Draft version August 20, 2012
Preprint typeset using LaTeX style emulateapj v. 03/07/07

WIDE-FIELD WIDE-BAND FULL POLARIZATION INTERFEROMETRIC IMAGING: THE WB A-PROJECTION ALGORITHM

S. Bhatnagar
National Radio Astronomy Observatory, Socorro, NM - 87801, U.S.A.

U. Rau
National Radio Astronomy Observatory, Socorro, NM - 87801, U.S.A.

K. Golap
National Radio Astronomy Observatory, Socorro, NM - 87801, U.S.A.

(Dated: Received: 08/08/2012; Accepted:)
Draft version August 20, 2012

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 forward gain in full polarization. In this paper we present the Wide-Band A-Projection algorithm (*WB A-Projection*) to also account for the effects of wide bandwidth and show that the algorithm can simultaneously correct for the time, frequency and polarization dependence of the PB. The algorithm also naturally interfaces with existing image-plane algorithms for similar time-, frequency- and polarization-dependent variations of the sky brightness distribution. In particular, we discuss the combined *WB A-Projection* and the Multi-term Multi Frequency Synthesis (*MT-MFS*) algorithm for simultaneously mapping the sky brightness distribution and its spectral structure across wide fields of view. As part of our investigation, we also explored alternatives to the *WB A-Projection* algorithm for wide-band wide-field imaging. Here we show the limitations of these alternatives and argue that out of all the approaches we investigated, *WB A-Projection* in combination with *MT-MFS* offers an optimal solution in terms of imaging performance and algorithm complexity.

Subject headings: Techniques: interferometric – Techniques: image processing – Methods: data analysis

1. INTRODUCTION

Many fundamental scientific questions in astrophysics today require high-sensitivity surveys covering large cosmic volumes in sky-coverage and redshift. 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, therefore 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 which afford such high sensitivities, impact the imaging performance and are significantly different from previous generation telescopes are: 1) the ultra-wide instantaneous fractional bandwidths, and 2) larger collecting area. The effects of wide instantaneous fractional bandwidths that classical calibration and imaging algorithms ignore, lead to errors higher than the sensitivity that these new telescopes offer. Examples, relevant for some of the telescopes already in operation like the Karl G. Jansky Very Large Array (VLA) (Perley et al. 2011) and ALMA (Wootten & Thompson 2009), include the effects of time and frequency variant primary beams, frequency dependence of the emission from the sky and antenna pointing errors. At low frequencies, the effects of wide fractional bandwidth and ionospheric phase screen limit the imaging performance of the GMRT (Swarup et al. 1991), VLA and LOFAR. Additionally, significant variations in the shape of the wide-band primary beams (PB) for LOFAR (Röttgering 2003) and other aperture array telescopes like LWA (Ellingson et al. 2009)/PAPER (Parsons et al. 2010) /MWA (Tingay et al. 2012) leads to errors of similar magnitude. All these effects form the general class of problems often referred to in the literature as "direction dependent effects" or DD effects. Both, wide fractional bandwidths and larger collecting area also lead to many orders of magnitude increase in the data volume, putting severe constraints on the run-time performance of the algorithms for calibration and imaging.

The errors due to these DD effects were smaller in magnitude than the thermal noise limit of the previous generation telescopes. The data volume from these telescopes was also significantly smaller due to narrower bandwidths. Classical calibration and imaging algorithms therefore have ignored these DD effects and were not designed to handle wide bandwidths and large data sizes. Albeit with limited success, these classical algorithms have been used in highly complex data processing schemes and applied to relatively small data sets for removing some DD errors. These schemes however fundamentally need multiple passes through the data, can be shown to be non-optimal in terms of utilizing the high signal-to-noise ratio of modern telescopes and often require human

Electronic address: sbhatnag@nrao.edu
Electronic address: rurvashi@nrao.edu
Electronic address: kgolap@nrao.edu

The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc.

intervention. These are therefore unsuitable for modern data sets which are already in the 100s of Gigabytes to 10s of Terabytes range and also require pipeline processing. The software complexity of these schemes is also significantly higher, which drives the cost of software development and maintenance higher. An integrated approach to simultaneously account for all time-, frequency- and polarization-dependent DD effects is required to fully utilize the sensitivity of modern radio telescopes at an affordable cost.

In the following sections we discuss the issues related to correcting for wide-band time-varying DD effects in general and describe the Wide-band A-Projection algorithm (*WB A-Projection*) to simultaneously correct for all known (modeled or measured) PB effects. We also describe the combined *WB A-Projection* and *MT-MFS* algorithm for simultaneously correcting wide-band wide-field instrumental effects and modeling the frequency dependence of the sky emission. As part of the investigations, we also explored alternatives which may allow the use of the existing *A-Projection* algorithm (Bhatnagar et al. 2008, henceforth referred to as Paper-I) to correct for the effects of time and polarization variability of the PB, and use the *MT-MFS* (Rau 2010; Rau & Cornwell 2011) algorithm to account for *all* of the wide-band effects in the data, from both the sky and the instrument. While some of these approaches looked promising in the beginning, they required considerable careful investigation to come to the conclusion that these approaches do not offer the required solution, and we report on our findings in Section 6.

2. PARAMETRIZATION: THE NATURAL DOMAINS FOR MODELING SKY AND INSTRUMENTAL PARAMETERS

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

$$\vec{V}_{ij}^{Obs}(\nu, t) = W_{ij}(\nu, t) \int \mathsf{M}_{ij}^{Sky}(\vec{s}, \nu, t)\vec{I}(\vec{s}, \nu) e^{\iota \vec{b}_{ij} \cdot \vec{s}} d\vec{s} \qquad (1)$$

where $\vec{V}_{ij}^{Obs}$ are the observed visibility samples measured by the pair of antennas designated by the subscript i and j, separated by the vector $\vec{b}_{ij}$ and weighted by the measurement weights W_{ij}. M_{ij}^{Sky} is the radio-Muëller matrix[1] representing the full polarization description of the complex *direction dependent* gain as a function of the direction $\vec{s}$, frequency ν and time t and $\vec{I}$ is the image vector. The vectors $\vec{V}$ and $\vec{I}$ are full polarization vectors in the data and image domain respectively. M_{ij}^{Sky} and $\vec{I}$ are the unknowns in this equation.

Removing the effects of the *direction-independent* terms is traditionally done as part of the calibration step. The calibrated data is then used to estimate the true, uncorrupted $\vec{I}$ – the primary scientific product. Such a separation of calibration and imaging is however not possible for removing the effects of M_{ij}^{Sky}, since it cannot be taken out of the Fourier integral and applied directly to the data independent of imaging. This makes it a fundamentally harder problem to solve, and conventional calibration and imaging algorithms have ignored it.

To facilitate the discussion on optimal algorithm design for wide-band wide-field imaging, we define the term *"natural domain"* as the respective domains for M^{Sky} and I where they can be adequately represented in the most compressed manner (i.e as parametrized models involving the least number of parameters). The image domain is clearly the *natural domain* for parametrized representation of I (or any other term related to sky brightness distribution in general). The primary purpose of a minor-cycle algorithm is to solve for an image-plane model of the sky brightness distribution (see section 3.1 and Rau et al. (2009) for the structure of imaging algorithms).

The image domain essentially corresponds to a vector average of the data (V_{ij} in Eq. 1), averaged with an appropriate kernel. Only the average values of otherwise antenna-based quantities in Eq. 1 (e.g. M^{Sky}) are available for manipulation in the image plane. Therefore, errors due to the variations in M^{Sky} with time, frequency or antenna cannot be fully corrected in the image domain or as part of the minor cycle. Furthermore, these errors must be corrected as part of the process of computing the residual image prior to triggering the minor cycle for the minor cycle algorithms to accurately solve for a model for I.

While the image domain contains only average quantities (averaged over time, frequency, polarization and baseline), the data domain samples coherence field along the time, frequency, polarization and baseline axes. Furthermore, while M_{ij}^{Sky} varies along these axes, it remains separable into antenna based quantities (Eq. 2) in the data domain, thus requiring only N models for $O(N^2)$ measurements. Therefore, not only is it possible to properly model variations in time, frequency and polarization in the data domain, the number of parameters required for this modeling are also reduced. The data domain is therefore the *natural domain* for a parametric representation of M_{ij} (or any other fundamentally antenna based effect in general). Hence, while the effects of antenna PB variations may be most easily seen and may be intuitively understood in the image plane, it is fundamentally an aperture plane effect and it is optimal to model and design algorithms to correct for its effects in the data domain.

The dominant sources of frequency dependence of the data come from the variations of sky brightness distribution and of the PB with frequency. The image-domain Multi-Term, Multi Frequency Synthesis (*MT-MFS*) algorithm models the frequency dependence of the sky emission as a polynomial in frequency (Sault & Wieringa 1994; Rau 2010; Rau & Cornwell 2011) . As part of an iterative image deconvolution algorithm, the *MT-MFS* algorithm affords simultaneous total power and spectral-index imaging covering the inner ∼30% of the antenna PB area (see Bhatnagar et al. 2011 for an example). Beyond this point, even in the main lobe of the PB, the variability of the PB with time and frequency becomes significant and the resulting residual errors in the image are higher than the thermal noise limit. An imaging algorithm that *simultaneously* accounts for the frequency dependence of the antenna PB and solves for the frequency dependence of the sky emissions is required. For the narrow-band case, the *A-Projection* algorithm in Paper-I corrects for the effects of time and polarization dependence of the PB. For the wide-band case, the variations in the PB as a function of frequency must also be accounted for during imaging.

3. THEORY

---

[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).

M_{ij}^{Sky} in Eq. 1 includes all DD gains in full polarization and can be written in terms of antenna-based quantities as:

$$\mathsf{M}_{ij}^{sky} = \mathsf{J}_i(\vec{s}, \nu, t) \otimes \mathsf{J}_j^*(\vec{s}, \nu, t) \tag{2}$$

where $\otimes$ represents an outer product. For the purpose of the discussion in this paper we define J_i is the full polarization voltage response of the individual antennas in the feed-polarization bases (i.e. circular or linear polarization). An appropriate unitary transform can be applied to convert the above equation to Stokes bases (see Hamaker et al. (1996) for details). Then, in order to benefit from the full available bandwidth, the data is averaged across frequency prior to inverting Eq. 1 for continuum imaging. All frequency dependence in the right hand side of this equation must therefore be removed either prior to averaging or accounted for during this averaging process.

3.1. *Imaging*

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 $\vec{I}$ is therefore typically done via iterative non-linear χ^2-minimization. Writing the observed data and the model data as vectors $\boldsymbol{V^{Obs}}$ and $\boldsymbol{V^M}$ containing all the data, $\boldsymbol{\chi^2}$ can be written as

$$\chi^2 = \left[\vec{V}^R\right]^\dagger \Lambda \left[\vec{V}^R\right] \tag{3}$$

where $\vec{V}^R = \vec{V}^{Obs} - \vec{V}^M$ and Λ is inverse of the noise covariance matrix. Using operator notation developed by Rau et al. (2009) and the symbols used in Paper-I, each element of the vector $\vec{V}^{Obs}$ can be expressed in terms of M_{ij}^{Sky} and the *true* visibility field vector $\vec{V}^\circ$ as

$$\vec{V}_{ij}^{Obs} = \mathsf{A}_{ij}^\circ \vec{V}^\circ \tag{4}$$

where $\mathsf{A}_{ij}^\circ = W_{ij} S_{ij} \mathsf{F} M_{ij}^{\circ sky} \mathsf{F}^\dagger$, F is the Fourier transform operator, $M_{ij}^{\circ sky}$ is the true DD radio-Muëller matrix and S_{ij} represents the sampling function. The operator A_{ij}° is therefore the data domain filter that describes measurements from a single baseline $i - j$. A° can be thought of as the DD version of the classical sampling operator S and can be constructed by replacing each non-zero entry of the classical S by the appropriate element of A°.

Minimization involves two stages of processing, described in the radio astronomy imaging literature as "major cycle" and "minor cycle". The major cycle corresponds to the computation of the derivative of the χ^2 with respect to the image model and the model for the sky emission is built in successive minor cycle iterations. The derivative of the χ^2, the update direction in the parameter space, is the residual image ($\boldsymbol{I^R}$) (Cornwell 1995) and forms the input to the minor cycle. For efficiency reasons, the minor-cycle algorithms are often themselves iterative. $\boldsymbol{I^M}$ is successively updated and subtracted from $\boldsymbol{I^R}$ in the minor cycle iterations until the residual image is no longer an accurate description of the update direction, at which point the derivative is recomputed (the major cycle) at the current location in the parameter space (the current value of $\boldsymbol{I^M}$).

The major cycle implements the forward and the reverse transforms between the data and image domains. First, the forward transform converts the current cumulative model image into model data. The model data is then subtracted from the calibrated data and the reverse transform is applied to the residual data to compute the residual image. As discussed in section 2, since calibration for DD effects cannot be separated from imaging (DD terms cannot be taken out of the integral in Eq. 1), correction for DD effects is optimally done as part of the major cycle. The narrow band *A-Projection* algorithm in Paper-I is an example of such a major cycle algorithm.

Using A as a model for A°, the residual image at each major cycle iteration is computed as:

$$\vec{I}^R = \left[\vec{I}^{wt}\right]^{-1} \mathsf{F}^\dagger \mathsf{A}_{ij}^\dagger \vec{V}_{ij}^R \tag{5}$$

where

$$\left[\vec{I}^{wt}\right] = det\left[\mathsf{F}^\dagger \overline{\left|\mathsf{A}_{ij}(\nu)\right|}\right] \tag{6}$$

The averaging in the equation above is done over time, frequency and baseline. A° represents the DD effect of interest. Since A° is part of the observations process and included in the data, we do not have any control on it. If however A is constructed such that, (1) it has the required properties for it to be used as a gridding convolution operator for the forward and reverse transforms, (2) $\left[\mathsf{A}^\dagger \mathsf{A}^\circ\right] \approx \left[\mathsf{A}^\dagger \mathsf{A}\right]$ (i.e. A is a good model for A°) and (3) $\left[\mathsf{A}^\dagger \mathsf{A}^\circ\right]$ is constant (at least approximately) as a function of the variable(s) describing the DD effect we want to correct (time, frequency, polarization, etc.), the reverse transform in Eq. 5 produces an image that is corrected for the effects of M^{Sky} (see Bhatnagar et al. (2008) for a more detailed discussion) . If A describes the DD effect accurately (to within a normalization constant – the constant is absorbed in $\boldsymbol{I^{wt}}$), it can be used in the forward transform to accurately compute model data from a model image including DD effects. Using such a forward transform and approximate reverse transform in the major cycle to account for non-image plane DD effects, along with the appropriate minor cycle algorithm to solve for image plane effects, an iterative minimization scheme can be devised to produce $\boldsymbol{I^M}$ that describes the true continuum sky brightness free of the effects of M^{Sky}.

4. THE A OPERATOR FOR WIDE-BAND FULL-POLARIZATION IMAGING

The A-term described in Paper-I is a full-polarization A-Projection operator for *narrow band full-polarization imaging* given by

$$\mathsf{A}_{ij} = \mathsf{A}_i \otimes \mathsf{A}^*{}_j \tag{7}$$

where A_i is the antenna based Jones matrix (here, the full-polarimetric antenna aperture illumination pattern). Each term of A_{ij} is a convolution of two Jones matrices and A_{ij} itself is a convolution operator. The effective *full polarization* operator, after proper normalization, is approximately unitary and when used in iterative deconvolution algorithms, corrects for the effects of the rotation of the azimuthally asymmetric antenna PB and associated variations with time and polarization. However, A_i (as described in Paper-I) is computed at a single reference frequency and does not account for the frequency dependence of M^{Sky}.

4.1. *Frequency-dependent prediction*

The far-field radiation pattern is the Fourier transform of the baseline aperture illumination pattern (A_{ij}), the width of which scales with frequency. This frequency dependence can be included in the forward transform by computing $\mathsf{A}(\nu)$ from a model (or measurement) of A at a reference frequency. Using $\mathsf{A}(\nu)$ in the forward transform at frequency ν

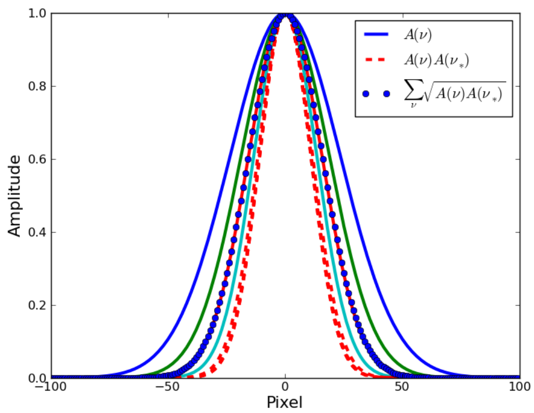

FIG. 1.— The solid lines represents $\boldsymbol{P(\nu)} = \mathbf{F}^\dagger\mathbf{A}(\nu)$ at four different frequencies between 1.0 GHz (outer most curve) and 2.0 GHz (inner most curve). The set of curves shown with dashed lines (red) represents the $P^2_{eff}(\nu)$ The curve marked with filled circle (blue) show the normalized average of $P_{eff}(\nu)$.

will accurately predict the data, including the effects of time-, frequency- and polarization-variability of the antenna aperture illumination pattern.

4.2. *The Conjugate frequency*

For the reverse transform, we consider the operator $\mathbf{A}_*$ given by

$$\mathbf{A}_*(\nu) = \mathbf{F}^\dagger\left[\frac{P_{ref}P^*_{ref}}{P(\nu)}\right] \tag{8}$$

where $P(\nu) = \mathbf{F}^\dagger\mathbf{A}(\nu)$ and P_{ref} is the desired effective frequency-independent PB. $\mathbf{A}_*{}^\dagger(\nu)\mathbf{A}(\nu)$ can be shown to be frequency-independent to high orders, and the use of $\mathbf{A}_*$ as a model for $\mathbf{A}(\nu)$ in the reverse transform can correct for the frequency dependence of $\mathbf{A}(\nu)$. However it also has a large support size, and therefore, in itself, is not an efficient reverse transform operator.

To explore usable approximations for $\mathbf{A}_*$, we define the *conjugate frequency*, denoted by ν_*, as

$$\nu_* = \sqrt{2\nu^2_{ref} - \nu^2} \tag{9}$$

where ν_{ref} is the reference frequency of the continuum image, and consider the following operator instead:

$$\mathbf{A}_* = \mathbf{A}(\nu_*) \tag{10}$$

For illustration, using a gaussian function for $\mathbf{A}(\nu)$ the width of which scales with ν, one-dimensional cuts through the model PB $P(\nu) = \left[\mathbf{F}^\dagger\mathbf{A}(\nu)\right]$ and effective PB $P_{eff}(\nu) = \sqrt{[\mathbf{F}^\dagger\,(\mathbf{A}^\dagger(\nu_*)\mathbf{A}(\nu))]}$ at different frequencies are shown in Fig. 1. The continuous-line curves show $P(\nu)$ at five equally separated values of ν between 1.0–2.0 GHz. The set of curves shown with dashed lines are for $P^2_{eff}(\nu)$ and the curve marked by filled blue circles is the square root of the frequency-averaged function given by $P_{ref} = \overline{P_{eff}(\nu)}$. While the curves for $P(\nu)$ scale with frequency (as expected), the set of curves for $P_{eff}(\nu)$ are largely independent of frequency.

Using the same model for $\mathbf{A}$ as used in Paper-I (i.e. a model for the VLA antenna PB), Fig. 2 shows the derivative of $P(\nu)$ (left panel) and $P_{eff}(\nu)$ (right panel) with respect to frequency per pixel. For reference, the overlaid contours trace the 80, 50, 20, 5 and 2.5% points of the instantaneous PB pattern. Note that the range of gray scale levels in the right panel are ∼10 times lower. One dimensional cuts through model PB and effective PB and their first and second derivatives with respect to frequency are shown in Fig. 3. $P_{eff}(\nu)$ is clearly 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 Eq. 9 and 10 is good enough for imaging data that are not sensitive to the higher order frequency dependent effects. This approximation is 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 total power and spectral index mapping. The frequency dependence in $\mathbf{A}(\nu)$ is reduced overall by an order of magnitude, effectively correcting for the frequency dependence of the PB at an accuracy sufficient for imaging. For future, more sensitive telescopes which will be sensitive to second order frequency dependent effects also, $\mathbf{A}_*$ as in Eq. 8 will have to be used.

To the limit that $\mathbf{A}(\nu)$ accurately models $\mathbf{A}^\circ(\nu)$ in Eq. 1, $\mathbf{A}(\nu_*)$ can be used for the reverse transform in Eq. 5 to make images corrected for the frequency dependence of $\mathbf{A}^\circ$. The effective PB for such an image would be P_{ref}, which is independent of frequency. The only frequency dependence left in the image thus made is due to the frequency dependence of the sky brightness distribution – a purely image-plane effect. Such an image is therefore an appropriate input for image-plane minor cycle algorithms. For wide-band wide-field imaging, this allows application of image-plane algorithms like the *MT-MFS* algorithm for mapping sky emission well beyond the inner ∼30% of the PB in an iterative image deconvolution scheme.

5. THE WB A-PROJECTION ALGORITHM

As described in Paper-I, $\mathbf{A}_{ij}$ can be used in the forward transform for baseline labeled by $i - j$ and $\mathbf{A}^\dagger_{ij}$ for the reverse transform to correct for the direction, polarization and antenna dependence of the PB[2]. To include wide-band corrections, it only remains to construct $\mathbf{A}_{ij}$ at the conjugate frequency (Eq. 9) as described in the section 4. Using $\mathbf{A}_{ij}(\nu)$ for the forward transform and $\mathbf{A}^\dagger_{ij}(\nu_*)$ for the reverse transform to process the data at frequency ν then gives a prescription for an algorithm to correct, in general, for all wide-band DD effects described by $\mathbf{A}$ in full-polarization.

An iterative image deconvolution algorithm for wide-field wide-band continuum imaging in full-polarization can be developed as follows:

1. Initialize the model and the residual images I^M and I^R.
2. **Major cycle:** For each baseline $i - j$ and polarization product of interest
    - Compute the model data at frequency ν as
    $$\vec{V}^M_{ij}(\nu) = \mathbf{A}_{ij}(\nu)\mathbf{F}^\dagger\vec{I}^M(\nu) \tag{11}$$
    - Compute the residual data $\vec{V}^R$ as
    $$\vec{V}^R_{ij}(\nu) = \vec{V}^{Obs}_{ij}(\nu) - \vec{V}^M_{ij}(\nu) \tag{12}$$

---

[2] $\mathbf{A}_{ij}$ is a visibility-plane filter, which in general is different for each baseline, and may change with time, frequency and polarization.

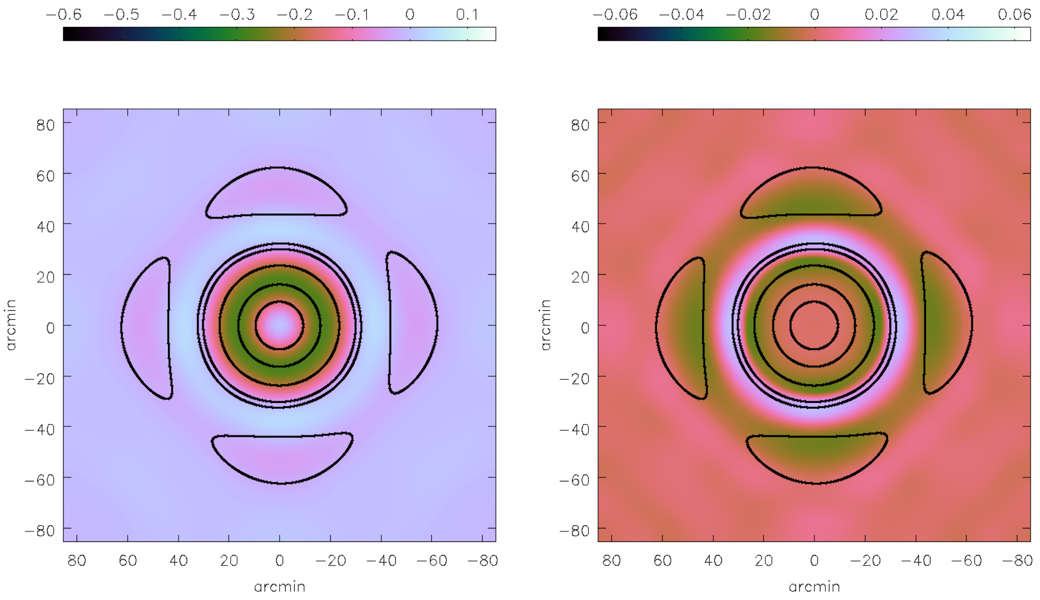

FIG. 2.— Figure shows the derivative as a function of frequency per pixel of the instantaneous PB pattern. The gray scale image in the left panel corresponds to $\boldsymbol{P(\nu)} = \mathsf{F}^\dagger\mathsf{A}(\nu)$ (frequency dependent) while the image in the right panel corresponds to $\boldsymbol{P_{eff}(\nu)}$ (the effective frequency independent gain pattern). Note that the gray scale in the right panel is ~10 times lower. The contours trace the 80, 50, 20, 5, and 2.5% points of the PB at the reference frequency.

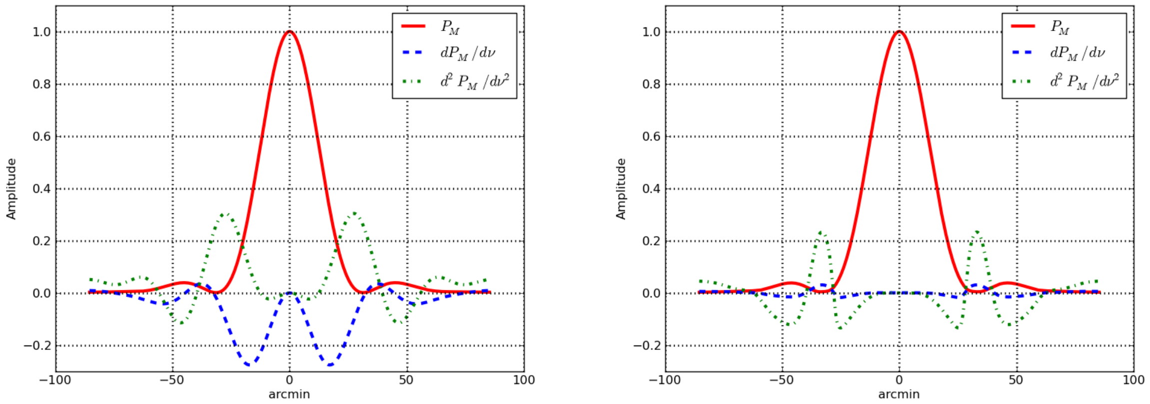

FIG. 3.— The plot in the left panels shows the one dimensional cuts through the PB model at a reference frequency (continuous red line) for the VLA and its first (dashed blue line) and second derivatives (dash-dot green line) with respect to frequency. The plot on the right shows the same cuts through the $\boldsymbol{P_{eff}(\nu)}$ (see Eq. 10) – the effective frequency-independent PB.

- Compute the continuum residual image as

$$\vec{I}^R = \sum_\nu \mathsf{F}^\dagger \mathsf{A}^\dagger_{ij}(\nu_*)\vec{V}^R_{ij}(\nu) \qquad (13)$$

  Also compute I^{wt} (Eq. 6) and normalize $\vec{I}^R$ by I^{wt}. I^{wt} is the PB pattern averaged over time, frequency and antenna-type and represent an estimate of the *ideal* PB.

3. **Minor cycle:** Invoke the appropriate minor-cycle algorithm using I^R to update and build the sky brightness distribution model $I^M(\nu)$.

4. If not converged, go to Step 2.

In practice, various optimizations are possible. When appropriate, A_{ij} can be computed at a single frequency and scaled to the required frequency. Time dependence for El-Az mount antennas can be included by an on-the-fly rotation of $\mathsf{A}_{ij}(\nu)$ by the Parallactic Angle. When the assumption of a homogeneous array is applicable, same A can be used for all the data. Finally, where necessary, A can be pre-computed at appropriate intervals in time, frequency and polarization and for each distinct antenna pairs $i-j$ and cached for later use.

6. RESULTS

The image deconvolution algorithm described in section 5 was tested using simulated wide-band data with 66% fractional bandwidth. The VLA C-array was used for antenna configuration and the observations covered Hour Angle range of $\pm 3^h$. 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

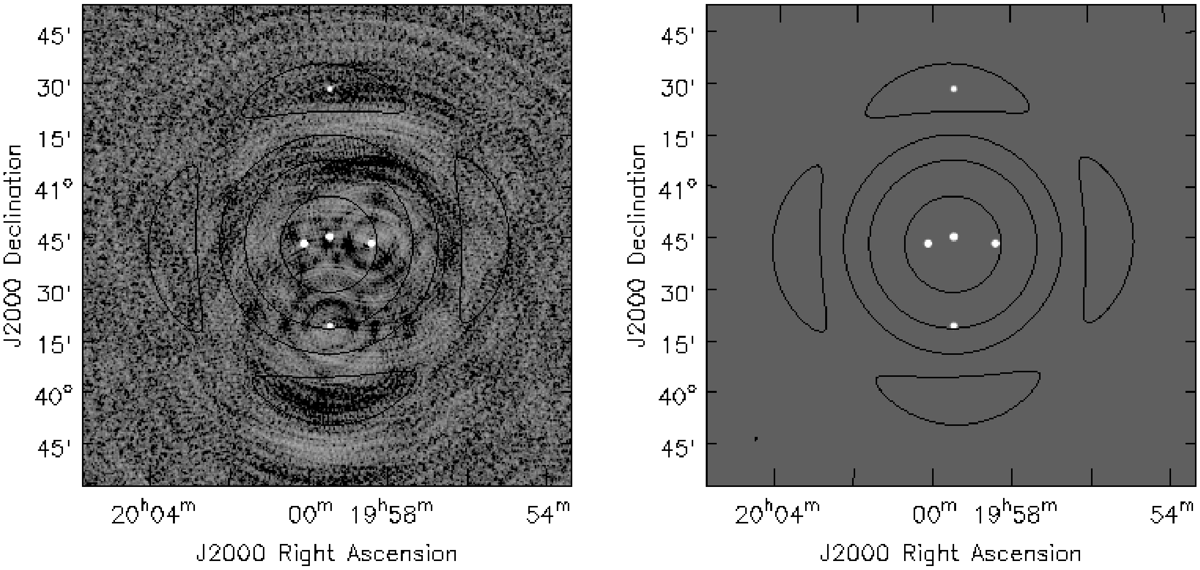

FIG. 4.— This figure shows imaging performance before and after applying corrections for the time and frequency dependence of the PB during imaging. The sky is assumed to have a flat-spectrum, and standard MFS imaging is done. Both restored images are shown at the same gray-scale, stretched to emphasize artifacts. Contours are drawn at the 0.02, 0.1 and 0.5 (HPBW) levels of the time-and-frequency averaged Primary Beam. No noise was added to the simulated visibilities, in order to clearly illustrate the noise-like artifacts produced by time-variable DD-effects.
LEFT : Standard MFS-imaging and deconvolution, using a prolate-spheroidal gridding convolution function. Dominant errors are due to the time and frequency variability of the PB. Off-source RMS : $\mathbf{4 \times 10^{-4} Jy}$, Peak Residual : $\mathbf{1.8 \times 10^{-3} Jy}$
RIGHT : MFS-imaging and deconvolution, using *WB A-Projection* to account for both time and frequency variability during gridding. Off-source RMS : $\mathbf{1.5 \times 10^{-7} Jy}$, Peak Residual : $\mathbf{7 \times 10^{-7} Jy}$

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 primary-beam 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 4 shows deconvolved images produced with and without *WB A-Projection* gridding. This comparison demonstrates that with an accurate model of the Primary Beam, it is possible to correct-for its *time- and frequency-variability* down to numerical precision levels.

The next section discusses the combination of this method with *MT-MFS* for the more realistic situation of a source spectra that are not flat.

7. COMBINATION WITH WIDEBAND-IMAGING (*MT-MFS*)

To test the algorithm described in section 5 with non-flat source spectra, we used the sky brightness distribution as before, but with a spectral index of $\alpha = -0.5$ assigned to all sources such that $I(\nu) \propto (\nu/\nu_o)^{\alpha}$. The location and amplitudes of all sources were the same as in the previous test.

Figures 5 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 an alternate algorithm designed to absorb all frequency-dependence into the minor cycle solver is also shown for comparison. Figure 6 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.

7.1. *MFS + SI (Standard Imaging)*

The image in the top left panel of Fig. 5 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. 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 the PB-shape. 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.

7.2. *MT-MFS + SI*

The image in the top right panel of Fig. 5 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 2-term Taylor-polynomial approximation. Average PB-spectral effects are absorbed into the sky model, and the dominant remaining error is due to the time-variability of the primary-beams. A post-deconvolution correction of the continuum intensity and spectral-index are accurate to within a few percent in intensity and ±0.1 in spectral index out to approximately the half-power point. Beyond this field-of-view, errors increase (to $\pm\mathbf{0.4}$ or more in spectral index) primarily due to the use of a time-averaged PB spectrum for the correction, in regions of the image where the PB-spectrum changes by 100% with time.

7.3. *MT-MFS + A-Projection*

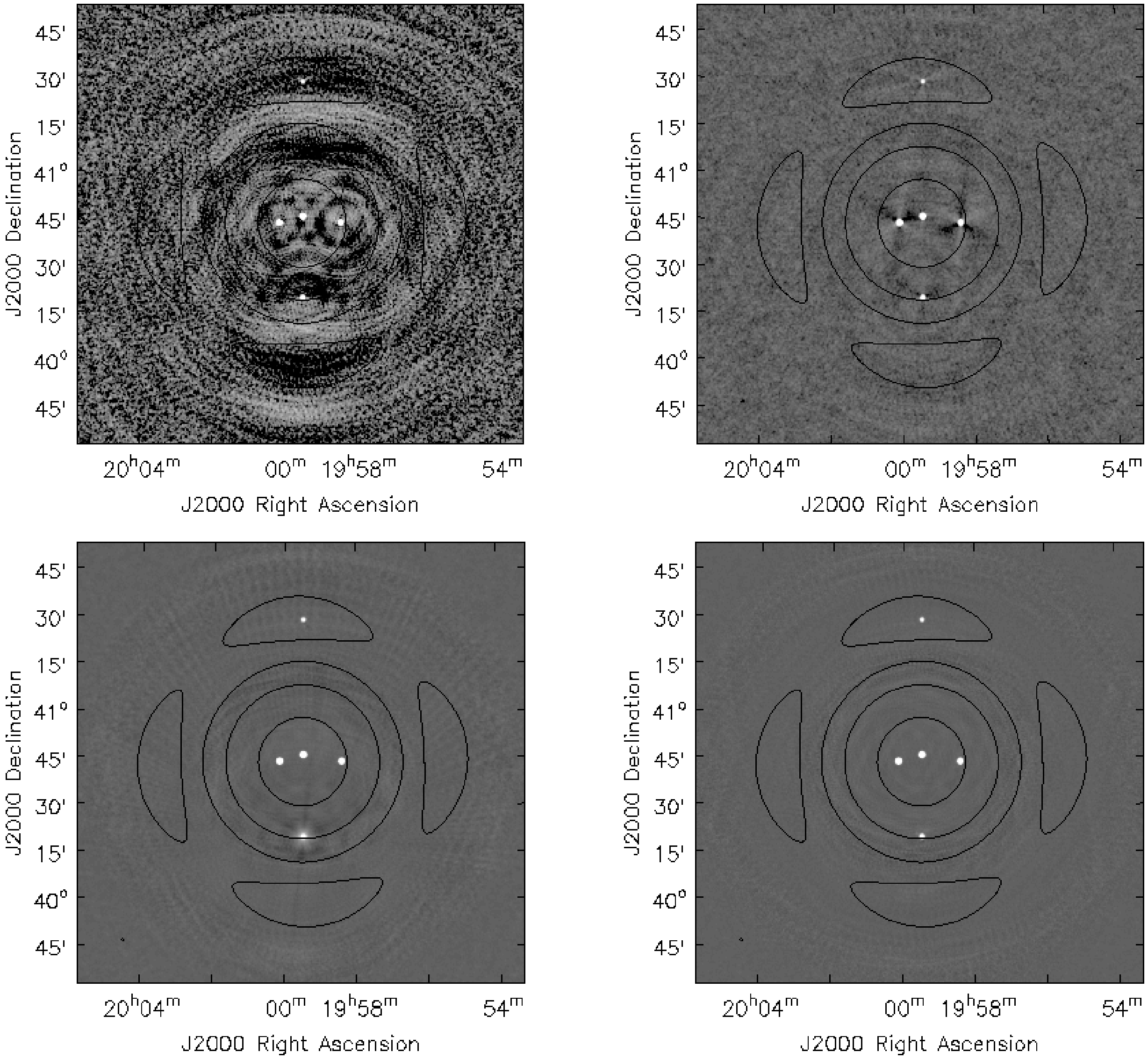

FIG. 5.— These figures compare the imaging performance before and after applying corrections for the time and frequency dependence of the PB during imaging. All restored images are shown at the same gray-scale, stretched to emphasize artifacts. Contours are drawn at the 0.02, 0.1 and 0.5 (HPBW) levels of the time-and-frequency averaged Primary Beam. Results from four algorithms described in Sec. 7 are compared here (MFS+SI, *MT-MFS*+SI, *MT-MFS+A-Projection*, *MT-MFS+WB A-Projection*). RMS and peak-residuals are listed in Table 1.

TABLE 1
TABLE OF QUANTITATIVE COMPARISON OF IMAGES IN FIG. 5

Panel	Algorithm	Description	RMS (Jy/beam)	Peak Residual (Jy/beam)	Comments
Top Left	MFS + SI	Standard Wideband Imaging	**6×10^{-4}**	**2.3×10^{-3}**	Ignore time & frequency dependence. Artifacts due to time and frequency variations of the PB.
Top Right	*MT-MFS*+ SI	Multi-term Imaging with Standard Gridding	**1×10^{-4}**	**5×10^{-4}**	Ignore time dependence. Absorb time-averaged frequency dependence in MT-MFS. Artifacts due to time-variability of the PB.
Lower Left	*MT-MFS+ A-Projection*	Multi-term Imaging with narrow-band AW-Projection gridding	**4×10^{-5}**	**8×10^{-4}**	Account for time variability of PB, and absorb the resulting $\boldsymbol{PB^2}$ 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	**3.5×10^{-5}**	**2×10^{-4}**	Account for PB time- & frequency-dependence in *WB A-Projection*. Account for static sky-frequency dependence in MT-MFS. Minimal artifacts.

The image in the bottom left panel of Fig. 5 is the result of *MT-MFS* in the minor cycle (2 terms), but with the narrow-band A-Projection gridding, using frequency-

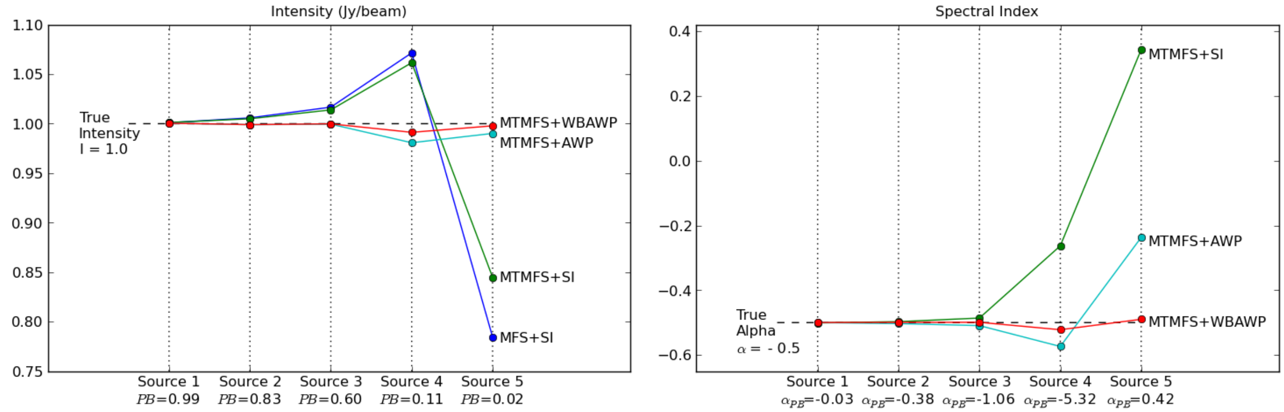

Fig. 6.— This figure compares the accuracy of the PB-corrected intensity (LEFT) and spectral-index (RIGHT) for the five simulated point sources, using the four methods whose results are shown in Figure 5. The labels "`AWP`" and "`WBAWP`" are used for *A-Projection* and *WB A-Projection* in the figure. 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.

dependent gridding-convolution functions $\mathbf{A}^{\dagger}(\nu)$ (as described in Paper-I). The flat-noise [3] image-domain normalization by the average PB does not account for the extra frequency-dependence introduced by the gridding process. The minor cycle therefore models the spectrum of $\boldsymbol{I(\nu)P^2(\nu)}$ in addition to continuum intensity. 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 non-linearity of the $\boldsymbol{P^2(\nu)}$ spectrum away from the pointing center, for which a two-term Taylor-polynomial approximation is insufficient. A run with 3 terms partially alleviates the 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-SNR regions.

Errors also arise from the high time variability of the PB-spectrum, which is ignored because only time-averaged PB-spectra are used for spectral-correction. A post-deconvolution correction of the spectral-index map for $\boldsymbol{P^2(\nu)}$ results in errors at the $\boldsymbol{\pm 0.3}$ level beyond the ~50% point.

7.4. *MT-MFS + WB A-Projection*

The image in the bottom right panel of Fig. 5 is the result of *MT-MFS* in the minor cycle (2 terms), and *WB A-Projection* gridding (section 5). Artifacts around all sources are gone, and the dominant errors are numerical (at the floating-point precision level). No post-deconvolution PB-correction is required, and 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 primary-beam than previously possible. The main difference between this method and all others, is that PB-variability has been corrected for in the data domain, before any averaging of effects (across antennas, time, or frequency) happens in the process of making images 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 Fig. 6 indicating that if the primary beam can be accurately modeled, its time and frequency variability can be corrected for during gridding, resulting in an accurate reconstruction in the minor cycle.

7.5. *Other methods*

1. *Cube imaging (standard or A-Projection transforms) + Cube Clean + Post-deconvolution PB-correction per channel:* This traditional method of wide-band PB-correction is the simplest, and will suffice if the sky-emission is strong-enough to be imaged and deconvolved unambiguously at each frequency channel, if antenna primary beams do not rotate with time, and if all antennas have identical beam patterns. However, the angular resolution of the images is restricted to that of the lowest frequency in the band, and the minor-cycle (deconvolution) cannot benefit from multi-frequency-synthesis.

2. *Cube imaging (standard or A-Projection transforms) + Pre-deconvolution PB-correction per channel (flat-sky) + MT-MFS:* This method is a hybrid of cube-cleaning and *MT-MFS*. Major-cycles use narrow-band *A-Projection* gridding per channel, and a time-averaged PB is divided out from each channel's dirty image. This is the flat-sky normalization of *A-Projection* (Paper-I), applied per frequency. Then, *MT-MFS* continuum residual images are formed in the image-domain as Taylor-weighted averages of the residual images at individual frequencies, and the *MT-MFS* minor cycle proceeds to model the sky spectrum only. Note that strictly, a flat-sky normalization for the minor-cycle does not satisfy a convolution equation, but this method may be robust away from the edge of the PB pattern, especially

---

[3] Definition of flat-noise and flat-sky conventions in A-Projection : The gridding process convolves the measured visibility function with the Fourier transform of the primary beam. In a flat-sky normalization, the resulting image (inverse Fourier transform of gridded visibilities) is divided by $\boldsymbol{PB^2}$. The minor cycle model represents true-sky flux values, but the input to the minor cycle does not strictly satisfy a convolution equation and the noise is not flat across the image. In a flat-noise normalization, the image is divided by $\boldsymbol{PB}$. The minor cycle models the product of the sky and an average PB, but the input to the minor cycle satisfies a convolution equation and the noise remains flat across the image.

for mosaic imaging. It would however limit the depth of each set of minor-cycle iterations, and perhaps require more major cycles. Also, this method has minimal errors only for antennas whose frequency dependence does not vary much with time (e.g. with non rotating PBs of equatorial mount or three-axis mount antennas – although even these antenna mounts are always subject to time-varying antenna pointing errors).

8. CONCLUSIONS

The narrow-band A-Projection algorithm published earlier (Bhatnagar et al. 2008) is a prescription for a full-polarization A-term and its use in an iterative image deconvolution algorithm to correct for the *known* (via a model or measurements) variations in the antenna forward gain pattern as a function of time, polarization and antenna. In this paper we described the *WB A-Projection* algorithm where the frequency dependence of the PB pattern is also included in the A-term. The reverse transform reduces the frequency dependence of the PB by an order of magnitude (Fig. 2) and an *WB A-Projection*-based iterative deconvolution scheme converges. The *WB A-Projection* algorithm is therefore an algorithm for wide-band wide-field imaging in full polarization.

With the large data volumes from modern telescopes and the high cost of software development, run-time performance and algorithmic complexity is now as important for reaching the science goals at an affordable cost as is the imaging performance of the algorithm itself. The approach to wide-band imaging presented here follows the mathematics of imaging and keeps the purely image plane effects (like the frequency dependence of the sky brightness distribution) decoupled from coherence-plane effects (like the effects of the antenna aperture illumination pattern). The models for image-domain and aperture-plane effects ($\boldsymbol{I}$ and M respectively) are also parametrized in their respective *natural domains* of representation, where they can be modeled in the most compressed manner (i.e., requiring least degrees of freedom). This has positive implications, both for designing optimal strategies for calibration and imaging wide-band data as well as on run-time performance of algorithms for calibration and imaging in the presence of direction dependent effects, particularly when applied to the large data sets from modern telescopes.

This decoupling also reduces the complexity of the software implementation of the full imaging solution. As discussed above, since the natural domain of representation of M is the data domain, correction for its effects is also optimally done in the data domain. The image domain is not the natural domain of representation, and attempts to model M and its effects in the image domain invariably leads to a proliferation of the required degrees-of-freedom, which has well understood undesirable effects and leads to non-optimal algorithms – both in terms of imaging performance as well as run-time performance and algorithmic complexity.

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-isoplanetic 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 codebase. We wish to thank the CASA Group for the underlaying libraries. Part of this work was funded by the ALBiUS work package of the European Commission Radionet FP7 program.

REFERENCES

Bhatnagar, S., Cornwell, T. J., Golap, K., & Uson, J. M. 2008, Astron. & Astrophys., 487, 419
Bhatnagar, S., Rau, U., Green, D. A., & Rupen, M. P. 2011, ApJ, 739, L20
Cornwell, T. J. 1995, The Generic Interferometer: II Image Solvers, Tech. rep., AIPS++ Note 184
Ellingson, S. W., Clarke, T. E., Cohen, A., Craig, J., Kassim, N. E., Pihlstrom, Y., Rickard, L. J., & Taylor, G. B. 2009, IEEE Proceedings, 97, 1421
Hamaker, J. P., Bregman, J. D., & Sault, R. J. 1996, Astron. & Astrophys. Suppl. Ser., 117, 137
Parsons, A. R., et al. 2010, AJ, 139, 1468
Perley, R. A., Chandler, C. J., Butler, B. J., & Wrobel, J. M. 2011, ApJ, 739, L1
Rau, U. 2010, PhD thesis, The New Mexico Institute of Mining and Technology, Socorro, New Mexico, USA
Rau, U., Bhatnagar, S., Voronkov, M. A., & Cornwell, T. J. 2009, Proc. IEEE, 97, No. 8, 1472
Rau, U., & Cornwell, T. J. 2011, A&A, 532, A71
Röttgering, H. 2003, New Astronomy Review, 47, 405
Sault, R. J., & Wieringa, M. H. 1994, Astron. & Astrophys. Suppl. Ser., 108, 585
Swarup, G., Ananthakrishnan, S., Kapahi, V. K., Rao, A. P., Subrahmanya, C. R., & Kulkarni, V. K. 1991, CURRENT SCIENCE V.60, NO.2/JAN25, P. 95, 1991, 60, 95
Tingay, S. J., et al. 2012, ArXiv e-prints
Wootten, A., & Thompson, A. R. 2009, IEEE Proceedings, 97, 1463