants AST CAREER-0847753, AST-1003314, AST-0457585, AST-1008353, AST-0908884, and PHY-0835713), the Australian Research Council (grants LE0775621 and LE0882938), the U.S. Air Force Office of Scientific Research (grant FA9550-0510247), the Smithsonian Astrophysical Observatory, the MIT School of Science, the Raman Research Institute, the Australian National University, the iVEC Petabyte Data Store, the Initiative in Innovative Computing and NVIDIA sponsored Center for Excellence at Harvard, and the International Centre for Radio Astronomy Research, a Joint Venture of Curtin University of Technology and The University

of Western Australia, funded by the Western Australian State government.

We acknowledge the Wajarri Yamatji people as the traditional owners of the Observatory site.

REFERENCES

- Bernardi, G., Mitchell, D. A., Ord, S. M., et al. 2011, *MNRAS*, **413**, 411
- Bhatnagar, S., Cornwell, T. J., Golap, K., & Uson, J. M. 2008, *A&A*, **487**, 419
- Carozzi, T. D., & Woan, G. 2009, *MNRAS*, **395**, 1558
- Cornwell, T. J., Golap, K., & Bhatnagar, S. 2008, *IEEE J. Sel. Top. Signal Process.*, **2**, 647
- Cotton, W. 2005, in ASP Conf. Ser. 345, From Clark Lake to the Long Wavelength Array, ed. N. Kassim, M. Perez, M. Junor, & P. Henning (San Francisco, CA: ASP) **337**
- Cotton, W. D., Condon, J. J., Perley, R. A., et al. 2004, *Proc. SPIE*, **5489**, 180
- Hales, S., Baldwin, J., & Warner, P. 1988, *MNRAS*, **234**, 919
- Högbom, J. 1974, *A&AS*, **15**, 417
- Large, M. I., Mills, B. Y., Little, A. G., Crawford, D. F., & Sutton, J. M. 1981, *MNRAS*, **194**, 693
- Li, F., Cornwell, T. J., & de Hoog, F. 2011, *A&A*, **528**, 31
- Lonsdale, C. J., Cappallo, R. J., Morales, M. F., et al. 2009, *Proc. IEEE*, **97**, 1497
- Mitchell, D., Greenhill, L., Wayth, R., et al. 2008, *IEEE J. Sel. Top. Signal Process.*, **2**, 707
- Morales, M., & Wyithe, J. 2010, *ARA&A*, **48**, 127
- Morales, M. F., & Matejek, M. 2009, *MNRAS*, **400**, 1814
- Myers, S., Contaldi, C., Bond, J., et al. 2003, *ApJ*, **591**, 575
- Ord, S. M., Mitchell, D. A., Wayth, R. B., et al. 2010, *PASP*, **122**, 1353
- Pindor, B., Wyithe, J. S. B., Mitchell, D. A., et al. 2011, *PASA*, **28**, 46
- Rohlfs, K., & Wilson, T. L. 2000, *Tools of Radio Astronomy* (New York: Springer)
- Schwab, F. R. 1984, *AJ*, **89**, 1076
- Smirnov, O. M. 2011a, *A&A*, **527**, 106
- Smirnov, O. M. 2011b, *A&A*, **527**, 107
- Smirnov, O. M. 2011c, *A&A*, **527**, 108
- Tegmark, M. 1997, *ApJ*, **480**, L87
- Tingay, S. J., Goeke, R., Bowman, J. D., et al. 2012, *PASA*, in press (arXiv:1206.6945)