SKA MEMO 46. 18 Feb 2004

SKA Imaging

Melvyn Wright

Radio Astronomy laboratory, University of California, Berkeley, CA, 94720

ABSTRACT

In this memo, we explore the imaging requirements and data processing options for the large N SKA. We discuss imaging from the sampled cross correlation function and direct imaging by beam formation. Cross correlation of all antennas provides the most complete sampling of the incident wavefront and allows imaging the full field of view of the individual antennas. Extrapolation of existing and planned radio astronomy correlators suggests that a 4000 antenna, 1 GHz bandwidth correlator is feasible by 2020. Direct image formation requires $\sim$ GHz data processing to phase the signals from all the antennas over a 10^6 pixel image. The calibration must be made in close to real time with the derived calibration parameters fed back into the real time system for multiple phase centers. The large data rate and data processing requirements suggests that the SKA should produce final, calibrated images as its normal output.

1. Introduction

A major goal of observational astronomy is to form images of the sky brightness distribution, $I(s, \nu, p, t)$, as a function of position, frequency, polarization, and time. In order to acquire sufficient angular resolution, radio astronomers have long been obliged to build radio telescopes from arrays of antennas leading to the well established and highly successful aperture synthesis imaging techniques in use at radio observatories worldwide.

Aperture synthesis imaging in radio astronomy has evolved from measurements with two-element interferometers. An interferometer measures the coherence function of the wavefront which can be used to derive simple source structures for compact sources, e.g. positions, angular diameters and double source structures. Measurements of the cross-correlation function between pairs of antennas at different spacings provides samples of the Fourier transform of the sky brightness distribution. With a sufficiently well sampled set of calibrated cross-correlations, we can perform a Fourier transform to make an image of the sky brightness distribution. The development of earth rotation aperture synthesis (Ryle 1962), with arrays of antennas has enabled the production of radio images with higher angular resolution than attainable at much shorter wavelengths. The success of aperture synthesis arrays has been enhanced by the development of sophisticated image processing techniques including self-calibration of the measured cross-correlation function using images of the sky brightness, and deconvolution to remove the effects of the finite sampling of the cross-correlation. For a concise review of the development of radio astronomy interferometers the reader is referred to Thomson, Moran and Swenson (TMS), 2001.

Despite these impressive achievements, there are significant problems extrapolating existing aperture synthesis techniques to the SKA. In this memo we discuss some of the problems for arrays with large numbers of antennas, and the problems associated with wide field imaging. At long wavelengths, there are also significant image plane corrections associated with finite isoplanicity, polarization corrections across the primary beam, etc. which preclude a simple Fourier transform relationship between the visibility function and the sky brightness distribution.

2. SKA specification

This section summarizes the SKA specifications relevant for this memo. These specifications were taken from the SKA Science Requirements version 6, 1 December 2003 (D.L.Jones, 2003)

Correlator: The current SKA specifications call for an input bandwidth of 25% at observing frequencies below 16 GHz, and 4 GHz per band above 16 GHz. Each band is to have 10^5 spectral channels with a minimum accumulation interval 0.5s. (The goal is 10^6 channels with an accumulation interval 0.05s.)

Imaging: The SKA should allow imaging up to 10^4 separate regions within the FoV with at least 10^5 beam areas at the maximum angular resolution. The image quality is specified by a dynamic range 10^6 and an image fidelity 10^4 between 0.5 and 25 GHz.

Configuration: The current specification has 20% of the collecting area within 1 km, 50% within 6 km, and 75% within 300 km, with maximum baselines of $\sim$ 3000 km. An array configuration with some sort of hierarchical clustering of antennas at large distances from the compact central cluster is envisaged as it would be cheaper and easier to build and maintain.

A strawman design which clusters the antennas into stations with correlations between the stations of phased antennas presents some problems:

1. The Fourier plane is less well sampled leading to an increased sidelobe level in the synthesised beam (this is true, independently of whether synthesised beams are formed from Fourier transform of the sampled correlation function, or by direct beam formation). The higher sidelobe level increases the data processing required by a factor N_{sta}, the number of antennas per station, as noted by Perley and Clark (2003).

2. The station beam forms the effective primary beam pattern for the array. The station beam has a sidelobe level $\sim 1/N_{sta}$. Moreover the station beam will be different for each station, depending on the configuration and the antenna weighting, and will be time variable due to projection

geometry, interference mitigation measures, antenna failures and calibration errors etc.

Thus, a correlation between stations of phased antennas will degrade both the primary beam, and the synthesised beam. For an SKA with ~ 4000 antennas, a compromise would be ~ 64 stations each with ~ 64 antennas giving a sidelobe level $\sim 1.6\%$ in both station and array beams. It is not clear if a dynamic range 10^6 or an image fidelity 10^4 can be obtained from such array configurations. An alternative model where the signals from all antennas are correlated, was explored in SKA memo 16 (Wright, 2002). The current plan is to correlate the inner ~ 2000 antennas, and to correlate phased stations at longer baselines. The large correlator needed to correlate 1000's of antennas, and the high data rates are both seen as problems.

3. Correlators

Figure 1 plots the total bandwidth of some existing and planned radio astronomy array correlators versus the year completed (or planned to be so). The solid line shows Moore's Law plotted to pass through an SKA correlator for 4400 antennas with 1 GHz bandwidth. Most of the planned correlators have continued to be scaled back in bandwidth and/or completion date since this figure was first presented in SKA memo 16. Using only the currently working correlators, the bandwidth has essentially remained constant since 1980. The planned correlators, shown in red, require a large increase in bandwidth to catch up with Moore's law. Assuming this is accomplished, then an SKA correlator capable of cross correlating 4400 antennas seems plausible by 2020. Most correlators significantly lag the current technology by the time they are put into service; the average time from conception to completion ~ 10 years.

A corollary is that for post processing the correlated *uv* data using general purpose computers which have kept up with Moore's law, it is now easy to keep up with the correlator data rates, but life will become more difficult if correlator bandwidths increase by $\sim 10^4$. The right axis gives the data rate assuming 2000 spectral channels per GHz of bandwidth, 4 bytes per channel, a 25% overhead, and a 10s sample interval. For the SKA with *only* 2000 spectral channels, the data rate is $\sim 10^7$ kbytes/s.

Figure 2–4 show multichannel imaging simulations for a range of spectral channels and image sizes. The plotted points show the timing for calibration and imaging on a single processor PC. The data rates are $\sim$ 1–3 Mbytes/s. If these data processing rates follows Moore's law over the next 20 years, we will get a $\sim 10^4$ increase, which will allow us, in theory, to process 2000 spectral channels from the SKA on a commodity PC by 2020 ! The problem then appears to be getting the data from the correlator into the PC, and getting the data out of the PC and into the astronomer.

4. Data Rates

The total input bandwidth is $N \times B$, where N is the number of antennas and B is the bandwidth. The output data rate from a correlator is $N^2/2t$, where t is the averaging interval in seconds. The time averaged output from the correlator typically also requires more bits than the input sampled bandwidth. Accumulating the data as a correlation function reduces the data rate if

$N < 2fBt$ ————————-(1)

where f, $\sim 4/32$, is the ratio of the number of bits per sample into and out of the correlator.

The input bandwidth is channelized, to provide spectral resolution, to facilitate interference rejection, and to reduce bandwidth smearing by using multifrequency synthesis (MFS), so the relevant bandwidth, B, in equ. 1 is really the spectral channel width. E.g. for extragalactic HI at 1.4 GHz, we want $B \sim 20$ km/s = 100 kHz, and for cold Galactic clouds $B \sim 0.2$ km/s = 1 kHz.

The Nyquist sample interval for the spatial coherence function is $D_{ant}/2$, where D_{ant} is the antenna diameter. For mosaic observations we should sample the data for each pointing center at the Nyquist rate. This leads a sampling rate $= D_{max}/(D_{ant}/2) \times N_p \times sdot$ s^{-1}, where D_{max} is the array diameter, N_p is the number of pointing centers which are time multiplexed, and $sdot = 7.27$ 10^{-5} s^{-1}.

For the ATA, with an array of ~ 300 6m-diameter antennas and a maximum baseline ~ 900 m, the Nyquist sample interval $\sim$ 30s. The factor $2Bt \sim 3\ 10^4$ for a 1 kHz spectral channel width. For the ATA, a low resolution array with a modest number of antennas, it makes sense to build a correlator and sample the correlation function averaged at the Nyquist rate.

For the SKA, with ~ 4000 12m-diameter antennas and higher sample rates needed for ~ 1000 km baselines, the decision is not so clear. The Nyquist sample interval is $\sim$ 10s for an 8 km baseline of 12m antennas, and $\sim$ 0.1s for an 800 km baseline for a single pointing center. On an 800 km baseline with a 1 kHz channel width, the output bandwidth from the correlator is ~ 100 times higher than the input bandwidth and it may be better to integrate the data in the image plane.

The required output sample rate from a correlator is increased by a number of factors including large field imaging and interference rejection, etc. A large number of antennas favors antenna based computations.

5. Correlation Function Imaging

The standard way of forming radio astronomy images is from the Fourier transform of the sampled correlation function between pairs of antennas (e.g., see TMS and references therein).

For a small field of view (FoV) a 2D FFT can be used. This works well at high frequencies, with a small FoV, compact sources, and a low density of background sources. Aperture synthesis using Fourier transform of the sampled correlation function works best for imaging compact source distributions at centimeter wavelengths, as typified by VLA measurements of strong extragalactic radio sources.

5.1. large field imaging problems

A small FoV is not so good for sources which are large compared with primary beam when mosaicing techniques must be used. There are substantial problems with large scale structure since spatial frequencies less than the antenna diameter cannot be sampled without antenna shadowing, and many astronomical sources are larger than λ/D_{ant}. Image fidelity is severely limited by the accuracy of single dish measurements needed to complement the interferometer measurements.

At decimeter wavelengths there are problems with non coplanar array geometry and non isoplanicity of the atmospheric coherence. When the density of sources is high, the whole FoV must be imaged in order to remove the sidelobes of all the sources within primary beam pattern of the antenna, and a 2D FFT cannot be used with the large FoV.

There are several approaches to these problems. The number of pixels in the z direction $\sim \lambda D_{max}/D_{ant}^2$. A 3D FFT is inefficient since most of the pixels are not on the celestial sphere and are empty. A common approach pioneered and developed by Cornwell and Perley (1992) is to mosaic the FoV with tangent plane images which are small enough that the 2D FFT approximation is good. The number of tangent plane images is also $\sim \lambda D_{max}/D_{ant}^2$. Each subfield still contains sidelobes from emission outside the subfield, and this emission must be subtracted from the uv data. The algorithm proceeds iteratively, with the mosaic of tangent plane images being re-computed when the residual noise in the deconvolved subfields is comparable to the sidelobe level from emission from sources outside each subfield.

Perley and Clark (2003) have developed scaling relations for the image processing requirements for large field imaging. For imaging the full FoV, they show that the processing cost scales as $\sim \lambda^{1.7} N_{sta} D_{max}^3/D_{ant}^6$, where N_{sta} is the number of antennas per station. If the antennas are clustered in stations the effective uv coverage is reduced, the synthesised beam has higher sidelobe levels, and the tangent plane images must be re-computed more often. For an array with a constant total collecting area, the cost scales as $\sim \lambda^{1.7} N_{sta} (N\ D_{max})^3$.

This could be construed as an argument in favor of a smaller number of larger antennas (such as, for example, the EVLA design). Such an array is well suited for mapping sources with high brightness compact emission which can be imaged at high resolution with a sparse uv coverage, but is less well suited for imaging extended emission where the image fidelity is severely limited by inadequate uv sampling and missing large scale structure which must be imaged using single dish observations.

As noted by Perley and Clark, these scaling relations only apply to full field of view imaging at the highest resolution. In many cases, the SKA will not image the full field of view at the highest resolution, and these scaling relations will not apply. For example, we consider two alternative observing modes.

(i) The SKA configuration provides high brightness sensitivity at lower resolution from the central core of collecting area. Conversely, at the highest resolution, only brighter, more compact regions can be imaged with the sparse uv coverage. For the proposed SKA configuration, the brightness sensitivity scales as $\sim D_{ant}^2/D_{max}$ allowing for the density of sampled uv points. So for imaging at constant brightness sensitivity, the data processing costs (using tangent plane faceting) scale approximately as N_{sta}.

(ii) Imaging fields of view larger than λ/D_{ant}, must use mosaics of smaller images. Mosaicing requires time shared observations (or array receivers). A mosaic process may also be required to merge subimages for different isoplanatic patches. The data processing for mosaicing scales roughly as the number of pointings, $N_p \sim D_{ant}^2$.

Since single antennas provide cleaner primary beam patterns than station beams from phased arrays, one could argue for a hybrid array with a large number of smaller antennas in a central cluster for imaging extended low brightness emission at lower resolution, and a smaller number of larger antennas at longer baselines to image compact higher brightness regions at higher resolution.

6. Direct Imaging

Direct imaging, or beam formation, solves a number of problems for large field images. The idea is to integrate the data in the image plane, and not as correlation functions, as outlined in ATA memos 1, 12 and 30.

In many cases it will not be necessary to image the whole FoV. We will have a-priori models of the sky brightness, and can identify regions which have bright emission. We need only image regions which are of interest or contain strong sources whose sidelobes corrupt the regions of interest.

For each direction, add up voltage V_k in phase from each antenna, k, and square the get the power

$I(s) = [\Sigma V_k(r) \exp(2\pi i/\lambda \ r.s)]^2$,

where s is the sky position vector and r is the antenna position vector.

For full polarization imaging, we must integrate all 4 cross products of the 2 polarizations from each antenna.

The summation, Σ, is the sampled 3D Fourier transform of the signals from the antenna array. For compact irregular arrays, it is possible to interpolate to a regular grid and use FFT techniques

(e.g. see Alan Rogers - ATA memo 12), but for imaging large fields with an unfilled array which does not adequately sample the wavefront, we can not interpolate the sampled data.

In any case we can write

$\exp(2\pi i/\lambda\ r.s) = \exp(2\pi i/\lambda\ r.s_o) \times \exp(2\pi i/\lambda\ r.s_i)$,

where s_o is the phase center and s_i is an array of sky positions in each region we wish to image.

The SKA specification is for imaging 10^4 separate regions each with 10^6 pixels (10^5 beam areas with 3 pixels/beam) within the FoV and up to 10^5 frequency channels. A possible architecture for direct imaging is a 1000x1000 matrix multiplier for each antenna which provides the phase factors for an array of positions, s_i. The phased data from each antenna is integrated into an image for each subfield and each frequency channel. In many cases it may be possible to use the same array of phase factors for multiple phase centers and frequency channels (within some reasonable bandwidth), since the image size can be scaled to the correct pixel size and frequency. In order to time multiplex the phased array multipliers for each subfield and frequency channel, the matrix multiply has to operate in ~ 0.5 ns.

So the questions: i) Is this feasible; 1 GHz DSPs are now becoming available ? ii) What are the implications if we do not record the averaged correlation function for each antenna pair ?

7. Calibration

For current aperture synthesis instruments most of the calibration of instrumental effects are made to the averaged correlation function - the *uv* data. Since most of the instrumental effects are antenna-based, we use data reduction software, such as *self calibration*, to derive antenna-based calibrations from the *uv* data and a model of the sky brightness distribution. These antenna-based calibrations are applied to the *uv* data before or during the imaging process. (So, currently we take data streams from N antennas, make $N^2/2$, cross correlations, from which we derive N antenna based calibration factors which we apply to the $N^2/2$ cross correlations before imaging!) At low frequencies there may be several isoplanatic regions within the FoV which require separate calibration. Large field images may also require substantial corrections in the sky plane.

With direct imaging we may not be able to measure individual antenna gains. Images can be Fourier transformed to the *uv* domain, but low spatial frequencies will be an average of many individual baselines. It is possible to use subsets of antennas to make direct images which contain non-redundant short *uv* spacings, but these images will have poorer fidelity and it is not clear how best to use these data to correct the antenna gains.

A correlator does provide a good way to derive antenna-based calibration factors, but it may not be necessary to cross correlate all antennas in order to calibrate the SKA. The measured correlation is $V^{'} = G(time, frequency, polarization...) \times V$, where V is the true visibility and G

is the calibration factor. If the correlator closure errors are small, the calibration factor can be expressed as antenna based calibrations, $G = g_i \times g_j$ for each antenna pair (i, j). Usually, the frequency (passband), and polarization calibrations are slow functions of time, and we can separate the calibrations into antenna dependent gains $g(time)$, passband and polarization calibrations. System components can be measured in the laboratory, but the final calibrations of the whole system are made using astronomical sources.

The high data rate from the SKA, and the need to apply gain (amplitude and phase), bandpass, polarization leakage, and Stokes conversions in order to average the data over antennas (for beam formers and station beams), time, frequency channels, or polarizations, strongly suggests that the calibrations should be applied as the data are acquired in close to real time.

Even if the antennas are clustered into stations, it is more efficient to send back the signals from each antenna, since the antennas in each station may be phased in up to 10^4 different directions. Sending the signals from each antenna is much more versatile than sending phased signals from each station. The data processing can be centrally located and computing resources allocated for calibration and imaging as needed. Calibration factors which are a function of antenna, frequency, polarization and direction in the sky, can be derived from cross correlation data, and used to correct the signals from the antennas in close to real time, before they are averaged to form beams and images.

Calibration of the SKA might proceed in a hierarchical way. Calibration of the antennas within each station cluster could be derived from a cross correlation of those antennas, and calibration between clusters (subsets of antennas in the central core of the SKA), can be derived from cross correlation of phased clusters. There is, of course, some loss in SNR in deriving the calibration if all antennas are not cross correlated, but in practice this may not be an issue since long and short baselines are responding to different sized source structures in the sky model used for calibration. For example, if compact sources in the FoV are used for calibration, we may wish to spatially filter out large scale structure and only use long baselines for deriving the calibration.

7.1. Self Calibration of Antenna Gains

The SKA will often use self calibration on the target source to determine the antenna gains, $g(time)$, from the measured correlations and a model of the source brightness distribution. In many cases continuum sources in the same field of view as the target source can be used to calibrate the antenna gains. However, the SKA is quite different from the VLA where strong compact continuum sources are often used for self calibration. For spectral line observations the signal may be distributed in complex structures over a range of frequency, e.g. Galactic HI observations. Multiple frequency channels can simultaneously be used to determine the antenna gains using a 3D image model.

Although self calibration works well for simple source models, it fails if the source model is too complex and there are too many free parameters to determine both the sky brightness distribution and the antenna gains in an atmospheric coherence time. It may be necessary to integrate the data for some time in order to acquire an adequate a-priori model of the sky brightness distribution which can be used to calibrate the antenna gains. In this case the image can be improved if the data are saved (either as correlation functions, or direct images with non-redundant baselines) so that gain corrections can be applied. The success of self calibration will depend on the individual situation, source structure, and scientific goals.

7.2. Confusion

RFI and sidelobes from strong sources also cause problems for calibration. A correlator can also be used for characterizing and removing interference. For confusing sources in the primary beam sidelobes, the signal will be time variable due to pointing, atmospheric, and at low frequency different isoplanatic patches. In this case, the confusing source could be treated as a time variable interference source, and removed using RFI mitigation techniques. E.g. some antennas could be pointed at the confusing source and the correlations between these antennas and the array pointed to the regions of interest used to remove the interference signal. For correlations between phased station beams and individual antennas, sources are illuminated by the product of the voltage patterns. Confusing sources in the sidelobes of station beams may be particularly troublesome.

8. Conclusion

1) A correlator provides a familiar and versatile mechanism for integrating the data in the *uv* domain. We have the advantage of over 40 years of development for calibration and imaging using *uv* data.

2) Large correlators are needed to correlate 1000's of antennas, and the high data rates are significantly increased by forming correlation functions.

3) The size of the correlator can be reduced by cross correlation of phased subsets of antennas. Whilst this is a possible solution for calibration and for imaging compact source structures, the narrow beamwidth and high sidelobe levels of phased station beams makes this an unattractive solution for complex structures.

4) Since the antennas may be phased in up to 10^4 different directions, sending the signals from each antenna to centrally located computing resources is more efficient and versatile than sending phased signals from groups of antennas.

5) Direct imaging, or beam formation, is appropriate for small fields, and, solves a number of problems for large field images. The development of $\sim$ 1 GHz data processing arrays may make direct image formation feasible.

6) A hybrid solution using direct imaging and correlators allocated to sub-arrays for calibration and imaging as needed, provides a flexible development path for the SKA.

7) Extrapolation of current correlators and imaging techniques suggests that the data processing requirements for an array with $\sim$ 4000 antennas can be satisfied, but experience with existing correlators shows that the development time is very long and may lag significantly behind Moore's law.

acknowledgements I would like to thank Douglas Bock and Larry D'Addario for their interest and comments on the manuscript.

9. References

Cornwell, T.J. & Perley, R.A., 1992, "Radio-Interferometric Imaging of Very Large Fields", A&A 261, 353

Jones, D.L., 2003, "SKA Science Requirements", SKA memo 45, version 6, 1 December 2003 (D.L.Jones, 16 December 2003, ISSC11-9.4)

Perley, Rick, & Clark, Barry, 2003, "Scaling Relations for Interferometric Post-Processing", EVLA memo 63

Rogers, A. E. E., 2000, "Array Processing Estimates for the ATA", Allen Telescope Array (ATA) Memo 12. (http://intranet.seti.org/docs/ata/Memo/index.html)

Ryle, M & Neville, A.C.,1962, MNRAS, 125, 39

Thompson, A. R., Moran, J. M. & Swenson, G. W., 2001, "Interferometry and synthesis in radio astronomy", 2nd ed. New York : Wiley, 2001

Wright, M.C.H., 1998, "Astronomical Imaging with the One Hectare Telescope", Allen Telescope Array (ATA) Memo 1 (http://intranet.seti.org/docs/ata/Memo/index.html)

Wright, M.C.H., 2001, "Astronomical Imaging with the ATA - III", Allen Telescope Array (ATA) Memo 30. (http://intranet.seti.org/docs/ata/Memo/index.html)

Wright, M.C.H., 2002, "A model for the SKA", SKA memo 16

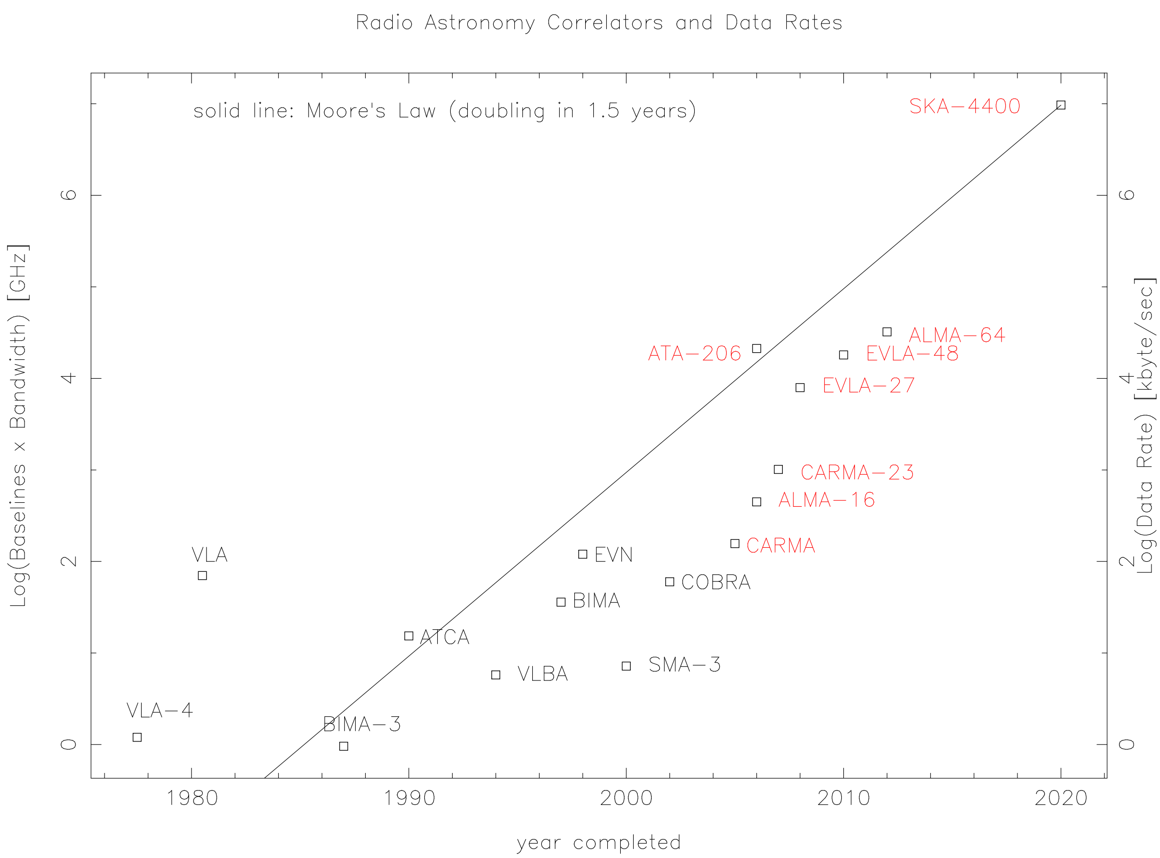

Fig. 1.— Bandwidth of Radio Astronomy correlators. Future correlators are shown in red. The solid line shows Moore's Law extrapolated from an SKA correlator for 4400 antennas with 1 GHz bandwidth. The right axis gives the data rate assuming 2000 spectral channels per GHz of bandwidth, 4 bytes per channel, a 25% overhead, and a 10s sample interval.

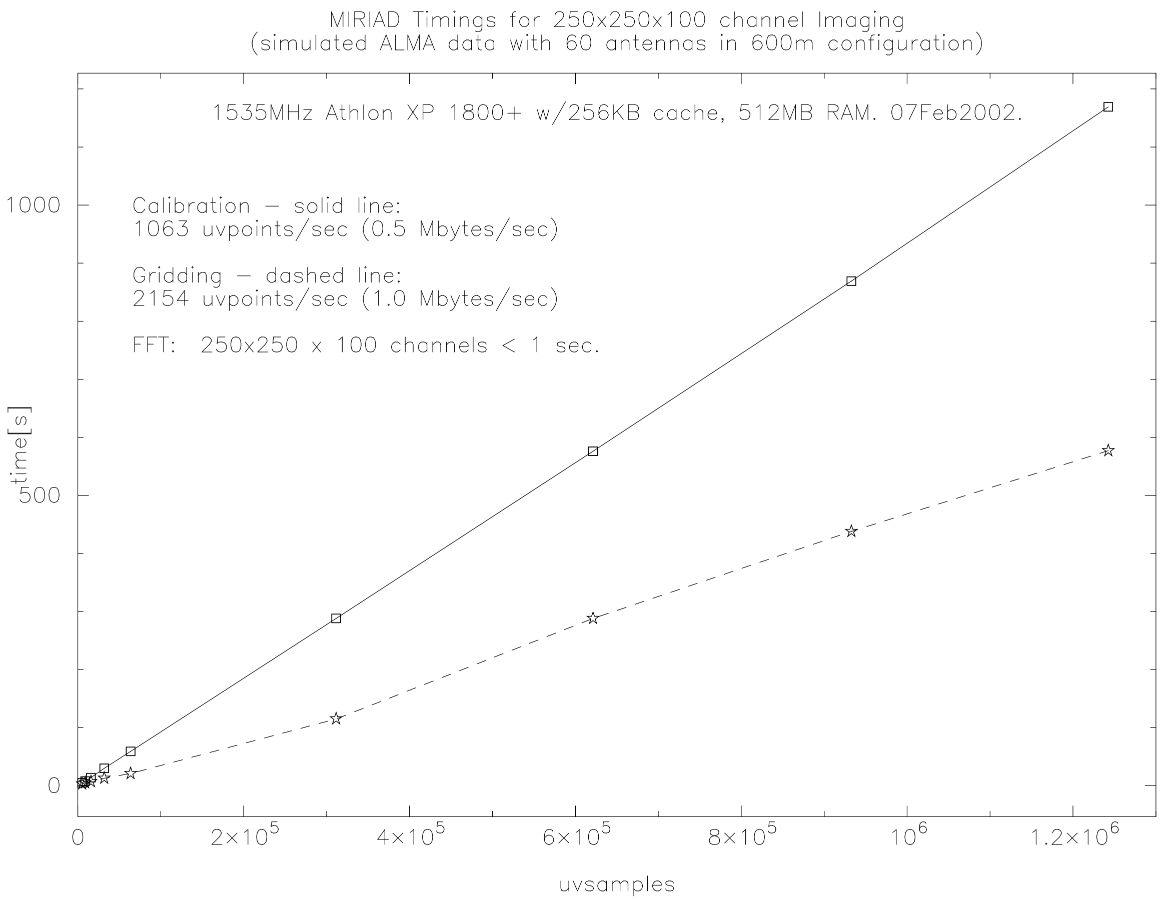

Fig. 2.— Multichannel Imaging Simulation. The plotted points and lines show the timings for Calibration and Imaging (see text)

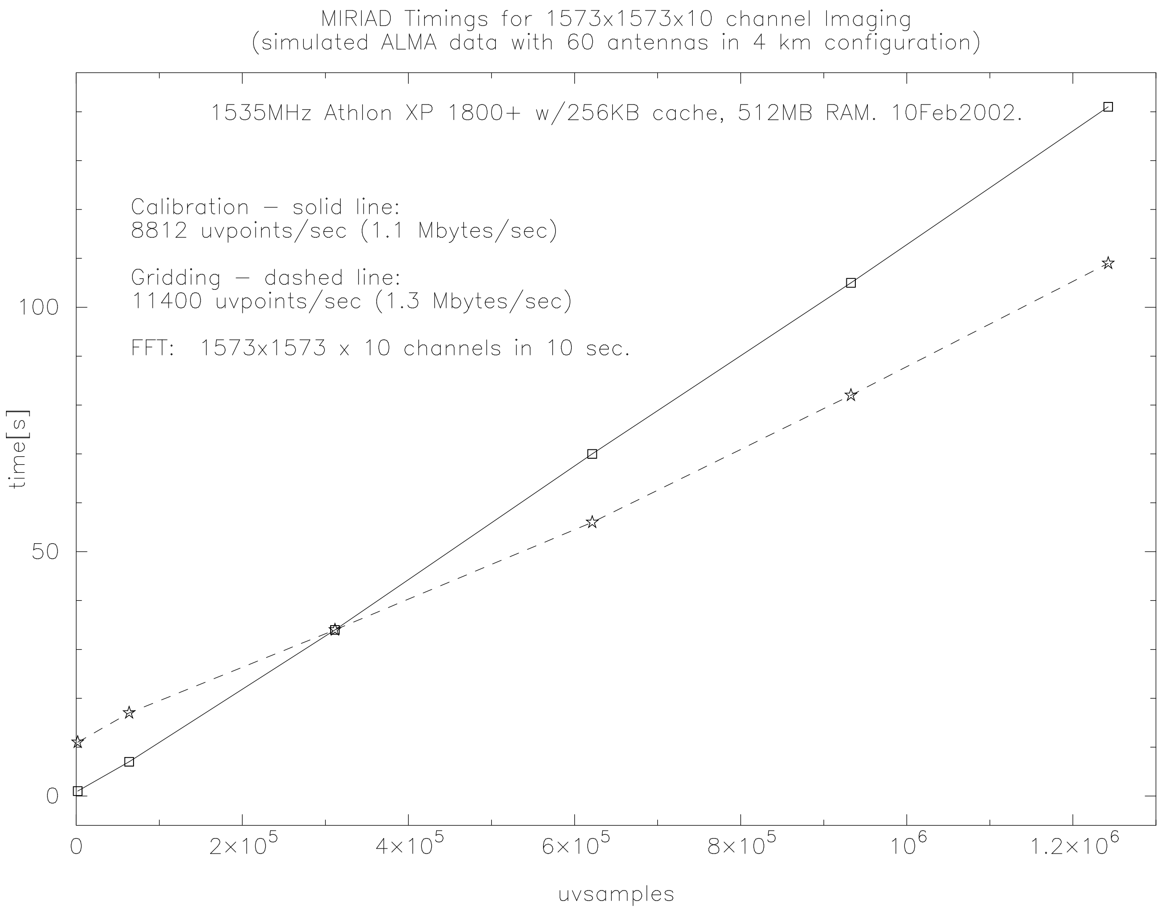

Fig. 3.— Multichannel Imaging Simulation. The plotted points and lines show the timings for Calibration and Imaging (see text)

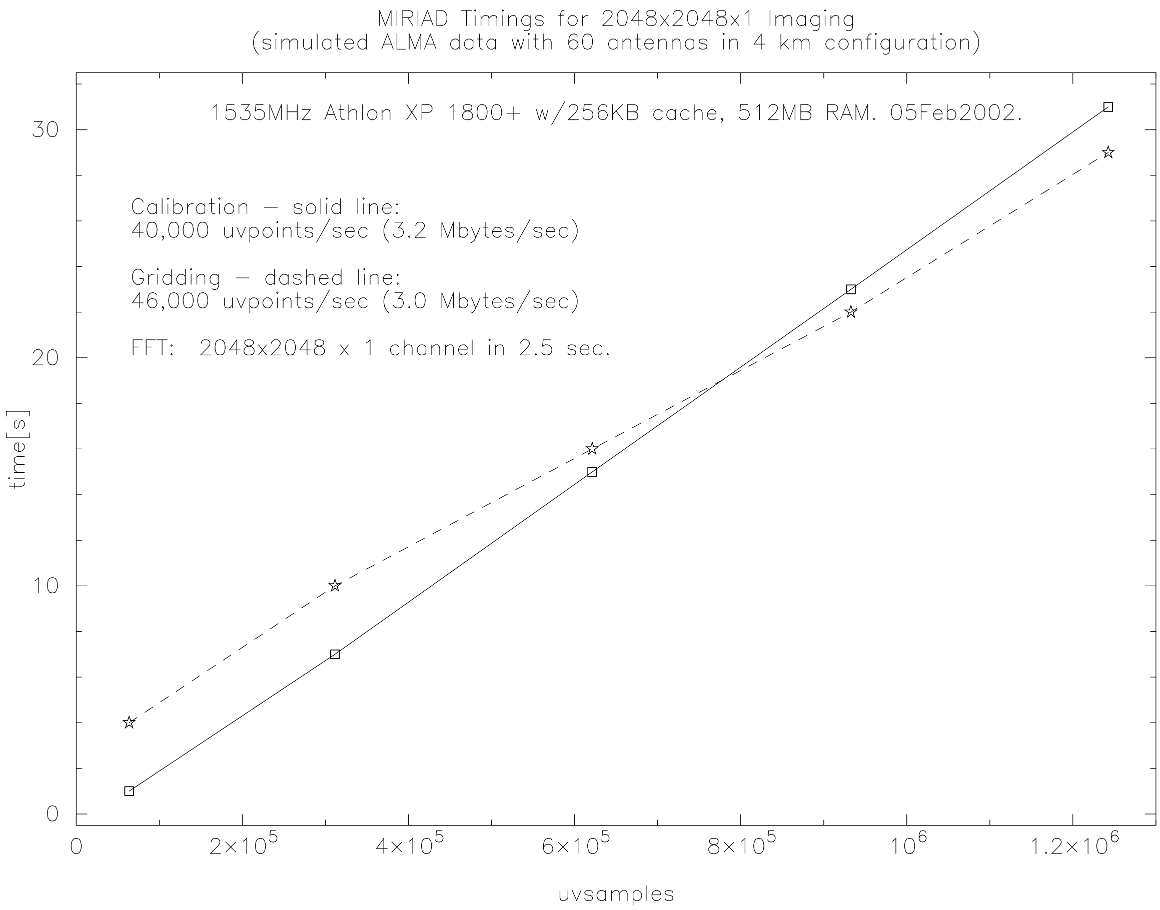

Fig. 4.— Multichannel Imaging Simulation. The plotted points and lines show the timings for Calibration and Imaging (see text)