WIDE FIELD IMAGING PROBLEMS IN RADIO ASTRONOMY

T.J. Cornwell, K. Golap, and S. Bhatnagar

National Radio Astronomy Observatory, PO Box 0, Socorro, NM, 87801

ABSTRACT

The new generation of synthesis radio telescopes now being proposed, designed, and constructed face substantial problems in making images over wide fields of view. Such observations are required either to achieve the full sensitivity limit in crowded fields or for surveys. The Square Kilometre Array [1] now being developed by an international consortium of 15 countries will require advances well beyond the current state of the art. We review the theory of synthesis radio telescopes for large fields of view. We describe a new algorithm, W projection, for correcting the non-coplanar baselines aberration. This algorithm has improved performance over those previously used (typically an order of magnitude in speed). Despite the advent of W projection, the computing hardware required for SKA wide field imaging is estimated to cost up to $500M (2015 dollars). This is about half the target cost of the SKA. Reconfigurable computing is one way in which the costs can be decreased dramatically.

1. THE SQUARE KILOMETRE ARRAY

The goal of the Square Kilometre Array is to provide a large increase in the sensitivity of our observations of the radio sky. The sensitivity will be approximately 50 times better than that of the currently most sensitive telescope, the Very Large Array. There are many technical challenges to be overcome in the design and construction of the SKA. Prime amongst these is the challenge of constructing collecting area for the cost of a few hundred dollars per square meter. There are various concepts for how this can be done, including phased arrays, small hydro-formed or stress parabolic antennas, huge flat antennas fed from a dirigible, and massive antennas constructed in depressions in the ground. Site selection for the array is also complex, since SKA requires both low radio frequency interference and baselines up to 3 000 Km. The digital challenges are substantial as well, ranging from transmission of the signals to a common location, the correlation of many wide band signals, and finally the construction of images at the full sensitivity of the array. It is this latter topic that we consider.

Associated Universities Inc. operates the National Radio Astronomy Observatory under cooperative agreement with the National Science Foundation

2. THE NON-COPLANAR BASELINES ABERRATION

The response of a narrow-band phase-tracking radio interferometer to spatially incoherent radiation from the far field can be expressed by the following relation between the spatial coherence, or 'visibility', $V(u,v,w)$, and the spectral intensity, or brightness, $I(\ell,m)$;

$$V(u,v,w) = \int I(\ell,m) e^{-2\pi i[u\ell+vm+w(\sqrt{1-\ell^2-m^2}-1)]} d\ell dm \quad (1)$$

In this equation, the baseline coordinates, (u,v,w), and direction cosines, (ℓ,m) have their usual definitions [2]. When the term $2\pi w(\sqrt{1-\ell^2-m^2}-1)$ is much less than unity, it may be ignored, and a two dimensional Fourier relationship results. The visibility function is then only a function of (u,v):

$$V(u,v) = \int I(\ell,m) e^{-2\pi i[u\ell+vm]} d\ell dm \quad (2)$$

When term $2\pi w(\sqrt{1-\ell^2-m^2}-1)$ is comparable to or exceeds unity, a two dimensional Fourier transform cannot be used. The value of this term is roughly $\frac{B\lambda}{D^2}$, where B is the maximum baseline length, D is the antenna diameter, and λ is the observing wavelength. Wide-field imaging is affected by this *non-coplanar baselines* effect when observing with small apertures, long baselines, or long wavelengths. In optics terminology, the effect is a *vignetting*: a limitation of the field of view due to the optical system. Sources are moderate distances from the image center are distorted, and sources farther away may vanish altogether.

Cornwell and Perley [3] reviewed the algorithms available to deal with this effect, and presented an algorithm in which the image space is divided into facets, over each of which a two dimensional Fourier transform may be used. The essence of W projection [4, 5] is to project w out of the problem, thus allowing a two dimensional Fourier transform

to a single image to be used. We rewrite equation 1 as a convolution between the Fourier transform of the sky brightness and the Fourier transform of an image plane phase term parametrized by w.

$$V(u,v,w) = \int I(\ell,m)G(\ell,m,w)\, e^{-2\pi i[u\ell+vm]} d\ell dm \tag{3}$$

$$G(\ell,m,w) = e^{-2\pi i[w(\sqrt{1-\ell^2-m^2}-1)]} \tag{4}$$

Applying the Fourier convolution theorem, we find that the key result that:

$$V(u,v,w) = \tilde{G}(u,v,w) * V(u,v,w=0) \tag{5}$$

This holographic relationship results because the original brightness is confined to a two dimensional surface (the celestial sphere).

In practice, interferometers often must measure the visibility function in the three dimensional space (u,v,w). An exception is when all the samples lie in a plane (this occurs, for example, for East-West arrays). In that case, a simple reprojection of the coordinate system restores a two dimensional Fourier transform [6, 7]. In the more usual "non-coplanar baselines case", equation 5 may be used to project the visibility function appropriately.

To understand the physical origin of the non-coplanar baselines effect, we can use a small angle approximation:

$$G(\ell,m,w) = e^{\pi i[w(l^2+m^2)]} \tag{6}$$

$$\tilde{G}(u,v,w) = \frac{i}{w} e^{-\pi i[\frac{(u^2+v^2)}{w}]} \tag{7}$$

The convolution function may be recognized as implementing Fresnel diffraction of the electric field sampled by one antenna to the plane of the other antenna [5], a necessary step that has been previously ignored.

3. ALGORITHMIC DETAILS

Imaging of the sky brightness by radio interferometric arrays typically requires deconvolution of the point spread function arising from the limited sampling of the visibility function in the (u,v) plane [2]. In the case of wide-field imaging, the sampling is still limited but occurs in the (u,v,w) space. To understand how W projection may be inserted into our typical deconvolution algorithms, we must describe how these are structured. We can write the measurement equation as a linear equation:

$$\mathbf{d} = A\mathbf{i} + \mathbf{e} \tag{8}$$

where $\mathbf{d}$, $\mathbf{i}$ and $\mathbf{e}$ are vectors for data, image, and noise, and A is the (non-square) observation matrix. In the usual case of simple radio interferometry, the elements of A are the cosines and sines of the Fourier transform (see equation 2). The observation matrix is usually singular and cannot be simply inverted. Instead, it is generally the case that non-linear iterative methods are used to solve this linear equation. At any one iteration of an iterative deconvolution process, we may use the deficit in the normal equation as a residual or update image:

$$\mathbf{i}^R = A^T(\mathbf{d} - A\mathbf{i}) \tag{9}$$

The typical image size (N by N) for current radio synthesis telescopes ranges between 1,000 by 1,000 to 10,000 by 10,000, and even up to 100,000 by 100,000 for the SKA. The number of sample points (M) can range from 10^3 up to 10^9. Consequently, direct solution of 9 is not feasible. direct application. Hence we usually adopt a two stage process in which the matrix A^TA is approximated in two different ways.

To start with, the CLEAN [8] algorithm is used to solve an approximate convolution equation for the update $\mathbf{\Delta i}$

$$\mathbf{i}^R = \left(A^TA\right)\mathbf{\Delta i} \tag{10}$$

In this step, A^TA, is approximated by a Toeplitz matrix with elements far the diagonal set to zero. CLEAN works iteratively by finding and removing the brightest point source in $\mathbf{i}$ from the peak in $A^T(\mathbf{d} - A\mathbf{i})$. This algorithm is highly effective for simple objects, but for more complex objects, CLEAN is less successful, and it has recently been extended to encompass multi-scale approaches [9]. There remains much room for improvement in this area.

Given an updated estimate of the image i, we must apply the product A^TA, and also calculate A^TD. These steps also require approximations. For small fields of view, the usual practice is to use convolutional gridding in conjunction with an FFT. Thus A is approximated by TF where F is an N^2 Fourier transform, and T is a block diagonal M by N^2 matrix, chosen as a compromise between computational costs (small number of non-zero elements) and aliasing (large number). Typically, T is constructed from a tapered prolate spheroidal wavefunction of support 7 by 7 or 9 by 9 elements (See the lecture by Schwab and Bridle in [10]).

In the case of wide field imaging using W projection, the observation matrix A is still compased of sines and cosines (see equation 1). The extra phase term in the transform prevents application of narrow field approximations. Itstead, the observation matrix is approximated by RTF where R is an M by N^2 matrix representing W projection. In practice, multiplication by RT is accomplished using a precalculated tabulation of the Fourier transform of $G(\ell,m,w)$ multiplied

by the appropriate spheroidal function. Calculation of both the predicted data $A\mathbf{i}$, and the residual image $A^T(\mathbf{d} - A\mathbf{i})$ is then straightforward.

Most of the computational cost occurs in the application of RT and its transpose, and is directly determined by the number of visibility samples and the size of the convolution kernel. Taking the field of view to be λ/D, we find that the number of pixels in the support region along each of the the u and v axes goes as $B\lambda/D^2$.

The costs for the standard facet based gridding [3] go as the total number of facets (*i.e.* the product of the number along each of the two spatial axes). We then have that the costs per sample go as:

$$t_{facets} = N^2_{facets}.N^2_T.t_{single} \quad (11)$$

$$t_{wproject} = 2(N^2_{wproject} + N^2_T).t_{single} \quad (12)$$

In this equation, N_T is the support of the normal gridding convolution function T in one axis (typically 9), and t_{single} is the time to grid a single sample to a single grid point. N_{facets} (the number of facets in one axis) and $N_{wproject}$ (the typical size of the R blocks) are both proportional to $B/(\lambda D^2)$ but with different proportionality constants. We have assumed that the sizes of the T and R add in quadrature. Note also that R is necessarily complex. If we take N_{facets} and $N_{wproject}$ to be roughly equal, then for large fields of view, the ratio of these times is roughly the total number of points in the blocks of T. Since this is typically 7 by 7 or 9 by 9, the *asymptotic* speedup is between 25 and 50. Allowing for different proportionality constants, *we could conservatively expect at least an order of magnitude speed advantage for w projection*.

More details of the W projection algorithm are to be found in [4, 5].

4. A SYNTHETIC EXAMPLE

To demonstrate the w projection algorithm, we simulated a low frequency (74MHz) observation of a typically full field. Data corresponding to a 74MHz VLA C-configuration full synthesis were calculated using analytical transforms, and should thus be fully accurate to machine precision. The resulting deconvolved images are shown in Figure 1.

The speed difference is substantial : 30488s for the facet based algorithm compared to 5419s for W projection. If the algorithms are configured for similar dynamic range, the difference in speed is usually about an order of magnitude.

5. COMPUTATIONAL COSTS FOR THE SQUARE KILOMETER ARRAY

Cornwell [11] has recently estimated the computational load for sensitivity limited full field continuum limited imaging at 1.4GHz with the Square Kilometer Array [1]. The costs scale as $\lambda B^3 D^{-8}$. A representative computing load to keep up with real time would be about 150PFlops. Scaling with Moore's Law with a doubling time of 18 months, Cornwell estimates that the dollar cost to be about $500M in 2015 - about half the target cost of the SKA. Without W projection, the costs would be about ten times as much.

Highly efficient parallelization is already assumed in this estimate, and there remains much work to be done in this respect. How then can we make further improvements? There are many options. Increasing the antenna diameter helps immensely but increases the antenna construction costs. Lonsdale *et al.* [12] advocate strong averaging of the data to limit the field of view, but they note that this probably has bad effects on the various self-calibration schemes likely to be used.

One promising computational strategy is to deploy W projection (and other related algorithms) onto reconfigurable computers. Compton and Hauck [13] provide an excellent overview of reconfigurable computers. The term "Reconfigurable Computing" is very broad and accomodates many different types of architecture and strategies. From our point of view, the most important is to augment traditional microprocessors with reconfigurable computational elements such as Fully Programmable Gate Arrays. The microprocessor does the general computing, and the FPGA is configured for special purpose operations to be executed within a few clock cycles. The gain in speed depends on the mix of operations but it seems likely that W projection is well suited to such an architecture. The convolution step would be implemented in the FPGA, thus returning the load roughly to that expected for coplanar arrays. Computers with this architecture are now coming to market (*e.g.* Cray XD1.) It thus seems plausible that with this approach, the costs could be reduced by up to an order of magnitude. The key advantage of W projection in this regard is that the global aberration, the non-coplanar baselines effect, can be countered by local operations in the Fourier space. As such, it provides a valuable model for the ongoing development of fast imaging algorithms to deal with other wide field effects [14].

6. REFERENCES

[1] SKA Consortium, "http://www.skatelescope.org," Tech. Rep., Square Kilometre Array Consortium, 2004.

[2] A.R. Thompson, J.M. Moran, and G.W. Swenson, *Interferometry and synthesis in radio astronomy*, Wiley, New York, 2001.

[3] T.J. Cornwell and R.A. Perley, "Radio-interferometric imaging of very large fields - The problem of non-

coplanar arrays," *Astron. Astrophys.*, vol. 261, pp. 353–364, 1992.

[4] T.J. Cornwell, K. Golap, and S. Bhatnagar, *EVLA memo 67*, National Radio Astronomy Observatory, http://www.nrao.edu, 2003.

[5] T.J. Cornwell, K. Golap, and S. Bhatnagar, "The non-coplanar baselines effect in radio interferometry," *Astron. Astrophys.*, vol. (submitted), 2005.

[6] W.N. Brouw, *Data Processing for the Westerbork Synthesis Radio Telescope*, Ph.D. thesis, Univ. of Leiden, 1969.

[7] R.N. Bracewell, "Inversion of Non-Planar Visibilities," in *Indirect Imaging. Measurement and Processing for Indirect Imaging. Proceedings of an International Symposium held in Sydney, Australia, August 30-September 2, 1983. Editor, J.A. Roberts; Publisher, Cambridge University Press*, 1983, p. 177.

[8] J.A. Hogbom, "Aperture Synthesis with a Non-Regular Distribution of Interferometer Baselines," *Astron. Astrophys. Suppl.*, vol. 15, pp. 417–+, 1974.

[9] T.J. Cornwell and M.H. Holdaway, "Multi-scale clean deconvolution," *Astron. Astrophys.*, vol. (submitted), 2005.

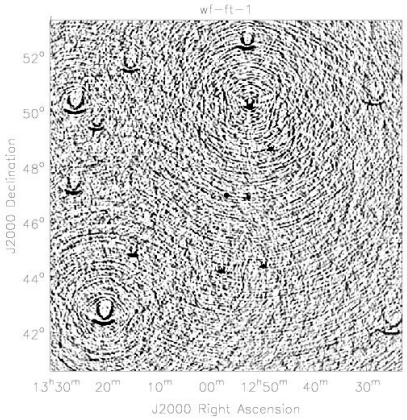
[10] G. B. Taylor, C. L. Carilli, and R. A. Perley, Eds., *Synthesis Imaging in Radio Astronomy II*, 1999.

[11] T.J. Cornwell, "Ska and evla computing costs for wide field imaging," *Experimental Astronomy*, vol. (in press), 2005.

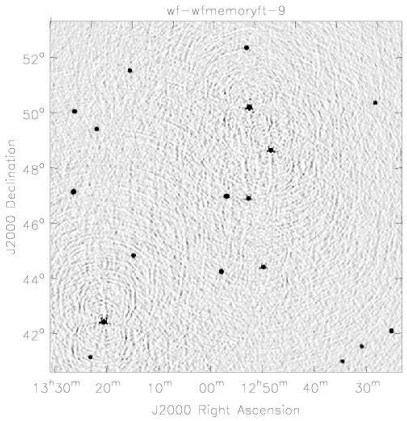
[12] Lonsdale C.J., S. Doeleman, and D. Oberoi, "Imaging strategies and postprocessing computing costs for large-n ska designs," Tech. Rep. SKA memo 54, Square Kilometre Array, 2004.

[13] K. Compton and S Hauck, "Reconfigurable computing: A survey of systems and software," *ACM Computing Surveys*, vol. 34, no. 2, pp. 171–210, 2002.

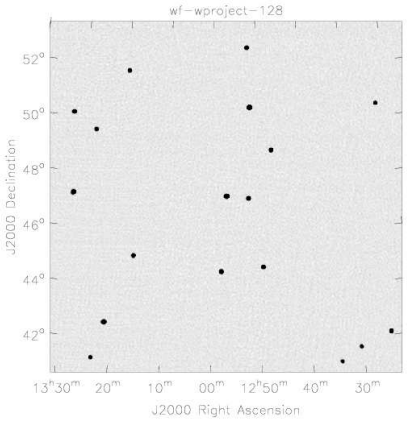
[14] S. Bhatnagar, T.J. Cornwell, and K. Golap, *EVLA memo 84*, National Radio Astronomy Observatory, http://www.nrao.edu, 2004.



Standard Fourier Transform



Facet-based algorithm (9 x 9)



w projection (128)

Fig. 1. Clean images for a VLA 74MHz simulation. The brightness range is -5 to +50 mJy/beam, and the peak brightness should be 47.2Jy. The peak sidelobes around the brightest sources in the uvw-space facets image are about 0.3%. Calculation of these images took 784s, 30488s, and 5419s respectively.