WIDE-FIELD WIDE-BAND INTERFEROMETRIC IMAGING: THE WB A-PROJECTION AND HYBRID ALGORITHMS

S. BHATNAGAR, U. RAU, AND K. GOLAP
National Radio Astronomy Observatory, Socorro, NM 87801, USA; sbhatnag@nrao.edu, rurvashi@nrao.edu, kgolap@nrao.edu

Received 2012 August 8; accepted 2013 April 17; published 2013 May 28

ABSTRACT

Variations of the antenna primary beam (PB) pattern as a function of time, frequency, and polarization form one of the dominant direction-dependent effects at most radio frequency bands. These gains may also vary from antenna to antenna. The A-Projection algorithm, published earlier, accounts for the effects of the narrow-band antenna PB in full polarization. In this paper, we present the wide-band A-Projection algorithm (WB A-Projection) to include the effects of wide bandwidth in the A-term itself and show that the resulting algorithm simultaneously corrects for the time, frequency, and polarization dependence of the PB. We discuss the combination of the WB A-Projection and the multi-term multi-frequency synthesis (MT-MFS) algorithm for simultaneous mapping of the sky brightness distribution and the spectral index distribution across a wide field of view. We also discuss the use of the narrow-band A-Projection algorithm in hybrid imaging schemes that account for the frequency dependence of the PB in the image domain.

Key words: methods: data analysis – techniques: image processing – techniques: interferometric

Online-only material: color figures

1. INTRODUCTION

Observations in the radio band offer distinct, and often times unique, scientific advantages in probing certain areas of astrophysical research (e.g., in the detection of the EoR signal, studies of the high-redshift universe in general, large-scale structure formation, early galaxies, etc.).

All next-generation radio telescopes, many in operation now, offer at least an order of magnitude improvement in the sensitivity and angular resolution compared to the telescopes operated in the past decades. The two key instrumental parameters that afford such high sensitivities, impact the imaging performance, and are significantly different from previous-generation telescopes are: (1) the wide instantaneous fractional bandwidths, and (2) larger collecting area. The effects of wide instantaneous fractional bandwidths that classical calibration and imaging approaches to full-beam, wide-band continuum imaging. We present a modification of the A-Projection algorithm, which we call the wide-band A-Projection, or WB A-Projection algorithm, where a modified A-term also compensates, to a large extent, the frequency dependence of the PB. We also discuss the use of the unmodified A-Projection algorithm (Bhatnagar et al. 2008, henceforth referred to as Paper I), which we call the narrow-band A-Projection, or NB A-Projection, along with various forms of image-plane normalizations for wide-band continuum imaging and the resulting issues and limitations.

2. THEORY

Using the notation developed by Hamaker et al. (1996), full polarimetric measurements from a single baseline calibrated for the effects of direction-independent (DI) gains, can be described by the following measurement equation:

\[ V_{ij}^{\text{Obs}}(v,t) = W_{ij}(v,t) \int P_{ij}^{\text{Sky}}(s,v,t)I(s,v)e^{ib_{ij}s} ds, \]  

(1)

where \( V_{ij}^{\text{Obs}} \) are the observed visibility samples measured by the pair of antennas designated by the subscript \( i \) and \( j \), separated by the vector \( b_{ij} \) and weighted by the measurement weights \( W_{ij} \). \( P_{ij}^{\text{Sky}} \) is the radio-Mueller matrix in the image domain representing the full polarization description of the antenna PB as a function of the direction \( s \), frequency \( v \), and time \( t \), and \( I \) is the image vector. The vectors \( V \) and \( I \) are full polarization vectors in the data and image domain, respectively. \( P_{ij}^{\text{Sky}} \) and \( I \) are the unknowns in this equation.

Equation (1) cannot be directly inverted as, in general, it is not a Fourier transform relation. It is also sampled only at a limited number of points, and therefore the data has insufficient information to allow an exact solution. Estimation of \( I \) is therefore typically done via iterative nonlinear \( \chi^2 \)-minimization.

1 This matrix as used in radio interferometric literature differs from that used in the optical literature only in that in radio it is written in the polarization basis (circular or linear polarization) while in the optical literature it is written in the Stokes basis. These radio and optical representations are related via a Unitary transform (Hamaker et al. 1996).
(Cornwell 1995; Rau et al. 2009). Below we briefly review the theory of imaging with A-Projection to correct for the time and polarization dependence of $p_{ij}^{\text{sky}}$ in narrow-band imaging and motivate the need for a WB A-Projection algorithm to also correct for frequency dependence of $p_{ij}^{\text{sky}}$ in wide-band imaging.

2.1. Imaging with A-Projection

To clarify the full-polarization nature of the A-Projection algorithm, we define the outer-convolution operator and denote it by the symbol $\otimes$. The outer-convolution operator is similar to the outer-product operation used in the DI description of Hamaker et al. (1996) with a minor difference. The element-by-element algebra of the outer-convolution operator is the same as that of the outer-product operator, except that the complex multiplications in outer-product operator are replaced by convolutions. Using the outer-convolution operator and the subscripts $i$ and $j$ to explicitly denote the antenna pair for baseline $i-j$, the A-matrix used in A-Projection at a frequency $\nu$ and time $t$ can be written in terms of antenna based quantities as

$$A_{ij} = J_i \otimes J_j^*$$

where

$$J_i = \left[ \begin{array}{cc} E_i^p & E_i^{p-q} \\ E_i^{p-q} & E_i^q \end{array} \right].$$

$E^p$ and $E^q$ are the polarized antenna aperture illumination patterns for the two polarization states. The off-diagonal terms are the leakage patterns. $A_{ij}$ is the DD equivalent of the $4 \times 4$ DI Mueller matrix for a given antenna pair. The elements of $A_{ij}$ are the complex convolution of the two antenna aperture illumination patterns $(E_i^p \times E_j^p)$, $(E_i^p \times E_j^{p-q})$, etc. For comparison, the elements of $J_i \otimes J_j^*$ would be $(E_i^p \cdot E_j^p)$, $(E_i^p \cdot E_j^{p-q})$, etc.

To keep the notation simple, in the following description we use a single subscript $k \equiv (ij, \nu, t)$ to refer to a measurement from a single baseline $ij$, at a spot frequency $\nu$ and an instant in time $t$. The vectors $V$ and $\delta$ are full polarization vectors whose elements are two-dimensional (2D) functions in the visibility plane (the $uv$-plane). Elements of $V$ are the 2D visibility data and elements of $\delta$ are 2D Delta functions representing the $uv$-sampling function for the data sample $k$. The superscripts obs, $M$, and $o$ refer to the observed, model, and true values, respectively.

Using the notation described above, the $\chi^2$ can be written as

$$\chi^2 = \sum_k V_k^R \otimes A_k V_k^M,$$

where $V_k^R = V_{k}^{\text{Obs}} - V_{k}^{M}$ and $A_k$ is inverse of the noise covariance matrix. The vector $V_{k}^{\text{Obs}}$ can be expressed in terms of $A$ as

$$V_{k}^{\text{Obs}} = (A_k^* \cdot V^o) \delta_k.$$

Note that, as mentioned before, the elements of $A^o$, $V^o$ and $\delta_k$ are 2D functions. The symbol “$\otimes$” represents the element-by-element convolution. $V^o$—without a subscript—represents the true continuous Coherence function. $V_{k}^{\text{Obs}}$ represents a sample of this Coherence function measured at the parameters represented by subscript $k$.

The calibration matrix for Equation (5) to correct for the effects of $A_k^o$ is $A_k^{-1}$ given by

$$A_k^{-1} = \frac{\text{adj}(A_k)}{\det(A_k)}.$$
4. If not converged, go to Step 2.

Since $A_k$ does not change from one major cycle to another, accumulation of $A^M$ is done in the first major cycle and cached for use in subsequent major cycles.

2.1.2. A-Projection: A Direction-dependent Gain Correction Algorithm

The antenna illumination pattern is essentially a DD description of antenna based complex gains in the data domain. $A_k$ is a DD generalization of the G-Jones matrix--the DI Mueller matrix for antenna gains in the Hamaker et al. (1996) formulation and since the A-Projection algorithm corrects for the effects of $A_k$, it can be thought of as an algorithm for DD calibration.

To establish the equivalence between DD corrections via A-Projection and DI antenna-based complex gain correction, we note that Equation (5) is the DD equivalent of DI measurement equation given by

$$V_{ij}^{\text{obs}} = G_{ij} \cdot V_{ij}^0.$$  (8)

$A_k$ in Equations (2) and (5) is the DD equivalent of $G_{ij}$ and the outer-convolution (“⊗”) and the “*” operators are the DD equivalent of the outer-product (“⊗”) and element multiplication. Calibration for $G_{ij}$ is done by multiplying Equation (8) by $G_{ij}^{-1}$ given by

$$G_{ij}^{-1} = \frac{G_{ij}^*}{|G_{ij}|^2}.$$  (9)

For calibration, Equation (6) is the DD equivalent of the above equation for the DD calibration.

For an intuitive understanding, we note that for the simpler case where the off-diagonal elements of $J$ are negligible, $\text{adj}(A_k) = A_k^\dagger$ and $\det(A_k) = \text{trace}(A_k) = \sum_i E_{ik}^P E_{ik}^{P'^*}$. For this simpler case, examination of Equations (6) and (9) shows the equivalence between DI gain calibration and the A-Projection algorithm more clearly. The process of imaging using the $\text{adj}(A_k)$ and normalizing the resulting image by $\det(\sum_k A_k)$ therefore, even for the general case, is the DD generalization of the antenna-based complex gain calibration.

3. THE WB A-PROJECTION ALGORITHM

$\text{adj}(A_k)$, when used in the A-Projection algorithm, corrects for the polarization and other DD effects that can be encoded in the phase of $A^*$ (e.g., time-varying gains due to polarization squint, ionospheric phase-screen, etc.). It is, however, not a conjugate operator for variations along the time or frequency axis.

The image domain effects of the time varying gains are largely in the amplitude scaling only (i.e., they do not disperse the flux in the image domain). Since the minor cycle algorithms typically assume that the sky brightness distribution is time-invariant and do not parameterize the model image in time, the effects of such variations can be ignored in the transform from the data to image domain. If the deviations from the average value are small (e.g., for antenna arrays and long integrations where time variability is cyclic, or antennas with three-axis mounts where beams do not rotate on the sky) the model prediction stage, which properly includes these effects, corrects for time variability in an the iterative deconvolution scheme.

Frequency dependence of $A^*$ also varies with time (and direction). This variation is not cyclic and its maximum deviation from the average value increases with fractional bandwidth.

Due to this, as for time varying gains, while the frequency dependence of $A^*$ in the inner part of the main lobe of the PB can also be corrected via the model prediction stage, the convergence is significantly slower (requiring more major cycles and hence higher computing). Alternatively, this frequency dependence can be absorbed, to some extent, in the multi-term MFS (MT-MFS) minor cycle algorithm, which solves for the time-invariant frequency dependence in the image plane. But this requires more Taylor-terms, which also is inadvisable (see Section 4.3.3).

Correction for the time-variable frequency dependent effects of $A^*$ requires a wide-band version of the $A^M$ matrix such that $\text{adj}(A^*_k) \star A^*$ results in a function which does not vary with frequency (i.e., $\text{adj}(A^M)$ is also a conjugate operator for frequency). The frequency dependence will then be projected out prior to accumulation, resulting in an image corrected for the frequency dependent effects of $A$.

3.1. The Wide-band A Operator

For the reasons given above, as well as to keep the parameters for modeling the instrumental effects in the data domain separate from parameters for modeling the sky brightness distribution, we need to construct the $A_{v}(v)$ matrix, which projects out the dependence on frequency during imaging. One such possibility is the following:

$$A_v(v) = F^{-1} \left[ P_{\text{eff}}(v) P_{\text{eff}}^*(v) \right] P(v) ,$$  (10)

where $P(v) = F^* A(v)$, and $P_{\text{eff}}(v)$ is the desired effective PB, which varies minimally with frequency. In the data domain, adj($A_v(v)$) $\star A(v)$ can be shown to be frequency-independent to high orders, and this use of $A_v(v)$ as a model for $A(v)$ in the reverse transform can correct for the frequency dependence of $A(v)$. However, it also has a large support size, and therefore, in itself, is not an efficient reverse transform operator.

3.2. The Conjugate Frequency

Equation (10) is valid for any frequency dependence. For the special case of PB scaling with frequency, we explore usable approximations for $A_v$, we define the conjugate frequency $\nu_\epsilon$, given by

$$\nu_\epsilon = \sqrt{2 \nu_{ref}^2 - \nu^2}$$  (11)

where $\nu_{ref}$ is the reference frequency of the continuum image, and examine the effects of choosing $A_v(v) \equiv A(\nu_\epsilon)$.

Using the same model for $A$ as used in Paper I (i.e., a model for the Very Large Array, VLA, antenna PB), one-dimensional cuts through the model PB given by $P_{\text{ref}}(v) = FA(v)$, the effective PB given by $P_{\text{eff}}(v) = F [A(v) \star A(v)] / [FA(\nu_{ref})]$, and the first and second derivatives of $P_{\text{eff}}(v)$ with respect to frequency are shown in Figure 1. A comparison of the derivatives of $P(v)$ and $P_{\text{eff}}(v)$ with frequency, shown in the two panels as blue dashed lines, shows that the effective PB is frequency independent to the first order. While it changes in structure, the maximum second derivative remains almost the same in magnitude. These figures show that the approximation in Equation (11) and use of $A(\nu_\epsilon)$ is good enough for imaging data which are not sensitive to the higher order frequency dependent effects. This approximation is

\footnote{This expression is arrived at by using a Gaussian approximation for the PB and imposing the condition that $P(\nu_\epsilon) P(v) = P_{\text{eff}}(v)$.}
Figure 1. Plot in the left panels shows the one-dimensional cuts through the PB model at a reference frequency (continuous red line) for the VLA and its first (dashed blue line) and second derivatives (dash-dotted green line) with respect to frequency. The plot on the right shows the same cuts through the $P_{\text{eff}}(\nu)$—the effective frequency-independent PB.

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

useful since it can be easily implemented, is appropriate for the sensitivity of current telescopes and covers a large fraction of scientific observations for simultaneous Stokes-$I$ and spectral-index mapping. The frequency dependence in $A(\nu)$ is reduced overall by an order of magnitude. When used in the A-Projection algorithmic steps in Section 2.1.2, it effectively corrects for the frequency dependence of the PB prior to the accumulation along the frequency axis in Equation (7). For future studies, more sensitive telescopes that will be sensitive to second order frequency dependent effects also, $A_\nu$, as in Equation (10) may be required. However, when software implementation itself requires partitioning along the time and/or frequency axis, hybrid approaches discussed in Section 4 may be more efficient.

In practice, there might be multiple terms that make up the $A$-term, some of which may scale with frequency while others may not. For example, aperture illumination will scale with frequency, but the pointing-offset term or resonance effects will not. The terms that do not scale with frequency can be applied as multiplicative terms to $A(\nu_\psi)$ during gridding. We have tested this approach to work for parallel-hand polarizations. While we have not tested it, we think this may also work for cross-hand polarizations.

To avoid confusion and for brevity, we will refer to the algorithm in Paper I as the NB A-Projection algorithm, and refer to the use of $A(\nu_\psi)$ (instead of $A(\nu)$) for the data-to-image domain transform as the WB A-Projection algorithm.

3.3. Algorithm Validation

The image deconvolution algorithm described in Section 3 was tested using simulated wide-band data with 66% fractional bandwidth. The VLA C-array was used for antenna configuration and the observations covered an hour angle range of $\pm 3^\circ$. The model for the PB used in Paper I was scaled by frequency and rotated with parallactic angle to simulate time-varying frequency-dependent effects. To clearly highlight the effects of time and frequency dependence of $A$, we used a model of the sky consisting of five point sources located at 0.99, 0.83, 0.60, and 0.11 levels of the PB within the main lobe and one source located in the first side lobe (PB gain of 0.025). All the point sources were assigned a flux of 1 Jy with flat spectra. The effective spectral-indices due to the PB at the five locations are $-0.026, -0.38, -1.0, -5.32$, and $+0.47$, respectively. No noise was added to these simulations, and all imaging and deconvolution runs were with a loop-gain of 0.2.

Figure 2 shows deconvolved images produced without (first panel) and with (second panel) WB A-Projection gridding. This comparison demonstrates that with an accurate model of the PB, it is possible to correct for its $time$ and $frequency$ $variability$ down to numerical precision levels in wide-band, wide-field imaging.

4. APPLYING NB A-PROJECTION FOR WIDE-BAND IMAGING

NB A-Projection was designed for a single reference frequency and does not automatically account for the frequency dependence of PB during gridding. However, it can still be used for wide-band imaging, as long as the frequency dependence of the far-field pattern is known and characterized by $P_{\nu}$. Several algorithmic options exist, all with different numerical approximations and computing load. This flexibility allows the implementation to be tuned according to the available computing resources, architecture of the hardware platform, and the desired imaging accuracy.

4.1. Cube Imaging + Cube Deconvolution

The simplest approach is spectral-cube imaging where each frequency channel (with its limited $uv$ coverage) is treated separately. NB A-Projection is applied as-is per channel, and the minor cycle run independently per channel. A continuum image is later constructed by adding together deconvolved and restored images from all channels.

The residual image per channel can be approximately rewritten from Equation (7) in the image domain, for the case where the aperture illumination functions are identical for all baselines and times, and only one polarization-pair is being imaged:

$$I^R_{\nu} = P_{\nu} \times \left( I^{\text{psf}}_{\nu} \star (P_{\nu} \times I^{\text{sky}}_{\nu}) \right)/P_{\nu}^2,$$  

(12)

where $I^{\text{psf}}_{\nu} = F^{-1} \sum_k \delta_{k,\nu} (k \equiv ij, t)$ is the point-spread function (PSF) for one channel and $P_{\nu} = F^{-1} A_{\nu}$. In this
expression, the division by $P_p^2$ implies a flat-sky\(^3\) normalization, but a flat-noise\(^4\) normalization may be used instead.

This method is straightforward and will suffice for modest imaging dynamic ranges and uncomplicated spatial structure (point sources). However, the angular resolution of the continuum image and any estimate of the sky spectrum will be limited to that of the lowest frequency in the band. Also, reconstruction uncertainties may be inconsistent across frequency when there is insufficient $uv$-coverage per channel or complicated spatial structure, leading to spurious spectral structure.

In this paper, we are focusing on imaging problems that require more accuracy and dynamic range than what the above offers. In the next two sections, we discuss A-Projection in the context of MFS where the combined $uv$-coverage is used for model reconstruction (minor cycle).

### 4.2. Cube Imaging + MFS Deconvolution

The simplest extension of NB A-Projection for MFS is to grid, Fourier transform, and normalize each frequency channel (or sets of channels) independently, and then produce a continuum image by an image-domain accumulation, before the minor cycle. When the sky spectrum is not flat, or when flat-noise normalization is used per channel, a wide-band minor-cycle algorithm such as MT-MFS can be applied to simultaneously solve for the sky intensity and spectrum. Taylor-weighted residual images are constructed as follows, before proceeding to the minor cycle:

$$I^R_\nu = \sum_v w^R_\nu I^R_\nu = \sum_v w^R_\nu \left( \frac{P_\nu \times (I_\nu^{\text{psf}} \ast (P_\nu \times I_\nu^{\text{sky}}))}{P_\nu^2} \right)$$

where $w^R_\nu = (\nu - \nu_0)^0$ are weights that represent Taylor polynomial basis functions (Rau & Cornwell 2011). The interpretation of the output Taylor coefficients depends on the choice of normalization as follows.

1. With flat-sky normalization per channel, the output Taylor coefficients will represent $I_\nu^{\text{sky}}$.

2. For flat-sky normalization per channel, but a regularized flat noise minor cycle, the Taylor-weighted residual images can be multiplied by a $P_\text{ref}$ after the frequency summation and before deconvolution. The output Taylor coefficients will represent $P_\text{ref} I_\nu^{\text{sky}}$ and a post-deconvolution division by $P_\text{ref}$ will be required for the intensity image. No corrections are needed for the spectral index map.

3. With flat-noise normalization per channel before Taylor-weighted averaging, the output Taylor coefficients will represent $P_\nu I_\nu^{\text{sky}}$, and a post-deconvolution polynomial division of the PB spectrum will be required to correct both the intensity and the spectral index.

Qualitatively, the reverse transform with flat-sky normalization per channel followed by frequency-averaging and a flat-noise regularization by $P_\text{ref}$ before the minor cycle, is equivalent to the use of WB A-Projection during gridding. It requires multiple images (one per channel) and fast Fourier transforms (FFTs) and has repeated beam divisions and multiplications that can increase numerical errors. However, it is naturally parallelizable, making it an attractive option for extremely large data sets and imaging goals where inaccuracy in low gain regions of the PB can be tolerated.

#### 4.2.1. MFS Imaging + MFS Deconvolution

To optimize on memory use and FFT costs (especially in a non-parallel imaging run), MFS gridding can be done, where...
averages over baseline, time, and frequency are accumulated onto a single grid, followed by a single FFT and normalization by an average PB (or its square). Here, attention must be paid to the consequences of averaging over frequency before normalization. The use of \text{adj}(A(\nu)) as the gridding convolution function (the NB A-Projection algorithm), introduces an additional frequency dependence \( P_\nu \), that gets averaged over before it can be removed. Once gridded, this extra frequency dependence is locked in, and can be accounted for only as an artificially steeper spectrum in the minor-cycle of the MT-MFS wide-band imaging algorithm.

Multi-term Taylor-weighted residual images must be constructed as follows:

\[
I_i^{R} = \sum_\nu w_i^\nu \left\{ P_\nu \times \left( I_{\nu}^{psf} \ast \left( P_\nu \times I_{\nu}^{\text{sky}} \right) \right) \right\},
\]

(14)

and the output Taylor coefficients will depend on the choice of normalization as follows:

1. For flat-sky normalization, the model intensity represents \( I_{\text{ref}}^{\text{sky}} \), but the spectrum represents the product of the sky spectrum and the square of the PB spectrum.
2. For flat-noise normalization, the model intensity represents \( I_{\text{ref}}^{\text{sky}} P_\text{avg} \) and the spectrum represents the product of the sky spectrum and the square of the PB spectrum.

In both cases, appropriate wide-band post-deconvolution corrections for the average PB and two instances of the PB spectrum must be applied.

Such corrections are inelegant, and are susceptible to numerical instabilities in low-gain regions of the PB. Section 4.3 shows a comparison of some of these methods with WB A-Projection, for a simulation with source spectra that are not flat and therefore require MT-MFS imaging and deconvolution.

### 4.3. Comparison of Hybrids with WB A-Projection

To test the algorithm described in Sections 3 and 3.1 with non-flat source spectra, we used the sky brightness distribution as in Figure 2, but assigned a spectral index of \( \alpha = -0.5 \) to all sources such that \( I(\nu) \propto (\nu/\nu_0)^\alpha \).

Figure 3 shows deconvolved images produced with and without time-dependent and frequency-dependent PB corrections during gridding, emphasizing the different types of error patterns that arise when one or more effects are ignored. An image formed from the hybrid method described in Section 4.2.1 to absorb all frequency dependence into the minor cycle solver is also shown for comparison. Figure 4 shows Stokes-\( I \) and spectral index values for these point sources after PB correction, to illustrate the accuracy to which different methods are able to recover the true-sky spectral index at various locations in the PB. The various methods tested and results obtained are described below. The rms and peak-residuals from using these algorithms are tabulated in Table 1.

### 4.3.1. MFS + SI (Standard Imaging)

The image in the top left panel of Figure 3 is the result of standard Cotton–Schwab CLEAN with MFS gridding using prolate-spheroidal functions as gridding-convolution functions, and a flat-spectrum assumption during the minor cycle:

\[
I_i^{R} = \sum_\nu \left\{ \left( I_{\nu}^{psf} \ast \left( P_\nu \times I_{\nu}^{\text{sky}} \right) \right) \right\}.
\]

(15)

Time and frequency variability of both the sky and the instrument are ignored, and for a 66% bandwidth, imaging artifacts around all sources away from the pointing center are dominated by spectral effects due to \( P_\nu \) present in the data. A post-deconvolution division by an average PB can recover the true source intensity to within a few percent, out to the half-power point of the PB, but errors increase with distance from the pointing center.

### 4.3.2. MT-MFS+ SI

The image in the top right panel of Figure 3 is the result of the MT-MFS algorithm in the minor cycle, with standard gridding (prolate-spheroidal functions). The minor cycle solves for the average intensity and spectrum of \( I(\nu) P(\nu) \) using a two-term Taylor-polynomial approximation:

\[
I_i^{R} = \sum_\nu \left\{ \left( I_{\nu}^{psf} \ast \left( P_\nu \times I_{\nu}^{\text{sky}} \right) \right) \right\}.
\]

(16)

Table 1

| Panel in Figure 3 | Algorithm      | Description                                   | rms (Jy/beam) | Peak Residual (Jy/beam) | Comments                                                                 |
|-------------------|----------------|-----------------------------------------------|---------------|-------------------------|--------------------------------------------------------------------------|
| Top Left          | MFS + SI       | Standard wide-band imaging                    | \( 6 \times 10^{-4} \) | \( 2.3 \times 10^{-4} \) | Ignore time \& frequency dependence. Artifacts due to time and frequency variations of the PB. Ignore time dependence. Absorb time-averaged frequency dependence in MT-MFS. Artifacts due to time-variability of the PB. |
| Top Right         | MT-MFS + SI    | Multi-term imaging with standard gridding     | \( 1 \times 10^{-4} \) | \( 5 \times 10^{-4} \)  |                                                                                                                                         |
| Lower Left        | MT-MFS + A-Projection | Multi-term imaging with NB A-Projection gridding                           | \( 4 \times 10^{-5} \) | \( 8 \times 10^{-4} \)  | Account for time variability of PB, and absorb the resulting PB² frequency dependence in MT-MFS. Artifacts due to stronger spectral structure. |
| Lower Right       | MT-MFS + WB A-Projection | Multi-term imaging with wide-band A-Projection gridding                         | \( 3.5 \times 10^{-5} \) | \( 2 \times 10^{-4} \)  | Account for PB time- \& frequency-dependence in WB A-Projection. Account for static sky-frequency dependence in MT-MFS. Minimal artifacts. |

Average PB-spectral effects are absorbed into the sky model, and the dominant remaining error is due to the time variability of the PB. A post-deconvolution correction of the continuum intensity and spectral index are accurate to within a few percent in intensity and \( \pm 0.1 \) in spectral index out to approximately the half-power point. Beyond this field of view (FoV), errors increase (to \( \pm 0.4 \) or more in spectral index) primarily because a time-averaged PB spectrum is not a good estimate in regions of the image where \( P_\nu \) changes by 100% with time as the beams rotate on the sky.
The Astrophysical Journal, 770:91 (9pp), 2013 June 20

Bhatnagar, Rau, & Golap

Figure 3. Comparison of the imaging performance before and after applying corrections for the time and frequency dependence of the PB during imaging. All restored images are shown at the same gray scale, stretched to emphasize artifacts. Contours are drawn at the 0.02, 0.1, and 0.5 (HPBW) levels of the time-and-frequency averaged PB. Results from four algorithms described in Section 4.3 are compared here (MFS+SI, MT-MFS+SI, MT-MFS+A-Projection, MT-MFS+WB A-Projection). The rms and peak-residuals are listed in Table 1.

Figure 4. Comparison of the accuracy of the PB-corrected intensity (LEFT) and spectral-index (RIGHT) for the five simulated point sources, using the four methods whose results are shown in Figure 3. The labels “AWP” and “WB AWP” are used for A-Projection and WB A-Projection in the figure. The algorithms compared are MFS+SI, MT-MFS+SI, MT-MFS+A-Projection and MT-MFS+WB A-Projection. Spectral-indices are shown only for methods using MT-MFS, with post-deconvolution (average) spectral-index corrections done for the SI and A-Projection runs. Results for the five sources are shown from left to right with increasing distance from the pointing-center. The reference-PB gain and effective PB-spectral-index at the locations of the five sources are listed on the x-axis. These plots show that outside the HPBW at the reference-frequency, methods that do not account for time-variable PB-spectra have considerably higher errors, and the combination of MT-MFS+WB A-Projection delivers accurate corrections even out in the sidelobe.

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

4.3.3. MT-MFS+ A-Projection

The image in the bottom left panel of Figure 3 is the result of MT-MFS in the minor cycle (two terms), but with the NB A-Projection gridding as described in Section 4.2.1 with a flat-noise normalization before the minor cycle, followed by a post-deconvolution correction of the intensity by $P_{\text{ref}}$ (average PB) and the spectral index by twice that of the PB. Artifacts due to frequency-independent time variability (antenna rotation) no longer exist within the HPBW (PB gain of 0.5), but new spectral artifacts appear away from the pointing center (beginning around the 10% level).
These errors are partly due to the increased nonlinearity of the $P^2(\nu)$ spectrum away from the pointing center, for which a two-term Taylor-polynomial approximation is insufficient. A run with three terms partially reduces this problem, indicating that errors in approximating the combined spectrum with a low-order polynomial dominates the errors, but higher order polynomials are inadvisable because of instability in low-S/N regions.

Errors also arise from the high time variability of the PB-spectrum, which is ignored because only time-averaged spectra are used for spectral-correction. A post-deconvolution correction of the spectral-index map for $P^2(\nu)$ results in errors at the $\pm0.3$ level beyond the $\sim50\%$ point.

4.3.4. MT-MFS + WB A-Projection

The image in the bottom right panel of Figure 3 is the result of MT-MFS in the minor cycle (two terms), and WB A-Projection gridding (Section 3). Artifacts around all sources are gone, and the dominant errors are numerical (at the floating-point precision level). The spectral-index map produced by MT-MFS is accurate to within 0.01 in the main lobe, and 0.05 out in the sidelobe. Such accuracies allows the recovery of source-spectra further-out in the PB than previously possible. The main difference between this method and all others, is that time and frequency variability of the PB has been corrected for in the data domain, before any averaging is done to construct a continuum image to send to the minor cycle. The minor cycle sees a flat-noise normalization, preserving the convolution-equation and allowing for deeper “cleaning” before triggering the next major cycle (i.e., faster convergence).

This method shows the lowest errors in Figure 4, indicating that if the PB can be accurately modeled, its time and frequency variability can be corrected for during gridding, resulting in an accurate reconstruction in the minor cycle.

4.4. Imaging Results with VLA L-band Data

Figure 5 shows the continuum intensity and spectral index distribution of the G55.7+3.4 Galactic supernova remnant (SNR), using the VLA in L-band and D-configuration and made with and without the WB A-Projection algorithm. The peak brightness is 6 mJy, and with an off-source rms of 11 $\mu$Jy, this is
We also analyzed hybrid schemes for wide-band imaging using the NB A-Projection and image-plane correction for the effects of PB. Our conclusion is that while for non-parallel implementation, WB A-Projection is required for wide-band imaging, for implementations that may require partitioning the data along time and/or frequency axis, hybrid approaches are also sufficient. However the use of WB A-Projection in implementations on parallel processing platforms allows the freedom to tune the distribution of the data to suite the available hardware and computing resources (e.g., this allows imaging smaller chunks of the total bandwidth in parallel, even if these smaller chunks need wide-band PB corrections. Without WB A-Projection, the data distribution is restricted to be partitioned in frequency such that each chunk can be imaged using NB A-Projection).

Comparisons show that the WB A-Projection plus MT-MFS enables simultaneous intensity and spectral index imaging throughout the PB in wide-field imaging. Moderate dynamic range imaging within the half-power point of the PB is possible where all frequency dependence in the image is absorbed in the solution of the MT-MFS algorithm. Beyond this FoV, errors increase because a time-averaged PB spectrum is not a good estimate in regions of the image where PB changes by 100% with time as the beams rotate on the sky. The FoV can be increased until the ~10% point of the PB by combining MT-MFS with NB A-Projection. The time-dependence of the PB is accounted for via A-Projection and its frequency dependence is absorbed in MT-MFS. Using larger number of Taylor-terms in MT-MFS improves the imaging performance for simpler fields, but is inadvisable because of instability in low-S/N regions.

Finally, we would like to note that while only the effects of the antenna PB were included in the A-term used in this paper, other antenna-based DD effect can also be easily included. The effect of non-isoplanatic ionospheric/atmospheric phases is comparable to the effect of PB for wide-band, wide-field imaging at low frequencies, particularly with aperture-array antenna elements. Similar effects come from the irregularities in the water vapor content in the lower atmosphere for imaging at high frequencies. The effects due to ionosphere/atmosphere and PB need to be corrected simultaneously, often for wide-band data in full polarization. It may be possible to extend the WB A-Projection algorithm presented here to include corrections for ionospheric effects. Work to test these extensions is underway and will be reported in future publications.

This work was done using the R&D branch of the CASA code base. We thank the CASA Group for the underlying libraries. We thank T. J. Cornwell for his useful and detailed comments/suggestions as a referee. Part of this work was funded by the ALBiUS work package of the European Commission Radionet FP7 program. The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc.

REFERENCES

Bhatnagar, S., Cornwell, T. J., Golap, K., & Uson, J. M. 2008, A&A, 487, 419

Cornwell, T. J. 1995, The Generic Interferometer: II Image Solvers, Tech. rep., AIPS++, Note 184

Hamaker, J. P., Bregman, J. D., & Sault, R. J. 1996, A&AS, 117, 137

Rau, U., Bhatnagar, S., Voronkov, M. A., & Cornwell, T. J. 2009, IEEE, 97, 1472

Rau, U., & Cornwell, T. J. 2011, A&A, 532, A71