An FX Software Correlator at the RPA

G. R. Harp and R. F. Ackermann
Nov. 8, 2001

Abstract

A prototype FX correlator is assembled at the Rapid Prototyping Array (RPA) and used to characterize the RPA's imaging capabilities. Movies are generated with the antennas held in fixed position while bright radio sources sweep across the field of view. The synthetic beam sidelobe pattern plays a dominant role in the appearance of the images. The efficiency of the correlator is discussed and the data processing limits of this architecture are elucidated.

Hardware / Choice of FX Approach

The Rapid Prototyping Array (RPA) is a small radio observatory consisting of seven off the shelf 3.6 m satellite dishes. These antennas have a prime focus design and are equipped with two orthogonal linearly polarized receivers sensitive in the 1400-1630 MHz range. Electronics perform a two-stage down conversion to baseband resulting in a net 10 MHz nominal bandwidth centered anywhere in the RF band. The baseband signal (two polarizations) for each antenna is then digitized at 30 MS/s by an ADC in a Sun Ultra-60 computer – with one computer per antenna. These computers are interconnected on a 100 Mb/s local area network comprising several other computers. The seven ADC computers are labeled RPA1 – 7 (numbers associated with their respective antennas), and specifically can communicate over the network with another Ultra-60 labeled RPA0.

Given the existing hardware arrangement described above, we considered how we could most easily build a spectral line correlator, and decided on an FX approach. Since the data stream from each antenna passes through a dedicated computer, it makes sense to have that computer perform the maximal amount of processing before shipping the data out for correlation. Hence FX processing is advantageous over XF or direct imaging.

Design of the Correlator

Data Decimation

It was immediately recognized that the maximal processing speed of the computers and network fall far below the acquisition rate of the ADC's. The total acquisition rate is 30 MS/s × 8 bits/Sample × 2 polarizations × 7 antennas = 3360 Mb/s. By comparison, one expects of order 50 Mb/s net transfer rate from our network.

Likewise, benchmarks showed that the Ultra-60's (300 MHz) are incapable of performing significant processing on the 60 MS/s data stream from the ADC (only 5 instructions per sample!). For FX processing, it is the Fourier transform that dominates the processing time. Benchmarks with the FFTW library of floating point routines showed that the Ultra-

60's could process ~ 200 kS/s/channel (using 512 sample FFT's), where we use "channel" as shorthand for the data stream from a single antenna and polarization.

Both of these constraints drive the design toward one where only a small subset of the data are processed. Data reduction was accomplished as follows (see Fig. 1): for each channel, N consecutive samples (typically N = 512) are captured from the 30 MS/s data stream and then stored. Approximately 1000 × N samples are then skipped, at which time another N samples are captured (typical capture rate = 40 kS/s/channel).

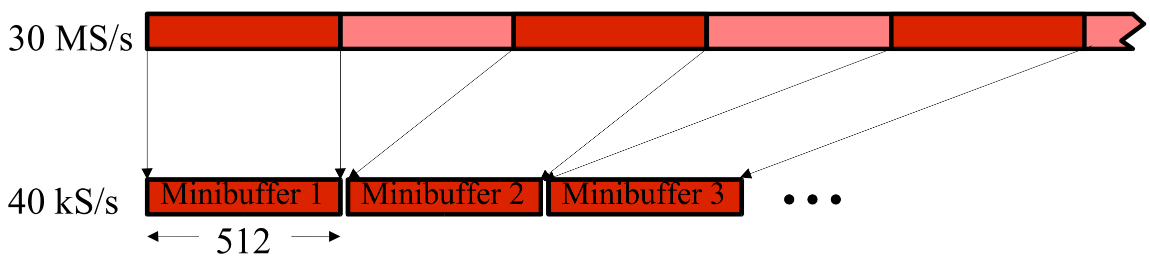

Fig. 1: Data decimation approach.

In a separate computing process (or thread) each 512 sample "minibuffer" is Fourier transformed to complex amplitude over the frequency range 0-15 MHz at baseband. The analog bandpass filter is such that the lowest and highest 2.5 MHz are corrupted by aliasing. There is no point transmitting this data over the network, so we chose (arbitrarily) to ship only half the Fourier transformed data, the range 3.75-11.25 MHz, for subsequent correlation (typical transfer rate = 20 kS/s/channel).

Problem with Trigger

The RPA is equipped with a trigger device that can be used to synchronize the start of data acquisition on all seven antennas. Initially, we had planned to use this mechanism in the correlator by triggering, acquiring a small amount of data, and then resetting the trigger. This process would be repeated frequently. However, Geoff Bower discovered some problems with the trigger, and ultimately it was shown that shortly after trigger, the ADC cards would frequently drop anywhere from 0-7 samples. This means that timing / phase calibrations must be repeated *every* time the trigger was used.

More detailed examination of the ADC performance showed that after the first μs, all the ADC's performed well and never dropped any samples (except RPA5 which drops about 1 per hour). So the data acquisition software was written to use the trigger only once at the beginning of an observation to get the data streams into rough alignment. After that, 30 MS/s data acquisition runs continuously as a background process on each of RPA1-7 for the duration of the experiment.

Data Servers, Data Queue, and Correlator

The multi-threaded process that acquires, decimates, and FFT's the data is called the Data Server (Fig. 2). This program reduces the 30 MS/s × 8-bit data stream from a single channel to 20 kS/s × 32-bit floating point. This data is time-tagged and then shipped across the network in blocks comprising 32 mini-buffers or 16384 samples.

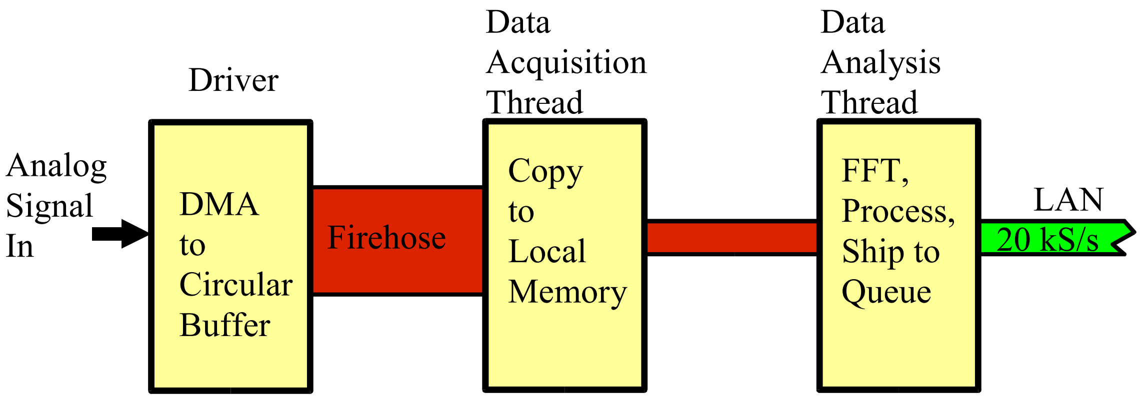

Fig. 2: Block Diagram of the Data Server running on RPA1-7.

The data are then captured by the Data Queue program running on RPA0 (Fig. 3). This program assembles the data from all antennas into a data structure. When this structure is complete, the data is shipped to various Data Subscribers, including the Correlator and Data Monitor programs.

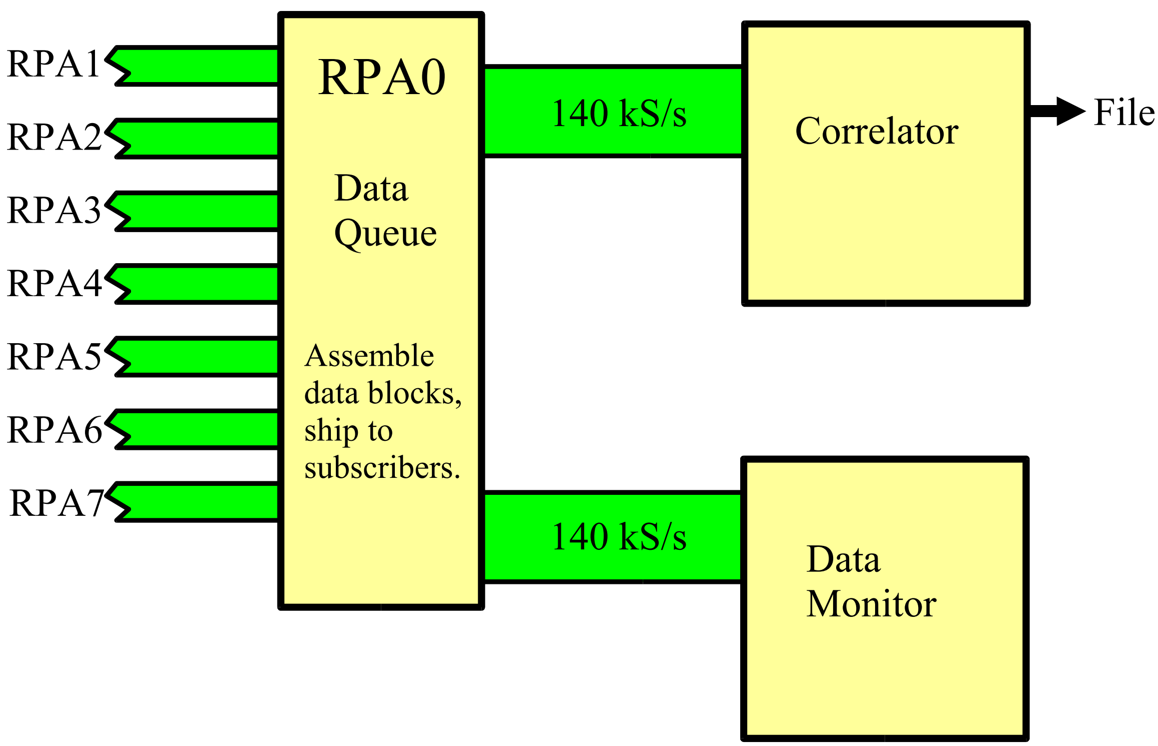

Fig. 3: Block diagram including the Data Queue, Correlator, and Data Monitor.

The Data Monitor provides real-time diagnostics. The Correlator takes the array of data blocks from the Queue and performs cross correlations and autocorrelations for all polarizations and antennas, for a total of 112 correlations times the number of frequency bins.

The Data Server was written in C++ while all other programs were coded in Java.

Experiment

The experiments described here represent the simplest imaging experiments conceivable, requiring the least amount of pre-correlation calibration. Because the antenna drive motors generate RF noise, it is necessary to blank the data processing while the antennas are moving. Therefore one simplification was to perform measurements with the antennas held in fixed position, while the object of interest sweeps by. Another simplification is to ignore the frequency-dependence and simply sum the correlated power over all frequency bins, giving the white light image. Finally, we chose image fields containing only a single symmetrical source. In future work, we intend to go beyond all these simplifications.

As demonstrations, we present data from a GPS satellite, the sun, and Cassiopeia A.

Calibrations / Data Analysis

Gain

The data were calibrated post-correlation as follows. First, the overall gain of each data channel was normalized by first calculating the integral, g_i, of the power in that stream over the duration of the measurement:

$$g_i = \int_{meas.} \mathcal{V}_{ii}(t) - D_i \quad dt\,,$$

where $\mathcal{V}_{ii}(t)$ is the measured autocorrelation and D_i is the average "dark power" of that data stream (measured power when observing dark sky). Here and in the discussion that follows, a single subscript indicates the channel (i.e. both the antenna and polarization). All of the autocorrelations, $\mathcal{V}_{ii}(t)$, and cross-correlations, $\mathcal{V}_{ij}(t)$, are normalized to the g_i as follows:

$$\mathcal{V}'_{ij}(t) \equiv \frac{\mathcal{V}_{ij}(t)}{\sqrt{g_i g_j}}\,.$$

Phase

Phase calibration is performed by using knowledge of the source under investigation. Since we are observing a solitary symmetrical source, and we know that when this source passes through image center (time t_o) the phase at every antenna must be the same. We impose this condition upon the data by expressing each correlation as

$$\mathcal{V}'_{ij}(t) \equiv \left|\mathcal{V}'_{ij}(t)\right| e^{i\phi_{ij}(t)}\,,$$

and then adjusting the phase of every $\mathcal{V}'_{ij}$ so that at time t_o the phase of that correlation is zero:

$$\mathcal{V}''_{ij} \equiv \left|\mathcal{V}'_{ij}\right| e^{i\left(\phi_{ij}(t)-\phi_{ij}(t_0)\right)}.$$

A little algebra shows that this is equivalent to all the antennas having equal phase at time t_o.

GPS Satellite

As a first example, we examine the case of a GPS satellite sweeping across the field of view. The time of maximal intensity approximately coincided with transit for the satellite at azimuth = 100.3° and elevation 70.1°. The IF electronics were tuned to 1575 MHz coinciding with the GPS band center. The correlator was set to integrate the signal for about 1 s of real time before writing data to disk.

The measured autocorrelations are displayed in Fig. 4. The FWHM of the primary antenna beam is about 4°.[1] Because of the gain calibration mentioned in the previous paragraph, all plots have about the same peak intensity. However, the time of peak intensity varies significantly due to pointing errors. At present, we know the antenna pointing direction to no better than ±2° as measured by crude geometrical methods. Future electronic measurements could greatly improve the accuracy of our pointing model.

[1] See ATA memo #19 for more information regarding the primary antenna beam shape.

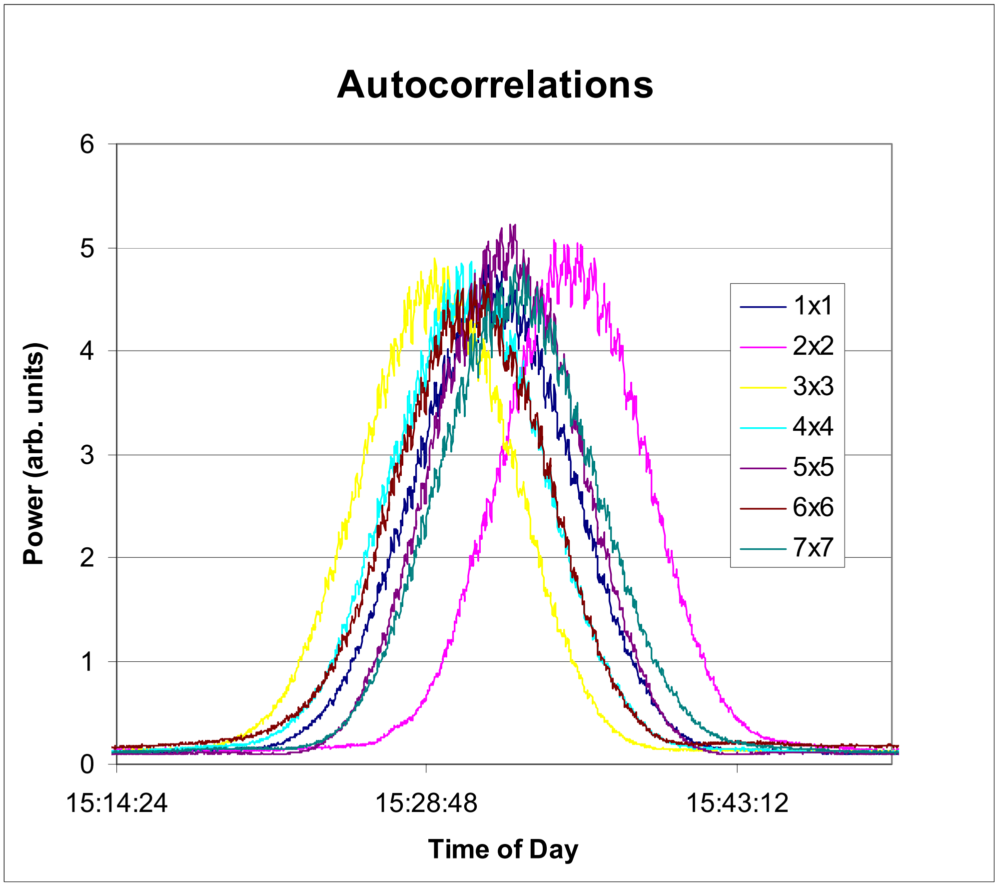

Fig. 4: The autocorrelations measured from a GPS satellite. The visible fluctuations are far above the noise level and are real fluctuations of the GPS power with time. The peak power occurs at different times for different antennas due to pointing errors.

The real part of twenty-one cross correlations from this satellite are plotted in Fig. 5. These correlations were generated by choosing a single (arbitrary) polarization on each antenna. Because GPS is strongly circularly polarized, these data represent a complete measurement of the GPS signal. At the right of the figure we indicate the numbers of the antennas used in that correlation. Clear differences in fringe frequency are observed depending on the baseline and direction of satellite motion.

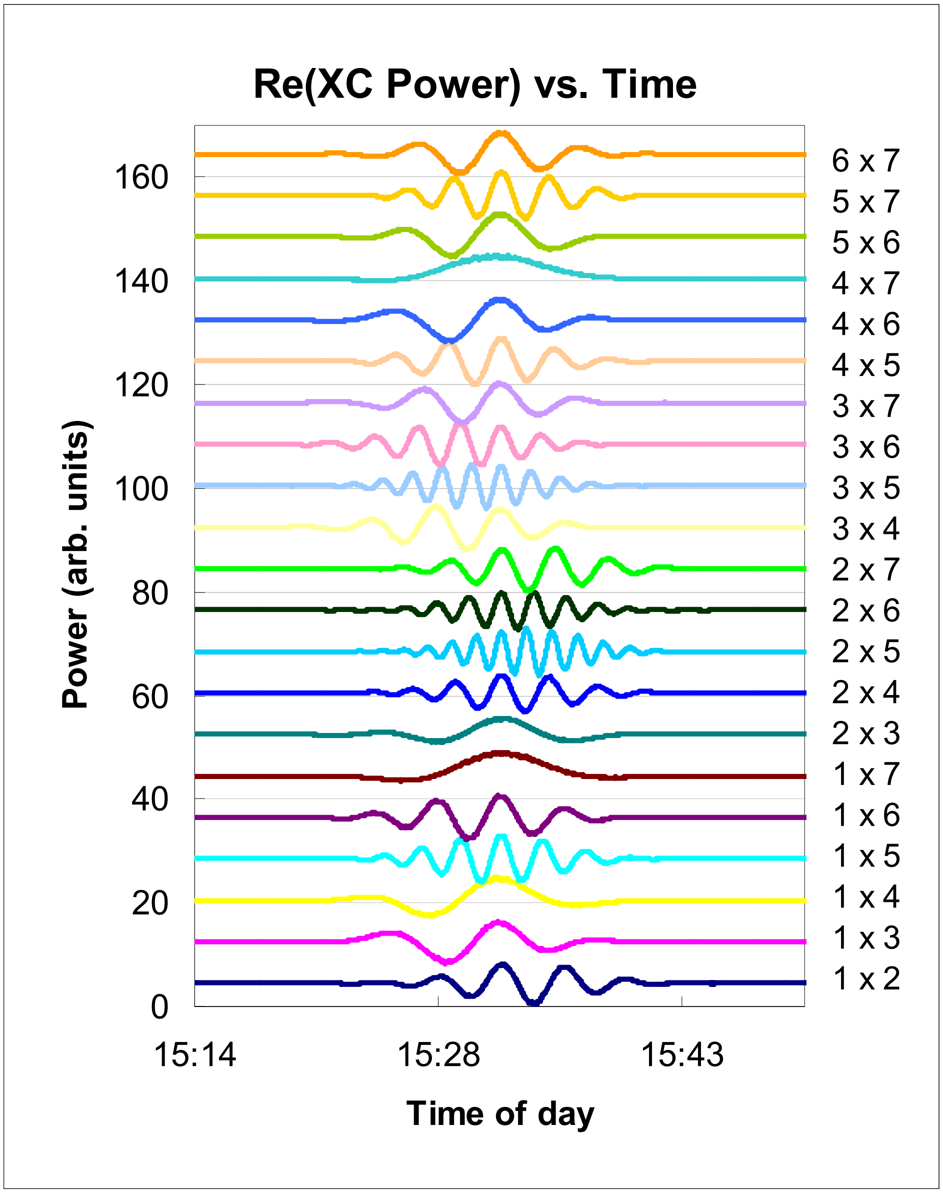

Fig. 5: Real part of the cross-correlated power of a GPS satellite including a single polarization from all antennas. Each plot is displaced vertically by 4 units as an aid to comparison.

Using the correlations in Fig. 5, about 500 images were generated and combined into an mpeg movie. This movie can be found on the ATA software website (http://intranet.seti.org/docs/ata/software/docs/). Fig. 6 shows four stills from this movie. Arrows indicate the instantaneous position of the satellite as it sweeps across the image. North is up and east is to the left in the images. The blue circle indicates the FWHM of the antenna primary beam.

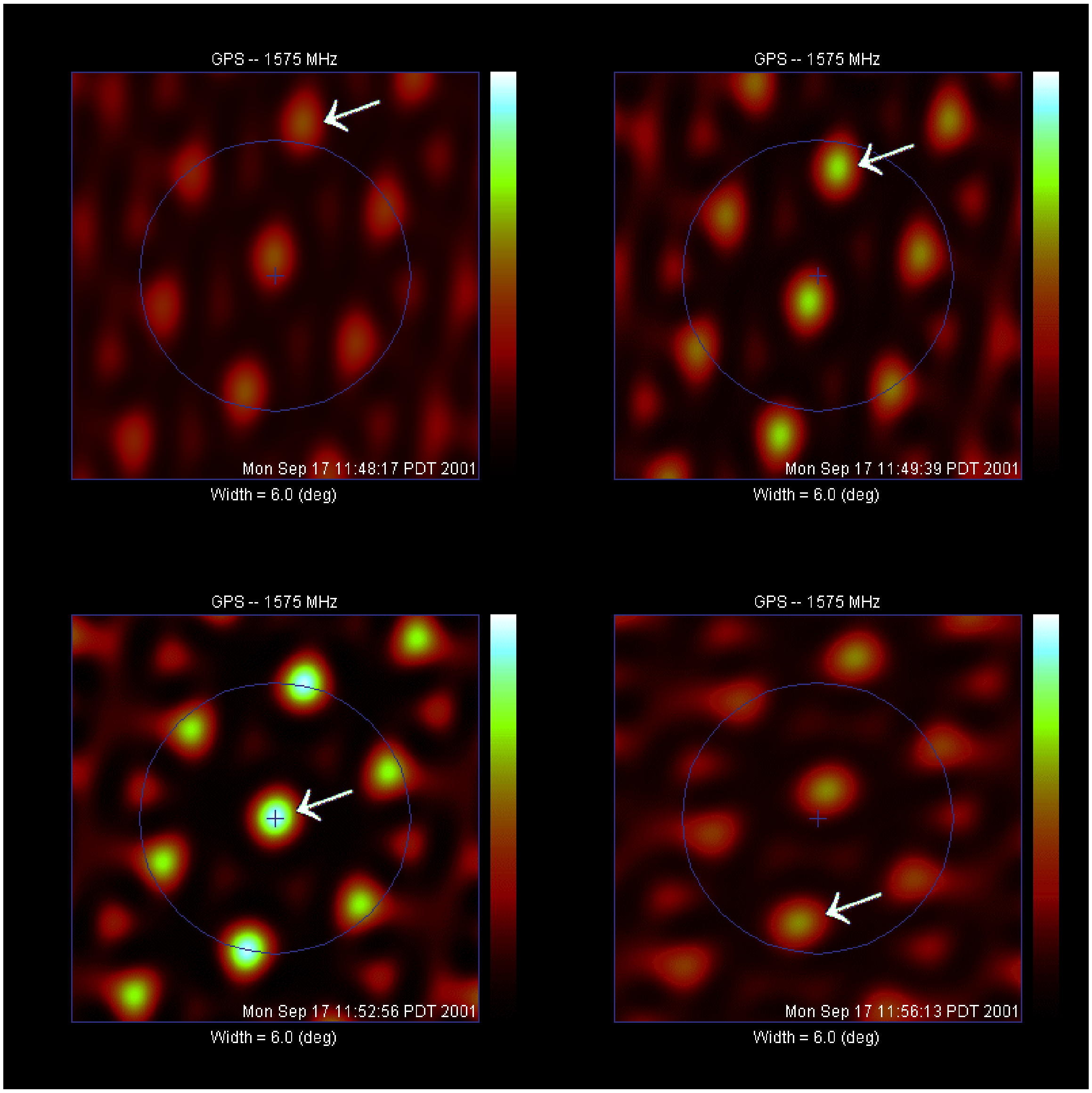

Fig. 6: Four stills of GPS satellite sweeping from north to south across image field. Satellite position is indicated by arrows. The blue circle represents the FWHM of the primary antenna beam.

Following the satellite as it sweeps through the primary beam, we observe the overall signal intensity varies strongly. When the satellite is far from the primary beam center, the intensity is low, and the intensity reaches maximum at the primary beam center. The satellite changes shape slightly depending on its position in the field of view. This is due to the pointing errors mentioned above which causes relative variations in the amplitude of different cross correlations depending on satellite position.

The Synthetic Beam

Strong sidelobes surround the satellite in all the images. The RPA dishes are arranged very nearly as a filled hexagon, giving rise to strong grating sidelobes. Fig. 7 plots the true RPA baselines (blue squares) in comparison with those from a hypothetical squashed hexagonal arrangement (purple dots) that maximizes the baseline degeneracy. The degeneracy associated with each purple dot is labeled with numbers.

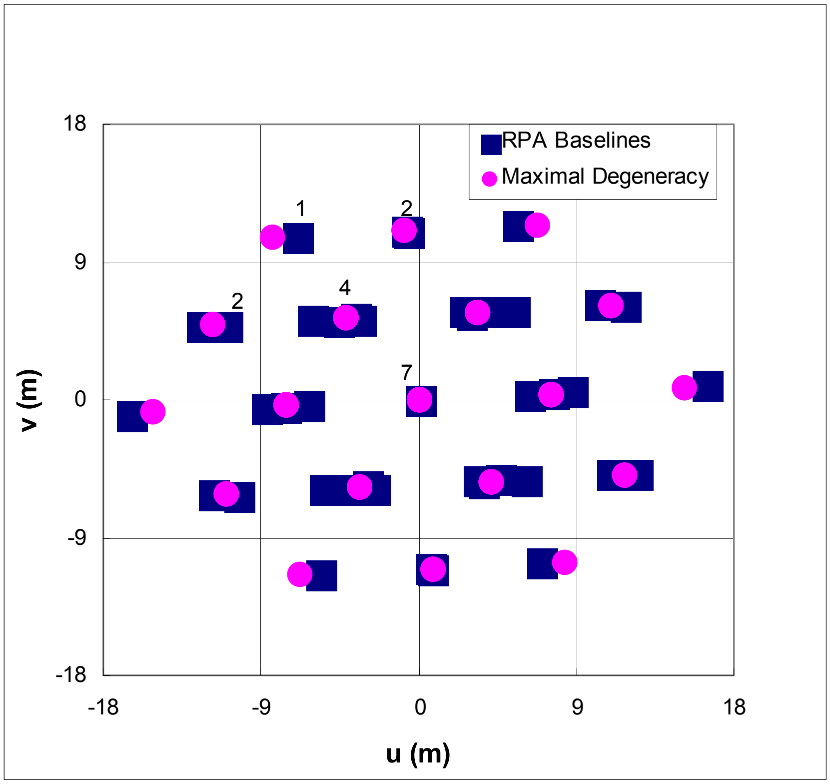

Fig. 7: Plot of the RPA baselines compared with those of a hypothetical squashed hexagonal array that maximizes baseline redundancy.

The images in Fig. 6 may be compared with Fig. 8 where the beam of an ideally pointed array is displayed. The grating sidelobes are strongly evident. If the RPA were arranged as a true grating (purple dots in Fig. 7) then the grating sidelobes would be perfect replicas of the beam center. In that case, the region bounded by the white hexagon would represent the non-aliased area of the image (according to the two-dimensional version of the sampling theorem).

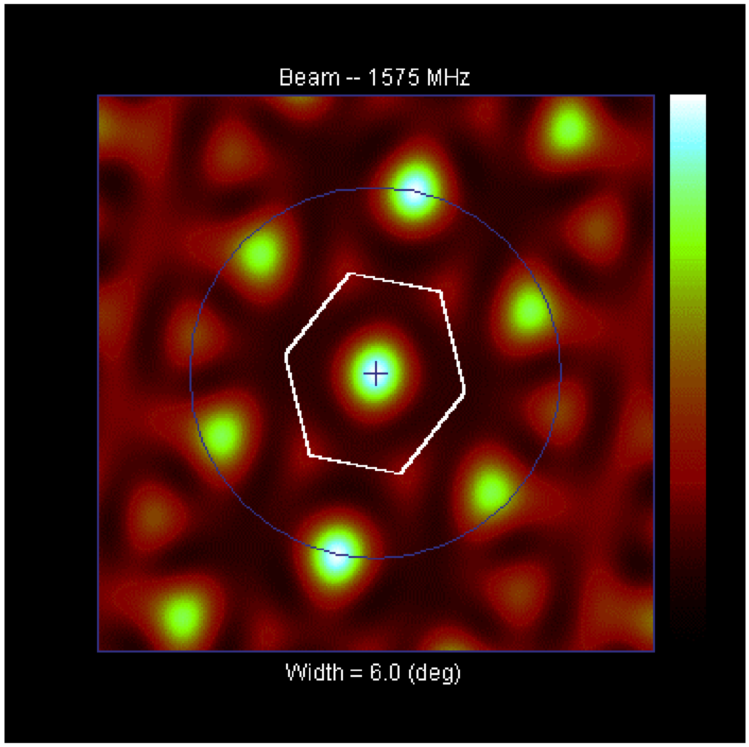

Fig. 8: Plot of the array beam under idealized conditions. The white hexagon represents the non-aliased area of the image for a hypothetical maximally degenerate array.

The grating responses extend far from the beam center as shown in Fig. 9, which is equivalent to Fig. 8 but over an expanded angular range.

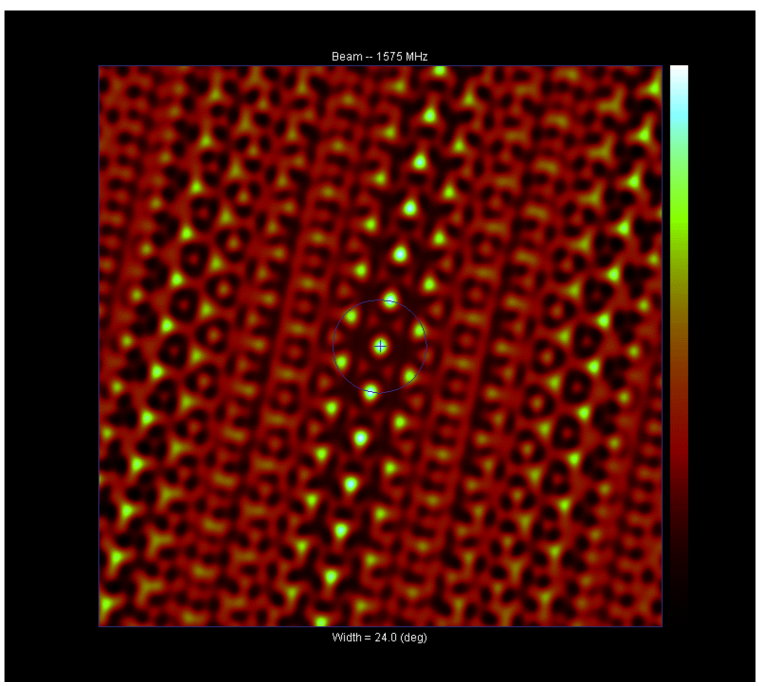

Fig. 9: Same as Fig. 8 but over an expanded angular range.

The Sun

Another set of images was generated from the sun. The parameters for data acquisition were the same as for the GPS satellite, except that the electronics were tuned to a center frequency of 1450 MHz, where little or no satellite interference has been observed at the RPA. Solar data was captured near meridian at azimuth 184.7° and elevation 50.1°. Because the sun's signal is unpolarized, and because the RPA antenna polarizations are arbitrarily rotated from antenna to antenna, it is important to measure all cross products for all polarizations. The images below were formed by summing the cross powers for all polarizations on a given baseline. For example, the cross correlation for antennas 1 and 2 was formed by the sum

$$\mathcal{V}_{12} \equiv \mathcal{V}_{1x\,2x} + \mathcal{V}_{1x\,2y} + \mathcal{V}_{1y\,2x} + \mathcal{V}_{1y\,2y},$$

where x and y refer to the two orthogonal (but otherwise arbitrary) polarizations on each antenna. (Note, sum is performed after phase calibration.)

Fig. 10 displays four stills from the sun movie. The image of the sun is noticeably broader than the GPS image, which is expected since we are barely resolving the sun. Arrows indicate the feature corresponding to the real sun image.

Notice that between the two lower images, only two seconds of time has passed. Yet the image sun has decreased significantly in amplitude and broadened, with a new intensity maximum appearing about one degree to the right of the true sun image. This sudden change is due to the "dropping" of one sample on RPA5. Effectively, 33 ns of delay were subtracted from the signal coming from antenna 5. The phase shift associated with this event is clearly visible in plots of the cross correlations for this data set, shown in Fig. 11.

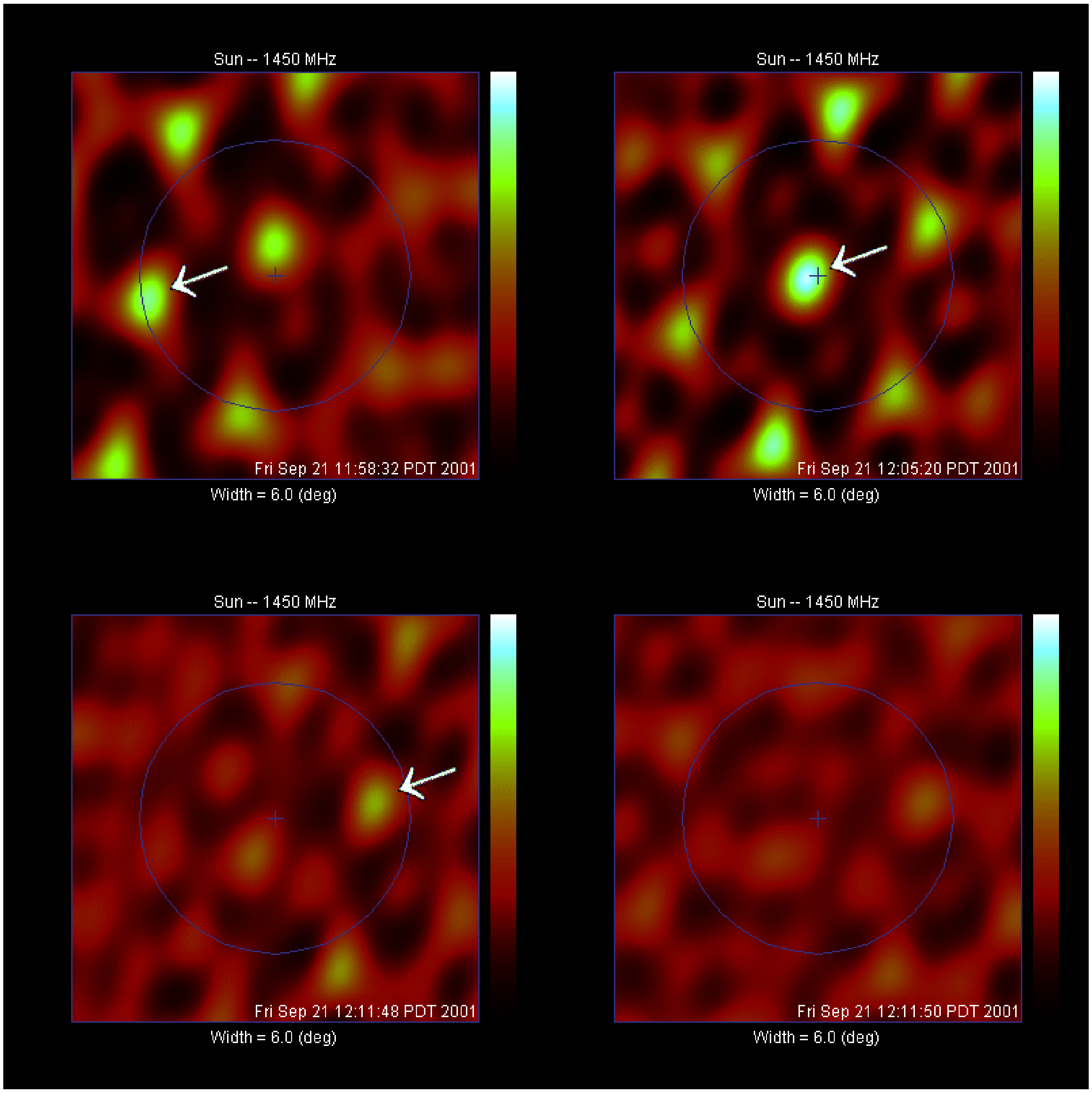

Fig. 10: Four stills from the movie of the Sun. Again, the mpeg movie can be found on the ATA software website. A dropped sample on RPA5 causes severe degradation of the image between the two lower stills, which are separated by only 2 s real time.

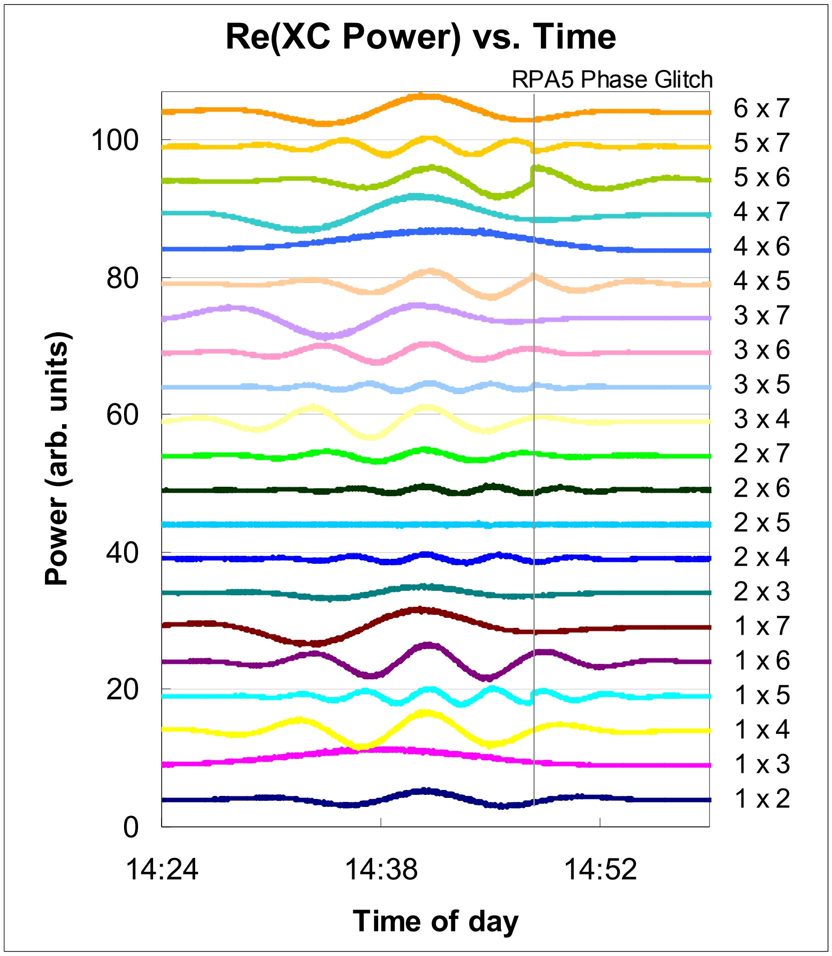

Fig. 11: Real part of the cross correlations from the sun data used in Fig. 10. At one point, RPA5 drops a sample, introducing a phase glitch in all cross correlations associated with that antenna.

Cross Polarization Leakage

By the nature of the antenna design, there is coupling between the two nominally orthogonal receivers on each antenna.[2] A measure of this leakage can be made by observing an unpolarized source (like the sun) and comparing the autocorrelation of one polarization with the cross correlation of both, as in Fig. 12.

Depending on the position of the sun in the antenna beam, there is a variable degree of cross talk between the two polarizations, with a peak value near 10%. The cross talk passes through zero close to the time when the sun has maximum amplitude. If the pointing model were improved, the zero crossing might be found to coincide with the point where the sun passes through the center of the antenna beam (as suggested by symmetry).

[2] See e.g. ATA memo #26 for another discussion of polarization cross talk at the RPA.

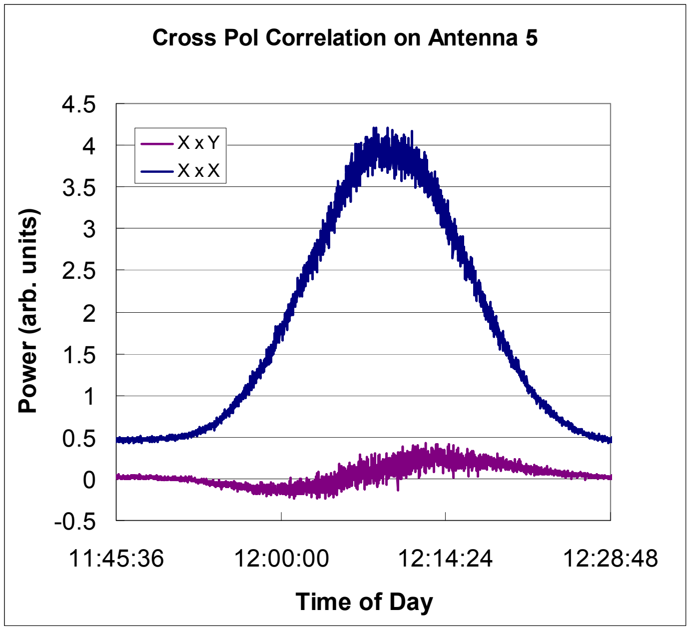

Fig. 12: Comparison of the autocorrelation of one polarization with the cross correlation of both polarizations on Antenna 5. The maximal cross polarization amplitude is about 10%.

Closure Relations

When observing point sources, the phase and amplitude closure relations are most easily interpreted, providing a quantitative measure of the stability of the array electronics. For a point source, phase closure implies that the sum of the cross correlation phases around any "closed loop" should be zero. For example, a closed loop including antennas 1, 3, and 5 is formed by

$$\phi_{13} + \phi_{35} + \phi_{51} \quad = \quad 0\,.$$

Using the GPS data of Fig. 5, we test the phase closure relation in Figs. 13 and 14. In Fig. 13 the correlation phases are plotted (colored lines, modulo π) and vary strongly with time as expected. However, their sum (black line) is very close to zero. In Fig. 14, we plot the same phase sum (on an expanded scale), but superimpose the product of the amplitudes of antennas 1, 3, and 5. The phase sum hovers near zero as long as the amplitude of the composite signal is large, as expected.

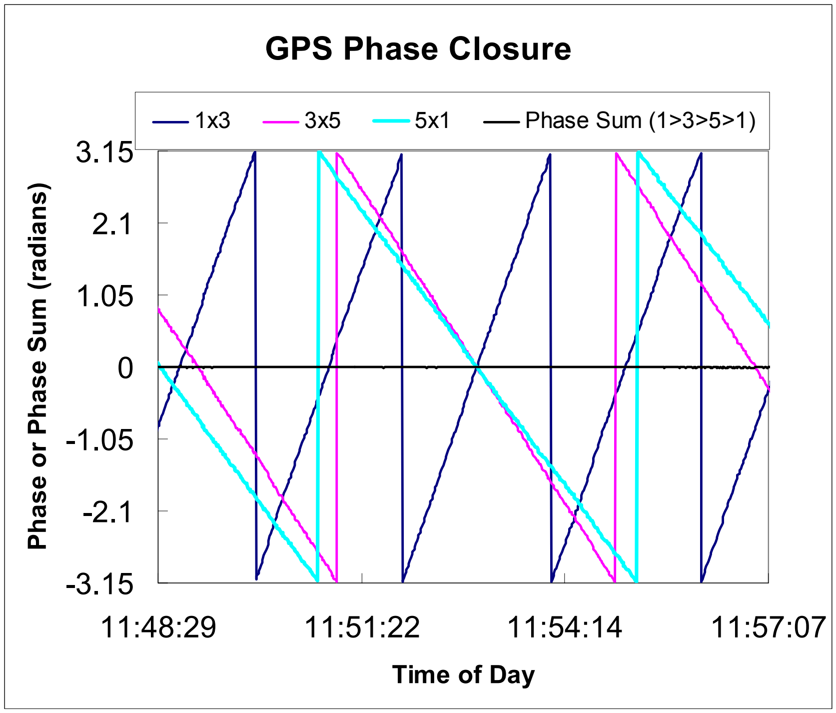

Fig. 13: The phases of three cross correlations (colored lines) are summed around the loop (black line) giving a value very close to zero independent of time.

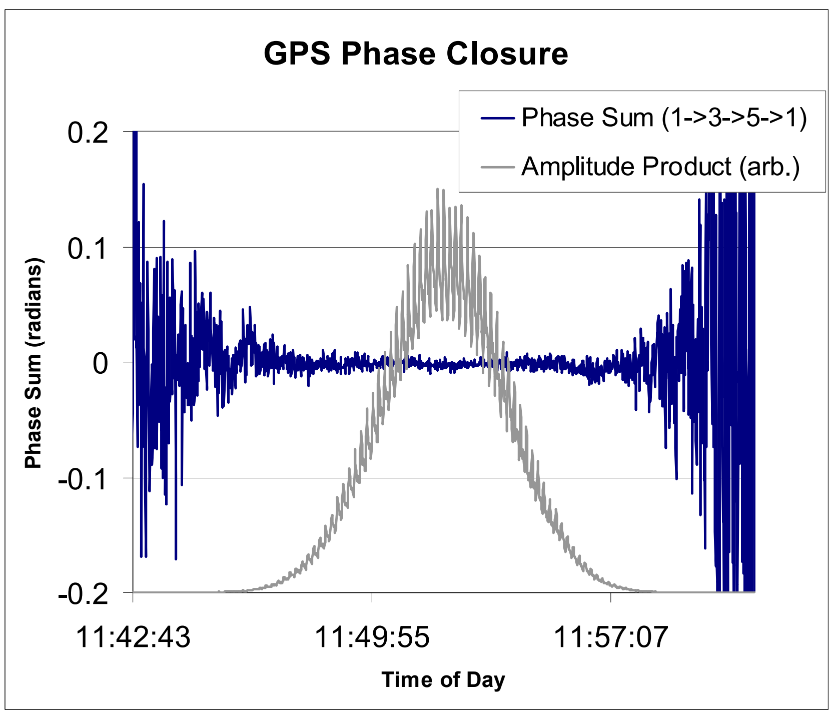

Fig. 14: The phase sum in Fig. 13 is plotted on different horizontal and vertical scales. Superimposed on this data is the product of the amplitudes of antennas 1, 3, and 5. It is seen that whenever there is sufficient signal, the phase sum remains small. The oscillations in the amplitude product are real variations in the GPS transmitted power.

Amplitude closure on a point source operates on the principle that the amplitude of the cross correlation of two antennas can be decomposed into the product of amplitudes associated with each antenna. Thus ratios such as the following must be unity:

$$\frac{|\mathcal{V}_{12}||\mathcal{V}_{34}|}{|\mathcal{V}_{13}||\mathcal{V}_{24}|} = 1.$$

One test of such a closure relation is shown in Fig. 15, where the measured ratio indeed hovers close to 1 as long as there is a strong signal from the satellite.

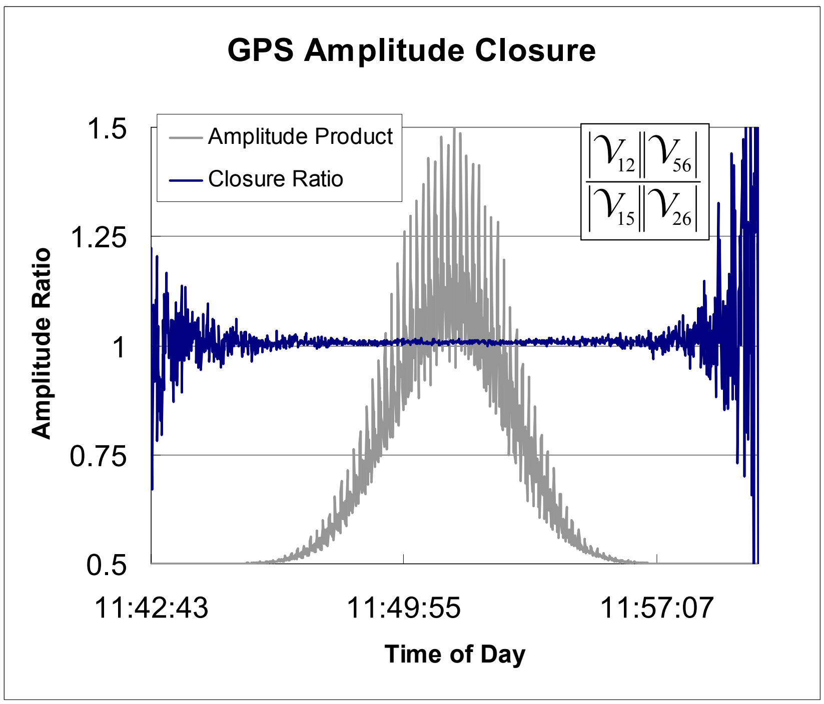

Fig. 15: Test of amplitude closure using antennas 1, 2, 5, and 6. The designated ratio remains within a few percent of unity for as long as the GPS signal remains strong.

Limitations of this Architecture

A few tests were performed to identify the bottlenecks in this correlator and estimate its ultimate speed. To begin with, using the FFTW library and current computers, the Fourier transforms limit the bandwidth to 200 kS/s/channel. Switching to an integer FFT, we expect this number could be improved by a factor of a few. Likewise the ADC cards fit a standard PCI bus, so they could be moved to 1.5 GHz Linux / Intel boxes at modest cost. This too should improve throughput by a factor of a few. Thus it seems reasonable to suppose that 1 MS/s/channel would be achievable with modest investment.

A more serious bottleneck is the network. Controlled testing has shown that the current RPA network becomes saturated when each computer is shipping ~ 50 kS/s/channel. Due to interactions with the ADC driver, the system cannot run for long periods at this rate, hence we normally run at 20 kS/s/channel to improve reliability. With finite development time, the ADC driver could be improved to maximize the transfer rate. Additionally the network bandwidth could be enhanced in three ways:

1) Instead of transmitting 32-bit floating point values, we could gain a factor of 2 by transmitting in a 16-bit representation.
2) Upgrade the 100 Mb/s network to 1 Gb/s, gaining a factor of 10.
3) Set up multiple networks, potentially having exactly one network per antenna, gaining a factor of 7.

Implementing these changes has the potential of raising the bandwidth to 7 MS/s/channel, which is higher than the figure deduced in the previous paragraph. Thus, with the current ADC's, an upgrade to the computer hardware and networks, and a finite software development investment, it is reasonable to expect that a correlator with 1 MS/s/channel bandwidth can be built on the present infrastructure.

Cassiopeia A

As a final demonstration, a movie was made of Cassiopeia A drifting through the beam (azimuth 0.9°, elevation 69.2°). Here the signal is only a small fraction of the system temperature, so longer integration times were used (10 s per frame). Even so, the signal to noise ratio is only about 3. A plot of the fringes is displayed in Fig. 16.

A few stills from the Cas A movie are shown in Fig. 17. Note that the for the time period when Cas A is close to the center of the image, the signal to noise ratio improves noticeably. The reason for this is related to the pointing errors. When the source is close to image center, all antennas have nonzero signal amplitude. Since the detected amplitude increases as the square of the number of useful antennas, this makes a big difference.

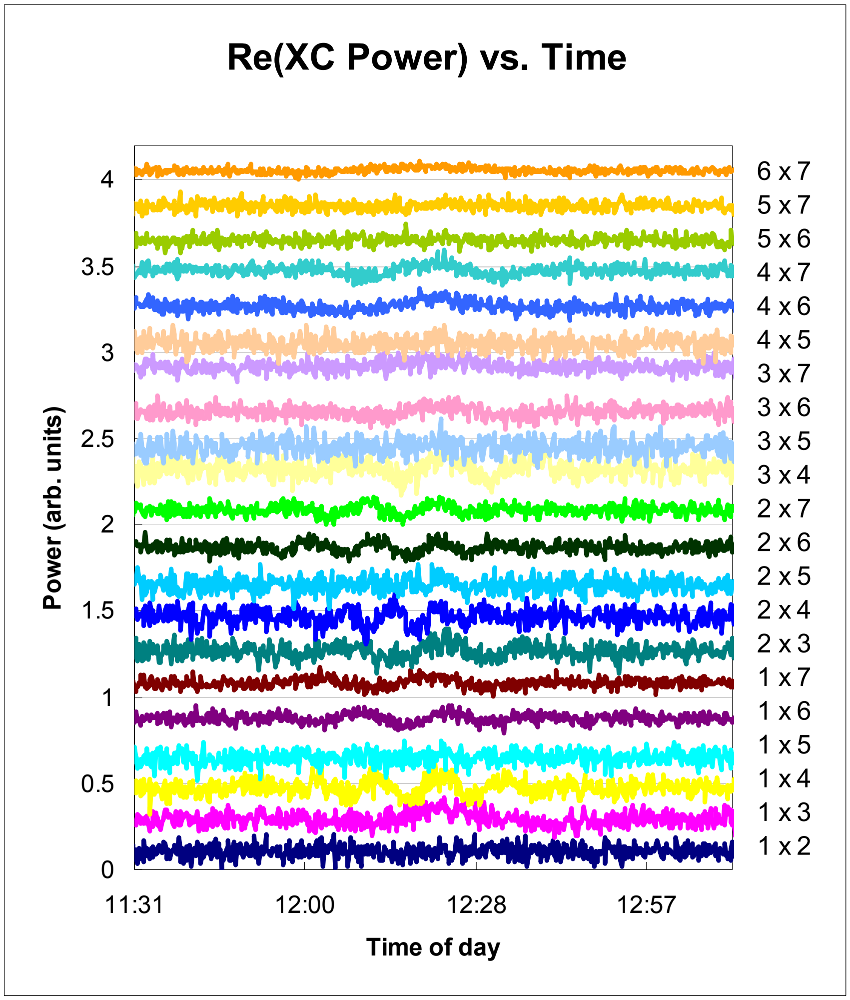

Fig. 16: Fringes on Cassiopeia A. Compared to Fig. 11, the integration time is increased by ~10 times. Even so, the signal to noise ratio is no more than a few.

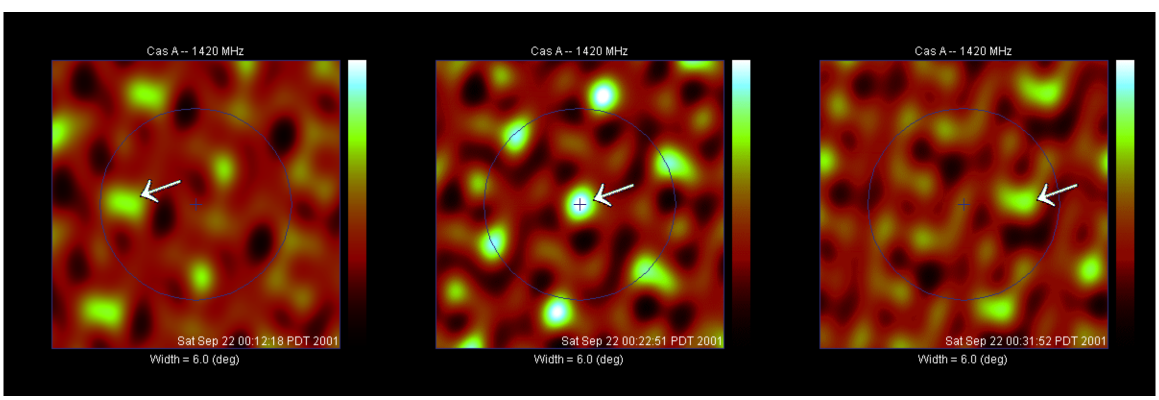

Fig. 17: Stills from Cas A movie. The low signal to noise ratio is especially apparent in still frames. In the movie, the eye integrates somewhat, giving a smoother impression.

Conclusions

The goal of this study is to demonstrate synthetic imaging at the RPA and to characterize this observatory as an imaging instrument. To this end, we have developed a prototype software correlator that operates continuously albeit at a modest data rate. We note that the present system could be readily extended to carry out beam forming with modest additional effort.

We acknowledge the fact that this instrument was not designed to be an imaging telescope, and indeed it is a poor one. All the images are dominated by strong grating sidelobes, which reduce the useful field of view to an area much smaller than the antenna primary beam. Even nonlinear techniques such as Clean will probably fail on these images since the strongest sidelobes are only about 5% weaker than the main beam.

Apart from the grating sidelobes, however, the telescope behaves well. Demonstrations of amplitude and phase closure suggest that the electronics are stable over periods of at least an hour. With the exception of RPA5 (which probably needs repair), the ADC's, computers, and network run reliably enough to perform useful measurements. We have not probed deeply for signals well below the system noise, but our initial results on Cassiopeia A make the prospect of detecting weaker sources appear promising.