EVLA memo 76

EVLA and SKA computing costs for wide field imaging

***T.J. Cornwell, NRAO*[1]**

tcornwel@nrao.edu

1. Introduction

Perley and Clark (2003) have recently derived a cost equation for synthesis arrays that includes the computing costs to counteract the non-coplanar baselines aberration. One conclusion from their work is that the cost equation should include a cubic term in the number of antennas. Consequently the minimum cost antenna diameter for fixed collecting area is increased over that derived while ignoring the costs of non-coplanar baselines. To determine how much the diameter is increased, the actual scaling coefficient must be known. In this document, I use AIPS++ to estimate the scaling coefficient of this cubic term.

The simulation in this memo is not designed to be overly realistic – the computational resources required are beyond those available.

2. Scaling behavior

Perley and Clark analyzed the time taken to clean an image afflicted by non-coplanar baselines smearing using the facet based algorithms (see Cornwell, Golap, and Bhatnagar, 2004 for more on the taxonomy of wide field imaging algorithms).

The w projection algorithm in AIPS++ outperforms the facet-based algorithms in AIPS, and AIPS++ by about an order of magnitude (Cornwell, Golap, and Bhatnagar, 2003, 2004), and so we choose to use it for these simulations. Calculation of the work required for w projection is straightforward. Only one image is made. The data are gridded onto the $(u,v,w=0)$ plane using a w dependent convolution function. Calculation of the gridding function is data independent, scaling as the number of w planes used. The area of the gridding function in pixels goes as $\lambda B / D^2$. The number of channels and the integration time both scale as B / D, and the number of baselines scales as $1 / D^2$. Hence the number of operations required to grid the data goes as B^3 / D^6 or $B^3 N^3$. The facet based algorithms have the same scaling but with a coefficient about 10-50 times larger. There is good reason to believe that all existing algorithms for non-coplanar baselines must have the same scaling, but with possibly different coefficients. At the moment, w projection is the best but it is possible that other better algorithms might be developed. In comparison, the scaling ignoring wide field effects is B^2 / D^4 or $B^2 N^2$. It is not too

---

[1] The National Radio Astronomy Observatory is operated by Associated Universities, Inc., under cooperative agreement with the National Science Foundation.

surprising that according for the third axis, w, increases the scaling behavior from square to cube.

The total cost also goes as the number of major cycles, which in turn is determined by the maximum exterior sidelobe (exterior to the beam patch in the Clark minor cycle). Although the rms sidelobe level should decrease as $1/N$ or D^2, the solid angle goes inversely as this behaviour. Hence the peak exterior sidelobe (which determines the number of major cycles in the Clean) will have much weaker dependence on the antenna diameter. To first order, we might expect that the number of major cycles is set by the desired dynamic range and not by the beam properties. We will investigate this point in the simulations but for the moment, we conclude that the total cleaning cost should go as the gridding cost.

3. Simulations

The simulations were performed using the AIPS++ (version 1.9, build 549) `simulator` and `qimager` tools, running on a Dell 650 Workstation (dual processor Xeon 3.06Ghz processors, 3GB memory, Redhat Linux 7.2, special large memory kernel). The script is given in appendix A.

Table 1 Details of simulation

Total collecting area	Equivalent to 1600 12m antennas within 10km
Antenna diameter	12.5, 15, 17.5, 20, 22.5, 25, 27.5, 30, 32.5, 35, 37.5, 40m
Number of antennas	Set by antenna diameter to achieve fixed collecting area
Array configuration	Random antenna locations
Frequency	1.4GHz, 50MHz bandwidth.
Observing pattern	60s at transit, integration time 10s, scaling as antenna diameter
Number of spectral channels	8 channels maximum, scaling inversely with antenna diameter
Array latitude	34 deg N
Source declination	45 deg
Source details	250 point sources per square degree with source count index –0.7. Peak strength = 1 (but two sources may be in same pixel).
Antenna illumination pattern	Unblocked, uniformly illuminated
Synthesis imaging details	1.5 arcsec pixels, natural weighting, with 4 arcsec taper.
Number of w planes in w projection algorithm	256
Clean details	Cotton-Schwab algorithm, loop gain 0.1, maximum 1000000 iterations, stopping threshold 0.0001, cyclespeedup=10000
Resolution	4.5 by 4.5 arcsec

Antenna diameter	Fresnel number	Ant	Int	Chan	Sizes... Sources	Image	Vis	MS	Times to ...			Clean Comps	Cycles	PSF min	Outer	Image properties minimum	robust
m						pixels	records	GB	construct	predict	clean						
12.5	13.4	1600	10	16	1000	5000	1.28E+07	11.16	4034.5	3404.8	38643.7	49960	7	-0.001	0.029	-4.29E-06	3.35E-07
15.0	9.3	1111	8	13	694	4166	4932840	3.802	524.5	591.3	9893.9	34223	6	-0.001	0.028	-8.53E-04	4.77E-07
17.5	6.9	816	7	11	510	3570	2327640	1.558	203.6	328.9	4176.7	25148	6	-0.002	0.037	-4.64E-06	5.81E-07
20.0	5.3	625	6	10	390	3124	1170000	0.664	71.0	142.4	1727.8	18690	5	-0.003	0.028	-6.72E-06	7.66E-07
22.5	4.1	493	5	8	308	2776	606390	0.285	36.2	58.6	1121.9	15132	6	-0.003	0.036	-1.68E-04	8.27E-07
25.0	3.4	400	5	8	250	2500	399000	0.187	23.3	44.2	858.8	12290	6	-0.004	0.051	-8.21E-06	9.61E-07
27.5	2.8	330	4	7	206	2272	217140	0.092	15.0	18.7	278.6	9969	6	-0.005	0.049	-8.25E-06	1.12E-06
30.0	2.3	277	4	6	173	2082	152904	0.057	9.3	10.2	204.0	8351	6	-0.006	0.042	-1.51E-04	1.32E-06
32.5	2.0	236	3	6	147	1922	83190	0.033	5.6	9.0	197.5	8054	6	-0.007	0.039	-1.39E-05	1.85E-06
35.0	1.7	204	3	5	127	1784	62118	0.021	3.9	4.1	135.1	16176	8	-0.008	0.033	-4.02E-05	4.55E-06
37.5	1.5	177	3	5	111	1666	46728	0.016	2.9	3.1	70.5	7596	7	-0.009	0.051	-2.06E-05	2.75E-06
40.0	1.3	156	3	5	97	1562	36270	0.013	2.2	4.8	92.3	6882	7	-0.010	0.054	-2.63E-05	3.20E-06

Table 2 Simulation results.

4. Results

The quantitative simulation results are given in Table 2. For the PSF, we show the minimum point, and the maximum "exterior" sidelobe (*i.e.* that outside the beam patch used in the minor cycle of the Clean). The image properties shown are the minimum (affected by cleaning errors around bright sources), and the median absolute deviation from the median (a robust statistic showing the off source error level). The robust statistic scales inversely as the number of antennas, as expected. The behavior of the minimum sidelobe level is also as expected (going as $1/N$), and the maximum exterior sidelobe is more or less independent of antenna diameter, as theorized above. The number of major cycles is *roughly* the number of decades of dynamic range required. Hence we have not been too concerned about the stopping point of the deconvolution. Going a factor of ten deeper requires only another major cycle, and is therefore only a 15-20% effect.

The times (wall clock) for measurement set construction, model data prediction, and cleaning are plotted as a function of antenna diameter in figure 1. In addition, we show predicted times for the cleaning scaling by the cube of the number of antennas. This confirms the result of Perley and Clark (and of section 2) that the cleaning time does scale as the cube of the number of antennas or inversely as the *sixth* power of the antenna diameter (for constant collecting area).

The Fresnel number $\lambda B / D^2$ is close to unity for the large diameters but significantly larger for small diameters. Thus the non-coplanar baselines effect is of marginal importance for the larger antennas in this simulation. In this situation, w projection reduces gracefully to ordinary imaging, apart from the constant cost step of calculating the convolution function. Even when the Fresnel number is large, the main effect for a snapshot is a coordinate distortion. However, we have included full w projection since it will be needed for longer integrations and baselines.

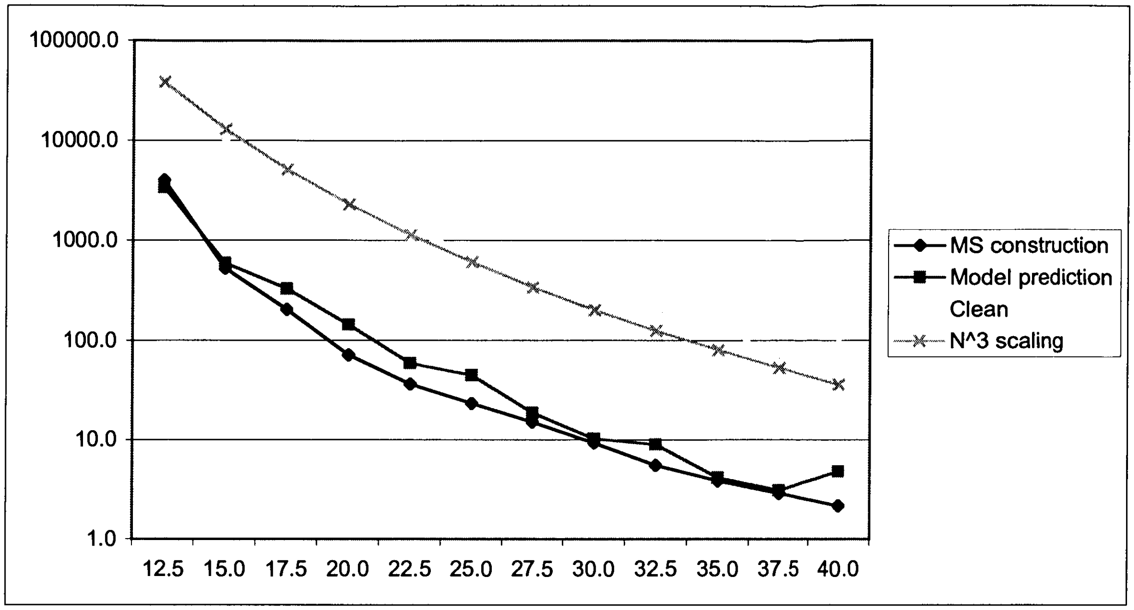

Figure 1 Wall clock times as a function of antenna diameter. The light blue curve (with x's) shows scaling by the cube of the number of antennas.

The stability of existing Clean algorithms for large numbers of resolution elements should be improved. The Clark algorithm that underlies the Cotton-Schwab algorithm has not been tuned for this context (or for modern processors). As a result, for larger source density (not shown here), the algorithm is prone to find spurious sources or to even diverge. A systematic study of tuning and convergence for relatively poor sampling (as occurs for large diameter antennas, despite the mitigating influence of the smaller field of view) would be timely. Even so, this instability reflects a real problem – the poor conditioning of the deconvolution for large antennas.

5. The scaling laws

We have confirmed the scaling behavior described in the Perley-Clark memo. We can also comment on the absolute value of the scaling coefficient.

For the SKA, the scientific specification on collecting area is 50% within 5km (Jones, 2003). Assuming Moore's Law for the cost of processing, we choose to write the scaling law as:

$$C_{SKA} \sim C_{12.5m}\left(\frac{0.1}{\nu}\right)\left(\frac{f}{0.5}\right)^{2}\left(\frac{B}{5km}\right)^{3}\left(\frac{D}{12.5m}\right)^{-6}\left(\frac{\lambda}{0.2m}\right)^{0.7}\left(\frac{\Delta\nu}{500MHz}\right)2^{\frac{2(2010-t)}{3}}$$

Or, for a constant collecting area:

$$C_{SKA} \sim C_{12.5m} \left(\frac{0.1}{}\right) \left(\frac{f}{0.5}\right)^{2} \left(\frac{B}{5km}\right)^{3} \left(\frac{N}{1600}\right)^{3} \left(\frac{}{0.2m}\right)^{0.7} \left(\frac{}{500MHz}\right) 2^{\frac{2(2010-t)}{3}}$$

The filling factor f is the fraction of collecting area within the baseline B. The efficiency of processing, η, is both very important and as yet unknown. It includes, for example, the cost of correcting for source spectral effects, and antenna primary beams, and the efficiency of parallel processing. We estimate a typical value for this efficiency as about 10%.

For a 12.5m antenna design, the ratio between observing time and real time in our simulation is roughly 750 so the efficiency is about 0.13% (for 25m, the ratio is ~ 17 efficiency is ~ 6%). The computer used in the simulations cost about $8000 in 2003. Solving, we find that the coefficient $C_{12.5m}$ is about $12M.

Since the antenna size for the EVLA has been chosen, we write the EVLA cost equation is:

$$C_{EVLA} \sim C_{A} \left(\frac{0.1}{}\right) f^{2} \left(\frac{B}{35km}\right)^{3} \left(\frac{}{0.2m}\right)^{0.7} \left(\frac{}{500MHz}\right) 2^{\frac{2(2010-t)}{3}}$$

Scaling appropriately from C_{SKA}, we find that C_A is $80K.

6. Implications for the EVLA

For the EVLA A configuration (baselines up to 35km), the cost of computing hardware required for wide-field processing is $80K in 2010, $8K in 2020. This is quite modest and not dissimilar to previous estimates. In phase II, the EVLA will have baselines up to 350km baselines, and the costs would be $80M (2010), and $8M (2020). Scientifically, we expect that such observations would be fairly rare and so the actual required duty cycle would be low.

Algorithm improvements help. The advent of w projection brings the cost down by about an order of magnitude, which is equivalent to a decade of Moore's law gains. Poor symmetry and stability of *e.g.* primary beams and pointing will hurt a lot by decreasing the efficiency (see *e.g.* Cornwell, 2003).

In addition, there remains a lot of software development to be done. It is clear that parallel processing using tens or hundreds of processors will be required to handle EVLA data. There has been relatively little work on parallelization of synthesis imaging algorithms. This has to be improved.

Finally, operational models of the EVLA will affect the cost estimates. If the most demanding observations occur infrequently and turnaround can be a few days or weeks (as is now often the case) then the computing costs can be reduced proportionately.

7. Implications for the SKA

The canonical case of SKA imaging with the 5km baselines at 20cm would require only $12M in 2010, and $1.2M in 2020. However, for the more interesting case of the 35km baselines, the costs rise to $10B and $1B. Increasing the antenna diameter to 25m brings the costs down to $140M and $14M. For 350km baselines, the cost increases to $140B and $14B, even with 25m antennas!

The cost disadvantage for small antennas is therefore large. For example, if the computing hardware budget were not to exceed $100M, we should not attempt 1", 20cm continuum imaging with 12.5m antennas until 2031. For 25m antennas, the corresponding date is 2022.

A key point is that the scaling behavior is very dramatic, as the cube of the baseline and the inverse sixth power of the antenna diameter. In comparison, the effects of more bandwidth and lower frequencies are quite mild. Thus the SKA computing budget will be determined by the emphasis placed on baselines in the range of 10km and longer.

The same comments about algorithm and software development made above for the EVLA apply in equal measure to the SKA.

Our major conclusion is that computing hardware is a major cost driver for the SKA, and much more attention is required before the concept cost estimates can be viewed as accurate. In addition, simulations should start to include the non-coplanar baselines effect, so as to raise awareness of the importance of the effect for SKA. In the specific case of the LNSD concept, the cost minimization with respect to antenna diameter should be repeated with these more accurate computing costs included.

References

Cornwell, T.J., 2003, EVLA memo 62, http://www.nrao.edu/evla.

Cornwell, T.J., Golap, K., and Bhatnagar, S., 2003, EVLA memo 67, http://www.nrao.edu/evla.

Cornwell, T.J., Golap, K., and Bhatnagar, S., 2004, *submitted.*

Perley, R.A., and Clark, B.G., 2003, EVLA memo 63, http://www.nrao.edu/evla.