"The MDC Systems Working Group Report"

A. R. Thompson et al

September 11, 1995

NRAO
MILLIMETER ARRAY
MEMO SERIES

NO. 142

REPORT OF THE SYSTEMS WORKING GROUP OF THE MMA DEVELOPMENT CONSORTIUM

A. R. Thompson (chairman), D. Bagri, B. G. Clark, J. E. Carlstrom, L. R. D'Addario, D. Emerson, R. P. Escoffier, J. Romney, S. Padin, R. A. Sramek, D. Thornton, and W. J. Welch.

September 11, 1995

The MMA system to be discussed here takes the signals from the front ends and performs all of the processing necessary to produce the visibility data at the output of the correlator. Thus it includes signal conversion to the intermediate frequency, transmission of the IF signals to the correlator location, filtering to obtain the required bandwidths, digital sampling, application of compensating time delays, and correlation. It also includes generation of all of the required local oscillator signals with appropriate phase changes for fringe rotation and phase switching. The antennas and the front ends including the feeds and quasi-optic components are also parts of the receiving system, but are the subjects of two special working groups.

The goal of the present report is to recommend ways in which the various electronic functions can best be implemented and to identify areas where prototype testing is required before choices can be made. It should be emphasized that recommendations are based on components and techniques that are currently available or within the state of the art, and thus the report provides a plan for immediate commencement of the development phase of the system described.

1.0 Specifications

Some specifications and guidelines that are relevant to the MMA system design are given below. These are taken from the MMA Proposal and from conclusions of meetings of the MMA Advisory Committee.

- Access to both sidebands of the SIS mixer responses and sideband separation at the correlator output are desired.

- Simultaneous operation in two bands (i.e. using two different front ends) is desirable, and is necessary for certain calibration procedures.

- Antenna spacings may be as large as 3 km.

- The correlator bandwidth for spectral line observations should be no less than 2 GHz total for a spectral resolution of 2 MHz. Finer spectral resolution should be available with reduced input bandwidth.

- A bandwidth greater than 2 GHz is desired for continuum observations.

- There should be at least four independently tunable input bands to the correlator for spectral line observations. The ability to select resolution bandwidths independently for these is very desirable.
- Fringe rotation is to be performed in the local oscillator system to allow for operation in a phased array mode.
- Single dish (total power) capability is required for both continuum and spectral line observations. Note that the term total power, as used in this report, includes switched power.

Avoiding unnecessary complexity in the system is an important consideration, particularly for electronics at the antennas. The array will be located at a high altitude site where physiological effects increase the effort of working outdoors on antennas, whereas it may be possible to mitigate such effects in the on-site electronics building. Also, the antennas will be relatively small and will not have such spacious rooms for electronics as the VLA and VLBA antennas.

2.0 Signal Transmission from the Front-Ends to the Correlator

We begin by considering the requirements for transmission of the IF signals from the front ends at the antennas to the correlator location, since this requires a number of choices that are critical to the overall performance of the array.

2.1 The Total Signal Bandwidth

The element of the system that mainly limits the overall bandwidth that can be processed is the correlator, which is the largest single item for which the cost is directly proportional to the bandwidth. The MMA proposal considers a total bandwidth for the correlator of 2 GHz for spectral line observations, with larger bandwidth for continuum observations desirable but unspecified. In section 6 and Appendices I and II of this report both lag and FX approaches to the correlator are discussed. A possible lag correlator design with 8 GHz continuum bandwidth is discussed in some detail in Appendix I. So far as we can tell, without doing a full design, the 8 GHz bandwidth is feasible with the present state of the art, and a further increase of bandwidth to 16 GHz should be feasible but would be challenging. Sixteen GHz is as large a bandwidth as it is useful to consider at this time, and we shall use it for the total bandwidth of the IF and transmission system. Note that 16 GHz is a goal, subject to verification in the more detailed stages of design.

2.2 Transmission Considerations

Single-mode optical fiber will be used for transmission of the signals from the antennas to the correlator, the maximum distance being about 2 km (unless the correlator is located off-site). We want to transmit over this medium the information corresponding to a signal bandwidth of 16 GHz. This can be sent as an analog signal, i.e. a noise waveform of the required bandwidth, B (= 16 GHz), or the same waveform can be sampled at the Nyquist rate (2B samples/sec) and transmitted as a digital signal of baud rate 2Bn where n is the number of bits per sample. The benefit that is offered by digital transmission is that the accuracy of the signal transmitted is much less dependent upon the frequency response of the transmission system or on instrumental noise introduced in the transmission, since to recover the signal it is necessary only to be able to distinguish between a limited number of voltage levels that represent digital bits.

For two-level quantization n=1 and the quantization degrades the signal-to-noise ratio by a factor of 0.64, reducing the effective bandwidth relative to the analog mode to 0.41B. For four-level quantization n=2, the reduction in sensitivity is 0.88, and the effective bandwidth is almost twice that for n=1. Thus four-level quantization is required to make effective use of the bandwidth. We shall assume that the correlator can handle a bit rate of 4B, corresponding to n=2, but note that this will need verification in the correlator design stage.

Optical systems designed for digital transmission with data rates of gigabits per sec. use two light levels only; we do not know of any currently available systems that use more levels. Thus such systems transmit only one bit at a time. A digitized noise signal in such a system has the form of a quasi-random squarewave with a minimum time interval between transitions of 1/4B sec. Such a waveform has a power spectrum of the form $[\sin(\pi f/2Bn)/(\pi f/2Bn)]^2$, where f represents frequency. It is not necessary to transmit the whole spectrum of the waveform extending from zero to infinity, and the usual practice is to truncate it at the first minimum, in which case the transmission bandwidth required is 4B. In the recovered waveform at the correlator location the limited bandwidth results in rounding-off of the sharp transitions of the waveform, but the information remains recoverable. Note that the bandwidth factor of four relative to analog transmission is a result of current engineering practice, not any fundamental characteristic of digital transmission.

Digital transmission systems can be purchased in which all of the engineering is complete and the user needs only to interface with an input connector at one end of the link and an output at the other. Most of these provide bandwidths considerably less than 2 GHz and a number of such units would be needed for each antenna. A quick look at prices suggests that this is not a feasible approach[1]. If we were to put together our own digital systems using two light levels the number of transmitters, receivers and fibers required would be four times that of an analog system. Attempting to develop a more bandwidth efficient technique than the industry has achieved at this time does not seem to be a sensible approach.

Before recommending against the use of digital transmission we should ask whether the possible

---

[1]For example, consider transmission of one 8 GHz-wide band. BCP (Broadband Communication Products) digital transmitter and receiver (51T-221and 51R-221) cost \$2.7k for the pair and can handle a little more than 1 Gb/s. Thirty-two would be needed to transmit 8 GHz, 4-level, with possibly two units sharing one fiber. Ortel analog transmitter and receiver (1530A and 2515A) cost approx \$13.5k for the pair and can handle up to 10 GHz, i.e. only one pair would be needed. The cost of fibers is not included in these figures.

benefits of this technique are sufficiently important to justify serious consideration of it. The main issue is whether the effects on the uniformity and stability of the frequency response that may be introduced by analog transmission seriously limit the required performance of the array.

2.3 Brightness Dynamic Range

One limitation on the achievable dynamic range in synthesis arrays is the variation of the instrumental frequency response from an ideal uniform level over the bandwidth of the signal. This effect (or more precisely the resulting differences in the instrumental responses of the signals from different antennas) results in closure errors in self calibration. The magnitude of these errors has been investigated by model analysis (Clark 1978, Thompson and D'Addario 1982) but relating these to dynamic range is not straightforward since the image processing involves nonlinear procedures. A way to relate instrumental effects to dynamic range is to examine the experience with the VLA. Some figures for the performance of the VLA in brightness dynamic range are given by Perley (1989). These indicate that use of self calibration results in a dynamic range of about 43 dB, and correction for closure errors about 49 dB. Phase errors in the quadrature networks of the digital samplers is the major limitation when using the continuum correlator, and use of the spectral correlator (with which the quadrature networks are not required) has resulted in a dynamic range of about 53 dB. At this level the limiting factors are not known but could include phase fluctuations (perhaps atmospheric) within the integration time and are not necessarily attributable to frequency-response variations.

Variations in the frequency response in synthesis arrays result from the electrical characteristics of the whole chain of electronics from the antenna to the digital sampler, including such minor items as cables, connectors, waveguide to coax adapters, etc. In the VLA, analog transmission is used and the TE_{01}-mode waveguide is an important contributor to ripples in the frequency response. In particular, the reflections between the coupler at each antenna location and the termination in the vertex room result in ripples of period a few MHz in the frequency dimension (Lilie, 1994), and typical peak-to-peak amplitude of 0.1-0.2 dB. In optical fiber, which will be used in the MMA, reflections are generally several orders of magnitude less than in waveguide or coaxial cable. One reason for this is that light reflected at a junction in a fiber is concentrated in a beam a few degrees wide, so by cutting the fiber surface at an angle the reflection is directed away from the return path down the fiber. Also, the runs from the antennas to the building will not be interrupted by couplers. Thus it can be expected that the performance of the MMA fiber system, using analog transmission, will be significantly better than that of the VLA waveguide. However, the dynamic range required for the MMA is in most cases expected to be less than for the VLA because of atmospheric effects, the lower collecting area of the MMA, and the greater difficulty in achieving low system temperatures and good phase stability at the higher frequencies. A goal of 50 dB is perhaps more than is required for the MMA, and based on the VLA experience this does not appear to require the use of digital transmission.

2.4 Spectral Fidelity

A second performance factor to be considered is the ability to separate spectral line features from continuum. The limit is set by the accuracy of calibration of the frequency response, since the effect of irregularities in the frequency response on the continuum can mimic true spectral features. We have used the term spectral fidelity[2] to denote this limit, expressed as a fraction of the total continuum level. Note that it is the variations in the response that occur over frequency intervals comparable to those of natural spectral features that are the most confusing, and slow, smooth variations that result from things like a gradual change in attenuation with frequency are a lesser problem. The fractional variation in power of the frequency response divided by the square root of the number of antennas should provide a rough measure of the spectral fidelity to be expected in an image.

Again, the most useful way to get some idea of the magnitude of the problem is to look at the VLA experience. A spectral fidelity of 50 dB is desirable for some H1 observations with the VLA and has been achieved, with great effort, in one particular case where a calibration source fell within the field being mapped (Van Gorkom et. al. 1993). Generally with the VLA a figure of 30 dB is obtainable with care, and 40 dB with great effort. At millimeter wavelengths the continuum is generally weaker and the lines are stronger than at centimeter wavelengths, so the requirement for high spectral fidelity should be less with the MMA than with the VLA. The ultimate requirement is likely to be for detection of a weak absorption line from gas in front of, say, a 10-Jy quasar, for which we estimate the requirement is about 40 dB. This consideration together with the better performance expected for the optical fiber again suggests that the desired performance should be achievable with analog transmission.

2.5 Discussion of the Digital-or-Analog Choice

The 16 GHz of analog transmission bandwidth desired for continuum observations would also accommodate up to 16 Gb/s of digital data with the current two-light-level technology discussed in section 2.2 Thus it could carry the 2 GHz of spectral line signals in digital form, which suggests the possibility of using analog transmission for the continuum and digital for spectral line signals. However, incorporating both transmission modes would be considerably complicate the system. All of the signal channelization would occur at the antennas, digital samplers would be required at both the antennas and the correlator location, and the local oscillator system would be significantly complicated. Another problem concerns multiplexing digital data from two or more IF channels into one data stream to feed to an optical transmitter, the bit frequencies being in the range of several Gb/s. Multiplexing ICs for such frequencies are very expensive and not widely available[3]. Overall, transmission of the IF data in digital form

---

[2]The term spectral dynamic range is sometimes used to refer to the same quantity (e.g. Van Gorkom and Ekers, 1989). However spectral dynamic range is also used to denote the limit set by intermodulation products formed from two lines by non-linearity in the electronic system, an effect which has not been identified in observations with the VLA.

[3]The company NEL of Japan makes a 4:1 mux (NLG4218) and a 1:4 demux (NLG4219), both of which operate up to 10 GHz but cost $3559 each. In principle another possibility would be to modulate the bit streams onto subcarriers at different frequencies and combine them as analog signals at the optical transmitter input. Use of QPSK (quadri-phase-shift keying)

would introduce a number of significant complications, and since this approach does not appear to be necessary it has not been pursued.

2.6 IF Signals and Channelization

It is proposed that the total IF bandwidth to be brought from each antenna to the correlator location be divided into four bands each 4 GHz wide. These would be independently assignable to the various front end outputs. This scheme would allow, for example, use of two frequency bands with two polarizations each, four bands with one polarization, and other combinations.

For spectral line observations it is desirable to be able to observe a number of lines simultaneously. This requires that the correlator can, in effect, be operated as a series of independent units, one for each line or closely spaced group of lines. An equal number of signal channels is required to provide the correlator inputs. For each such channel the bandwidth should be independently variable by factors of two as appropriate for the desired frequency resolution. A finely tunable LO will be required for each channel to set the center frequency. In this report we consider one such channel per IF signal, i.e. four in all but independently assignable to the IF signals. It would be possible to increase the number of input channels to the correlator by adding more channel hardware, but this can rapidly increase the cost of the system and can be considered an optional enhancement.

3.0 Suggested Design Details

Based on the recommendations discussed above we can now consider some design details and schematic block diagrams.

3.1 Signal Bandwidths from the Front Ends

A recent development at Caltech in which the first IF stage is integrated into the mixer unit has resulted in good performance over a 4 GHz bandwidth, with the IF covering approximately 0.5-4.5 GHz (Padin et al. 1995). It is believed that with the best InP HEMTs and the integrated design it may be possible to increase the bandwidth to as much as 8 GHz. Thus it would be good to consider a bandwidth of 8 GHz (in each sideband) for SIS front ends. Even if this is not achieved in the initial design it may well be possible for later front-end upgrades. As round numbers, 2-10 GHz will be used in this report for the IF output bands of SIS mixers.

The highest bandwidths of the HEMT front ends can be taken to be about 30 GHz, since the two highest frequency bands with these amplifiers are currently planned to be 60-90 and 90-115 GHz.

3.2 Intermediate Frequencies

For purely practical reasons it is very desirable to keep all intermediate frequencies well below 20 GHz. Switching of IF signals will be an essential function in selecting between different

---

modulation with truncation of the spectra at the first minima would require the same transmission bandwidth as digitally multiplexing the bit streams into one waveform, but would add a great deal of undesirable complication.

receiving bands. Coaxial switches, mechanical or solid state, are widely available up to about 20 GHz. Waveguide switches are almost the only alternative above this frequency and they are relatively clumsy and expensive. (Nevertheless it will be necessary to use some of them for switching LO signals.) Also, the upper frequency limit for most SMA connectors is 18 GHz. Other coaxial connectors are available for higher frequencies but they are much more expensive. Finally, as frequency increases the performance of mixers and amplifiers generally decreases, and prices increase. Any cost increase is multiplied by a large factor because there are 40 antennas and each requires four IF signals. Factors that tend toward increasing the intermediate frequencies are the need to separate the two responses at a mixer, and the difficulty in obtaining good matching between components over a large relative frequency range: note that 2-10 GHz covers more than two octaves. Within these constraints the aim is to keep the IFs as low as possible.

3.3 Bandwidth on a Fiber

The cost per unit bandwidth of analog systems (optical transmitter plus receiver) decreases a little as one goes to higher bandwidths. This points in the direction of using a small number of wide bandwidth links per antenna. Also, in an instrument in which bandwidth may increase with future upgrades, narrow band transmission components are likely to prove a poor investment. The bandwidths considered above could be handled by one fiber system of width 16 GHz per antenna. However, it is expedient to consider two fibers each carrying 8 GHz since then one does not completely loose an antenna if a fault develops in a fiber link.

3.4 Equipment at the Antenna

Figure 1 is a block diagram of a possible IF system at the antenna. The typical SIS front end is shown with four outputs, for two sidebands and two polarizations, each covering 2-10 GHz. (It is assumed that sideband separating SIS mixers will be developed for certain bands. For bands with conventional SIS mixers there will be just two IF outputs.) Any two of these outputs can be connected to the 2-10 GHz inputs of the two optical transmitters. Each transmitter also has separate inputs for the 2-6 GHz and 6-10 GHz halves of its band. This allows observations to be made using two front ends simultaneously. In this dual band mode the first LOs of the two front ends would be tuned so that the desired lines fall within 2-6 GHz in one front end output and 6-10 GHz in the other.

HEMT front ends have two outputs, one for each polarization. The lower 15 GHz of the HEMT band can be converted to 10-18 GHz using a first LO frequency tuning from f_c-25 to f_c-18 GHz where f_c is the center frequency of the front end band. Similarly the upper 15 GHz of the band can be converted to IF using a first LO tuning from f_c+18 to f_c+25 GHz. In each case the unwanted response (image) would fall outside the HEMT band[4].

---

[4]For a front end bandwidth of 30 GHz these numbers are chosen so that the nearest edge of the image is approximately one -3dB-bandwith from the center frequency. With a 6-pole filter response for the front end, the image should be approximately 40 dB down. Front end filters are not shown in the diagram.

The scheme in Fig. 1 is about the simplest collection of components that one can devise to fulfill the necessary IF functions at an antenna. It provides at least as much bandwidth as the most optimistic correlator design will be able to handle. It also provides for dual frequency observation, and for SIS front ends it requires only one frequency conversion per IF channel. There are no switches at frequencies greater than 10 GHz.

3.5 Channelization for Continuum Operation

The signals received at the electronics building of the array need to be filtered and converted to digital samples before going to the correlator. It will be assumed that digitizers with a sample rate of 4 GHz will be available for the MMA, and thus it is necessary to break down the IF bands into 2-GHz wide sections for the digital samplers. For continuum observations it is assumed that all of the IF band may be used. Figure 2 shows a system in which the 2-10 GHz band from an SIS front end is divided into four sections and digitized. Four fixed-frequency LOs are required, but since the equipment is all in one location (the electronics building) these frequencies would each be generated once and distributed by a network of power splitters and amplifiers. Two of the systems shown in Fig. 2 would be required per antenna, one for the output of each fiber link. The 0.1 GHz figure in the baseband frequency response represents the unavoidable low end roll-off, the actual value for which will depend upon the circuit details.

3.6 Channelization for Spectral Line Operation

For spectral line observations it is necessary to be able to select a band of frequencies centered at any desired point within the 2-10 GHz IF band, and to filter it to a bandwidth that is variable in factor-of-two steps from a maximum of 2 GHz. A commonly used scheme for selecting a small part of a wide IF band is to use a sideband-separating mixer (also called an image rejecting mixer) with a baseband (lowpass) IF immediately following it. This is done in the VLA and VLBA and the image rejection that is achieved in these systems is in the range 24 to 30 dB. For the MMA, which operates in a part of the spectrum where there are so many strong spectral lines, the minimum image rejection should be at least 40 dB. (This criterion follows from the same considerations as those for spectral fidelity in section 2.4) One way of increasing the image rejection would be to use 90° phase switching of the LO and separate the required sideband after correlation. Note, however, the switching does not suppress the image for total power observations. The alternative is to use filtering to obtain the rejection, as in the system shown in Fig. 3. Here several frequency conversions are used, and the IF after each mixer is sufficiently high that the unwanted sideband can be rejected by a filter. In Fig. 3 the required signal is first converted to a band from 10-12 GHz using an LO that is finely tunable from 14-20 GHz. As this LO is tuned over its range the lower sideband tunes across the 2-10 GHz input band. A 10 GHz fixed LO then converts the band to 0-2 GHz. The 10-12 GHz filter response must have steep enough edges that there is very little fold-over of the signal at the lower edge of the band, and the 0.1-2 GHz filter must be steep enough at the edge that very little aliasing results from the sampling. The signal is also passed through other filters with bandwidths progressively decreasing by factors of two. The center frequencies of these filters are also reduced to keep the percentage bandwidth 40% or greater. Two further frequency conversions are introduced to prevent the center frequencies from moving close to parts of the band where there may be some residual fold-over from previous frequency conversions. A range of bandwidths from 2 GHz to 31.25 MHz can be selected and digitally sampled. With the narrower bandwidths the 4 GHz sample rate provides redundant samples which can be discarded. If narrower bandwidths are required they can be obtained by digital filtering after sampling. Digital filter ICs are available for frequencies below about 50 MHz.

To obtain the desired 40 dB image rejection at each mixer in Fig. 3 careful selection of the filters and some adjustment of the frequencies shown will be required, but a design of this general type is certainly feasible. The disadvantage of the scheme is that the number of filters, mixers, etc. is expensive and tends to introduce gain and phase variations as a function of temperature. A scheme based on a sideband separating mixer alone is not good enough, but one incorporating 90° phase switching on the LO is certainly a competitor for this part of the system. Some development and testing will be required before a the best choice can be made.

The filtering and sampling of a unit such as that in Fig. 3 provides for observation of one line or group of lines with one polarization, and four such units are required as explained in section 2.6. Their inputs can be assigned to the optical receiver outputs in any way desired by the astronomer. A continuum system, as in Fig. 2, could also be fed from the same optical receiver to provide simultaneous line and continuum observations, limited only by the capacity of the correlator.

4.0 The Local Oscillator System

4.1 Reference Frequency for the First LO

The first LO, which provides the conversion from the frequency received at the antenna to the first IF, is the most critical one with respect to noise and phase stability because generating it requires a very large multiplication factor from a reference frequency. The reference frequency is transmitted from the electronics building to the antenna on an optical fiber, and some noise is added in the transmission process. This noise limits the frequency multiplication that can be applied at the antenna, i.e. it is a factor that places a lower limit on the transmitted reference frequency. The BIMA Array at Hat Creek uses a reference frequency of approx. 1.2 GHz (with a phase-locked YIG oscillator for noise filtering at the antenna) for LO frequencies up to 200-300 GHz. Thus something in the region of 2 GHz should be appropriate for the MMA in which the LO has to be multiplied up to 360 GHz, i.e. a multiplication factor of 180. With a fixed 2-GHz reference, the LO could easily be made tunable in 2 GHz steps. With a 4- or 8-GHz wide IF following the mixer, 2 GHz steps would allow any frequency to be set within and IF band, but might not allow a group of lines distributed over 4 GHz to be observed simultaneously. Greater tunablilty could be obtained at the expense of some complication by combining a small multiple of a second reference frequency, such as 100 MHz, into the LO frequency. However, the simplest way to obtain very fine tunability is to allow the reference frequency to be varied by a few percent. Again from the BIMA experience, a frequency synthesizer such as HP8662A should be a suitable reference source. Although a variable reference is frequently used in millimeter radio telescopes, two possible disadvantages should be pointed out. First, interaction between two oscillators or their harmonics can cause spurious responses if they fall within an IF band. With fixed frequency references such things are more easily found and corrected, but with tunable ones it is necessary to be able to predict potential problems in the oscillator settings and provide software that checks for them in the observing files. Second, a phase locked oscillator at the antenna provides a means of filtering the reference to reduce noise introduced by the transmission system, as well as providing sufficient signal level. The noise reduction requirement determines the bandwidth of the loop and for the greatest reduction the locked oscillator must be highly stable, implying a crystal controlled (i.e. fixed frequency) oscillator. (In the VLA a fixed reference is used to lock a 10 MHz crystal oscillator, which is necessary in this case because the two-way waveguide system requires that the reference be intermittent.) It is perhaps unlikely that these advantages of fixed references would be important in the MMA. If they are not important it would be a pity to avoid a variable reference, since it provides the best tunability and keeps the electronics simple. Two independently tunable first LOs are required at each antenna to allow two separate front ends to be used simultaneously, or to allow the two polarization inputs for any single front end to be independently tunable to extend the range of instantaneous coverage for a single band.

4.2 Phase Stability Goal

The MMA proposal document suggests a figure for the overall phase stability for the array of 0.3 radians at 345 GHz, which corresponds to a time stability of 0.23 psec. The Phase Calibration Working Group have suggested that a useful phase stability specification would be one based upon variations on a time scale of 1 to 100 sec. This is appropriate for observations in which a calibrator is observed at intervals of 100 sec, and we consider such a figure here. Without some testing of a protype system it is very difficult to estimate the level of phase stability that can be achieved. Instead of attempting such a prediction, we have considered what would be a desirable goal for phase variations on a 1 to 100 sec time scale. We start by considering an overall phase variation of 9° rms which results in a loss in correlation of only 5%. We then sugggest that variations from the electronics system would be tolerable if they contribute no more than 10% to the overall phase variation, and allocate 8.1°rms (i.e. 90% of 9°) to the combined contribution of the atmosphere and antennas. Then the electronics can contribute 4°rms (the rms combination of 8.1° and 4° is 9°) which corresponds to 11 μm of electrical path at 300 GHz or a time interval of 3.7×10^{-14} sec.

If, at an antenna, there are 10 m of unburied fiber with temperature coefficient of path length of 8×10^{-5} per °C (Frye et.al. 1995), then a change of 11 μm in path results from a temperature variation of 0.014°C. Antennas may be several km apart so relative temperature changes between them could be of this order over times of 100 sec. In practice the fiber will be thermally insulated to some extent, but clearly a calibration scheme will be needed to monitor changes in the electrical lengths of the fibers carrying reference frequencies to the antennas. One must also consider the effects of all of the rest of the electronics including amplifiers, multipliers and filters that cary the LO signals (see Frye et. al., 1995, for examples of phase stability). It is difficult to predict compliance with the phase stability specification because the time scale is critical and thus the rates of variation must be known. It would also be useful to know how much the temperature changes vary from point to point at a potential MMA site. Features such as a local peak in the landscape could cause differences in times of sunrise and sunset and modify wind patterns over a site, and thus affect the ambient temperatures at individual antenna locations. Such temperature differences could have serious effects on the structural dimensions of the antennas as well as the stability of the fibers.

To monitor the effective electrical length of a fiber a round-trip phase measurement system is used. The round-trip system on the fiber at Hat Creek can detect path length changes of order 10 μm (1° at 100 GHz), which is close to the goal suggested above for the for the overall instrumental phase. Thus it appears that the phase variation budget is extremely tight and just about practicable if other contributions are small compared with that from the first LO.

4.3 Round-Trip Phase Calibration

In the round-trip phase scheme it is necessary to be able to separate the frequencies traveling in the two directions. One can use two fibers, one for each direction, but it has to be assumed that they behave identically. An alternative is to use one fiber with a different laser wavelength in each of the two directions, but there can be a problem from the difference in dispersion in the fiber at the two wavelengths. It has also been suggested that one can use the same fiber and the same wavelength band in the two directions, but with different frequencies modulated onto the laser signals. Testing will be required to determine which of these three approaches is the best.

4.4 Single- and Double-Sideband Operation

With the SIS mixers both sidebands of the received signal that result from the frequency conversion can contribute to the IF signal. If the wanted signal falls within one sideband, the best mode of operation is to remove the other sideband so that it does not contribute noise or other unwanted signals to the output. The input contribution of the unwanted sideband, i.e. astronomical signal and noise from the antenna and atmosphere, can be filtered out using a Martin-Puplett interferometer and replaced by noise from a cold load. Alternatively, a sideband-separating mixer may be available in which case there are separate outputs for the different sidebands. In either case the unwanted sideband is likely to be reduced by about 10 dB only, which is enough to prevent most of the sensitivity loss that results from the noise received by the unwanted sideband. However, 10 dB is not enough to suppress unwanted lines, and for this the two sidebands can be separated at the correlator output by introducing a sequence of 90° phase shifts at the first LO, or the unwanted sideband can be eliminated by using an LO offset scheme suggested by Barry Clark[5]. Note that these two schemes are only effective for the correlated component of signals from different antennas, and do not separate the noise. Also they do not work for the autocorrelation outputs of the correlator that are required for total power observation of spectral lines.

---

[5]In Barry's scheme a small frequency offset δf is inserted into the first LO, and then for the wanted sideband it is taken out again at a later LO. As a result the unwanted sideband suffers an offset $2\delta f$. Then if δf is made slightly different for each antenna the unwanted sideband signals become uncorrelated from one antenna to another.

Double sideband operation, using an SIS mixer without Martin-Puplett filtering or sideband separation, is an alternative method of operation. For spectral line observations it is necessary to use the 90° phase switching scheme mentioned above to separate the sidebands. To compare the sensitivities in double and single sideband modes we need to define system noise temperature for three different cases:

T_1 = single sideband noise temp. (unwanted sideband terminated, or separated in the mixer)

T_2 = Double sideband noise temp. as measured with signal entering both sidebands

$T_3 = 2T_2$ = Double sideband noise temp. measured with signal entering only one sideband

First consider continuum observations. The single sideband sensitivity is proportional to $1/T_1$. The sensitivity for double sideband observation, when the sidebands are separated by 90° phase switching, is proportional to $1/T_3 = 1/(2T_2)$ for each sideband. For the combined outputs of the two sidebands[6] it is proportional to $1/(\sqrt{2}T_2)$. It has been pointed out by the Front-End Working Group that in the higher frequency bands the instrumental noise limit set by quantum effects exceeds the sky noise for excellent sites such as Chile. In circumstances like this, where the instrumental noise is dominant, T_1 is approximately equal to $2T_2$ and one is better off to use the double sideband mode. On the other hand, if the atmosphere dominates, then T_2 can approach T_1 and the single sideband mode yields better sensitivity.

For spectral line observations the single sideband sensitivity is again proportional to $1/T_1$, but the double sideband sensitivity with sideband separation is proportional to $1/(2T_2)$ since we are not combining the two sidebands. Then in cases where the receiver noise dominates and T_1 approximates $2T_2$ the sensitivities appear to be equal, but double sideband observing with sideband separation is actually better because it provides twice as much frequency coverage, and thus possibly twice as many lines. (A factor of two in frequency coverage is like two in observing time or $\sqrt{2}$ in sensitivity). When the atmosphere dominates and T_2 approximates T_1 single sideband observing has the advantage. Even in the best sites the atmosphere can dominate at frequencies near the edges of the atmospheric windows. Thus it is important to have the option of single or double sideband modes to allow the best choice for any situation.

4.5 Instrumental Considerations

There are also some instrumental considerations involved in the choice between single and double sideband operation. If one wants to observe in one sideband only, the fringe rotation and

---

[6]The sensitivity is also proportional to $1/(\sqrt{2}T_2)$ if the sidebands are not separated, but if sideband separation is possible, it is advantageous to use it in a double-sideband continuum observation, and then to combine the images in the brightness domain. This procedure reduces the bandwidth smoothing, and comparison of the individual images for the two sidebands can be an aid in detecting the presence of errors such as interference.

Doppler tracking can be introduced at any convenient LO in the signal chain. If simultaneous observation in both sidebands is desired, using 90° phase switching to separate the outputs after correlation, then Doppler tracking can be introduced only on the first LO. The fringe rotation on the first LO should be set to stop the fringes at a sky frequency equal to the first LO frequency. To stop the fringes completely it is then necessary to insert a second fringe rotation in a later LO, to take care of the residual fringe rate that occurs because the observing frequency differs from the first LO frequency. In single sideband observing the fringes can be completely stopped for one channel by fringe rotation on the first LO, but since the other three channels are centered on different parts of the IF they have residual fringe rates. These can be removed by fringe rotation on a later LO for each particular channel. (In the case of double sideband observation without sideband separation the fringe term is a product of two sinusoidal functions one of which is slowed to zero by fringe rotation on the first LO and the other by fringe rotation on a later LO; see, e.g. Thompson et al. 1986).

To deal with the residual fringe rates, fringe rotation can be applied to one of the later LOs on an individual basis for each channel. With four channels and 40 antennas this requires 160 residual-fringe-rate generators each with a voltage controlled oscillator and a phase-locked loop to offset an LO. If this is not done the residual fringe rates appear at the correlator output, modulating the visibility data. The maximum offset of the sky frequency from the first LO frequency allowed in SIS front ends in Fig. 1 is 10 GHz, and with the maximum baseline of 3 km the highest residual fringe rate that will be encountered is 7.3 Hz. This limits the integration time at the correlator output to something like 10 ms. In principle it is possible to perform fringe rotation in software at this point, and then further integration. However, with the large number of channels required in spectral line observing it is not clear that this approach is preferable to provision of hardware for fringe rotation of 160 LOs.

4.6 Phase-Locked Loop for the First LO

Figure 4 is a simplified block diagram of the phase locked-loop for the first LO. A synthesizing unit generates a frequency of a few MHz with the required frequency offset for fringe rotation and the phase shifts for phase switching. Phase switching includes both 180° shifts for removal of the effects of DC offsets in the digitizing and 90° shifts for sideband separation. In the frequency offset and phase shifts allowance must be made for frequency multiplication outside the loop. The synthesized signal is used as an IF reference in the loop and its frequency is added to that of a harmonic of the first LO reference, which is assumed to be tunable in this scheme. As much of the frequency multiplication as possible should be included within the phase-locked loop. In practice it may be necessary to have some multiplication outside the loop, as shown in Fig. 4, if the response of the harmonic mixer does not extend as high as the required LO frequency.

4.7 Local Oscillators at the Electronics Building

The mixers in Figs. 2 and 3 are all part of the equipment in the building in which the correlator is located. Each LO frequency needs be generated only once and distributed to the mixers using a network of power splitters and amplifiers. In the case of the 14-20 GHz tunable oscillator in Fig.

3, a commercial synthesizing signal generator can be used. The additional fringe rotation needed to remove the residual fringe rates discussed in section 4.5 could be inserted into one of the LOs in Figs. 2 and 3, a fixed-frequency one being the most convenient.

If digital transmission were used for the spectral line signals the components in Fig. 3 would be located at the antennas. The LOs in Fig. 3 would be required at each of the 40 antennas and the simplest thing to do would probably be to send them out from the electronics building on fibers. Two reference frequencies for the first LOs are required in any case, but it would probably not be good to put more frequencies on the fiber (or fibers) carrying them. Thus at least one more fiber link would be required to each antenna. At the antennas, the LO signals would need filters and amplifiers, or phase-locked oscillators, for amplification and cleanup. Alternatively the LOs could be synthesized at the antennas using fixed, phase-stable references. In either case the additional complication is another reason for preferring analog transmission.

5.0 Number of Fibers per Antenna

Two optical fibers will be required to carry the IF signals, one or two for the first LO references and one for other reference frequencies, timing and possibly monitor and control. A special fiber may be required for the return path of the round-trip phase calibration, and one spare should be provided. Thus for each antenna six or seven fibers will be required.

6.0 The Correlator

Two studies of possible correlator designs have been made, one for a lag (XF) correlator by Ray Escoffier and one for an FX correlator by Larry D'Addario. Memoranda giving the full details of these studies are given in Appendices I and II of this report. It is generally agreed that during the development stage of the project designs based on both approaches should be pursued, including detailed architecture of the custom VLSI chips required. A choice can then be made on the basis of predicted cost and performance. Brief descriptions of some of the main points are given below.

6.1 Lag Correlator

The lag correlator design is based on the specification of 1024 lags (channels) for a total IF bandwidth of 2 GHz without polarization cross products. By halving the input bandwidth the number of lags is doubled, and a table of bandwidths, numbers of lags, and spectral resolution for various configurations of this system is given in Appendix I. The design that is described in Appendix I has a total bandwidth of 8 GHz, made up of inputs from four 2-GHz wide bands. At maximum bandwidth the number of lags for each of these four bands is 256, and the corresponding spectral resolution is 7.8125 MHz. This is excellently suited for continuum observations for which one wants the widest possible bandwidth and the ability to omit parts of the spectrum where strong lines occur. A total input bandwidth of 2 GHz can be handled as four 500-MHz wide bands, each of which has 1024 lags, providing a total of 4096 lags and a resulting spectral resolution of 488 kHz. This exceeds the original correlator specification for 2 GHz bandwidth by a factor of 4 in resolution. It is therefore suggested that the number of lags in the individual correlators could be halved and the total bandwidth capability increased to 16 GHz.

This figure has therefore been used in considering the IF and signal transmission requirements for the system, as discussed in section 2.1.

The proposed lag correlator would use digital samplers with 4 GHz sample rate, four being required for the 8 GHz bandwidth design (or eight for the 16 GHz design). The outputs of the digital samplers would be demultiplexed into 32 signals with 125 MHz clock rate, which would be the speed of the correlator chip. Each of these 32 signals would contain 2 bits per sample, either 3- or 4-level, depending mainly upon the design of the digital sampler. The sample streams would then go to RAMs capable of supplying a range of delays up to 32 μs. From the RAMs the data would be transferred to the correlator chips in appropriate sequences. The correlator chip would be a 4x4 matrix of basic correlator circuits that run at 125 Mb/s, each providing up to 256 lags. It would also contain some short term (up to 1 ms) memory for integration of the outputs. For the 8 GHz design the number of correlator chips required is 12,800 and these could be mounted on about 200 to 400 boards. The interconnections to bring the 125 Mb/s signals into these boards involve a very large number of cables (see section VIII of Appendix I), and must be considered one of the limiting factors of the overall correlator size. It is estimated that the level of integration of the correlator chip is about twice that of the chip designed by J. Canaris that is currently being used for the construction of the correlator for the Green Bank Telescope. The price of this GBT chip is roughly estimated to be $200 each.

With 780 baselines, 1000 (or more) spectral channels, and integrators for both states of the 90° phase switching, each output dump of the correlator for spectral line observations involves about 1.6×10^6 complex numbers. It is desirable to keep the data rate within the capacity of a single recorder, so with a tape recorder of 8 Mbyte/sec capacity it is suggested that the visibility data be recorded at intervals of about 3 sec.

6.2 FX Correlator

The FX design of correlator offers the advantage that a large part of the computation required is carried out in FFT circuits the number of which is proportional to the number of antennas, whereas in the lag design all of the computation is done in circuitry in which the quantity requirement is proportional to the number of baselines. Thus the FX design offers the possibility of reduced chip count and reduced cost. Disadvantages of the FX approach are that it is less well adapted to VLSI implementation; the FFT stages convert the two-bit data samples into complex numbers that require about 16 bits, thus increasing the arithmetic requirements of the cross multiplying stages and the complexity of the interconnections to them; and the FX architecture is generally less adaptable to implementing different modes for polarization and other special measurements.

The discussion in Appendix II considers 2 GHz overall bandwidth for which the spectral resolution is 1 MHz, and adopts the same basic clock speed of 125 MHz as the lag correlator study. The required functions are considered as the signals progress through the correlator, but in this case we do not have a state-of-the-art design for a similar chip, such as the GBT chip provides in the case of the lag design. Nevertheless, it is estimated that the overall chip count is no greater than required for the lag design, and could possibly be lower by a factor of four. The FX design therefore merits further study.

6.3 Other Approaches

The idea of separate digital correlators for continuum and spectral line observations has been mentioned in some MMA discussions, but is rejected because it is clear that the large number of cross multipliers used for a spectral line system can also provide the large bandwidth at low spectral resolution required for continuum observations. Thus one instrument will do both jobs well, and for a given cost it will have more capacity than either of the separate ones. The idea of an analog correlator for broadband continuum use is also rejected, for two reasons. First the bandwidth has to be broken down into bands no more than, say, 500 MHz width to avoid smearing in the synthesized image. For 16 GHz total bandwidth and 780 baselines, 24,960 broadband multipliers would be required. Such replication is best handled with digital technology. Second, it is difficult to obtain the precision required for high dynamic range with analog circuitry, especially if an analog delay system is also used.

7.0 Total Power Observation

The specifications of the MMA call for the antennas to be instrumented for total power measurement (i.e. single-dish operation) as well measurement of cross correlation in the usual interferometer mode. One reason for this is to provide information at spacings shorter than the minimum feasible between two antennas. The resulting visibility measurements near the (u,v) origin are essential for broad field imaging and mosaicing (Cornwell et al., 1993). It has also been suggested that it may be useful to operate the array as 40 independent single dishes for certain low resolution measurements, and when the atmosphere is so turbulent that forming an array beam is impossible.

In total power operation the gain stability required to reach the noise limit is approximately $1/\sqrt{(\text{bandwidth} \times \text{integrating time})}$. We assume that for total power operation the available IF bandwidths should include all those provided for interferometry and that for spectral line observations the autocorrelation mode of the correlator will be used. Thus the bandwidth ranges from 6 kHz, which is about the narrowest required frequency resolution, to 16 GHz. For an integration time of one hour the range of required gain stability is approximately 2×10^{-4} to10^{-7}. Beam switching at a rate of about 1 Hz will be necessary to reduce the effects of atmospheric irregularities. Thus it is gain fluctuations on time scales of order 1 sec that will limit the sensitivity.

Values of gain stability achieved with the BIMA array are given by Frye et. al. (1995). For the BIMA IF system, from the output of the front end to the output of the power detector, the temperature sensitivity is 1.3×10^{-2} per °C. The temperature stability is 0.01°C over one hour and the corresponding gain stability over an hour is 1.3×10^{-4}. Also, in a gain stability measurement of one of the BIMA front ends at 100 GHz, a drift of 10^{-3} was observed over three hours. If the MMA achieves similar stability, and the fluctions have a spectral dependence proportional to 1/frequency, the expected variation on a time scale of 1 sec is of order 10^{-7}. Thus with great care in the temperature control there is hope of achieving the required gain stability. For continuum observations with the highest bandwidths, it may useful to put a power detector at the antenna to eliminate gain variations in the optical fiber system and the subsequent IF stages.

As noted in section 4.4, 90° phase switching and frequency offsets cannot be used to eliminate unwanted (image) sidebands for single dish observations since they only affect the cross correlation of signals between different antennas. Thus with SIS front ends the unwanted sideband will always be present, although it may suffer attenuation of order 10 dB as a result of filtering or use of a sideband separating mixer. We have not found any way to completely eliminate the unwanted sideband for total power observations, but some ways of mitigating its effect can be suggested. The problem is presumably most serious in spectral line observations when the unwanted sideband also contains lines. If a band that is free from spectral lines can be found a few GHz from the band under study (perhaps a remote possibility in most cases) the first LO can be set to a value midway between these two bands, so the unwanted response will fall in the line-free band. The first LO must be finely tunable, and another LO further down the signal path must also be finely tunable to allow the part of the IF band within which the signal falls to be selected by the final filters. A similar technique, known as sideband smearing and sometimes used on the 12-m antenna at Kitt Peak, is to sweep the first LO over a range that is smaller than the first IF bandwidth, and remove the sweep for the wanted sideband in a later LO. The unwanted sideband then suffers the combined sweep of the two LOs which smears any lines within it. Note that both of these techniques require that the first LO be finely tunable, which argues in favor of using a tunable reference rather than a fixed one, as discussed in section 4.1.

The most important use of total power measurements is likely to be for short spacing visibilities in wide field imaging which is fundamental to the operation of the array. The number of antennas and the amount of observing time required to obtain the short spacing visibilities is somewhat difficult to define concretely. The most detailed discussion on this subject is by Holdaway and Rupen (1995) who conclude that the amount of total power data required in an image depends upon whether the astronomer is primarily interested in the broad features or the fine structure. For good imaging of the broad structure Holdaway and Rupen estimate that with all 40 antennas operating in the total power mode the ratio of total power time to interferometric time required is typically about 1 to 4. This estimate is based on the equalization of sensitivity in the visibility data in the (u,v) plane. For observations where the small scale structure are of most interest the total power time could be reduced by a factor of 10-20. It seems that all antennas may need to be equipped for total power observation including the beam switching capability.

8.0 ALC and Gain Calibration

In the VLA an automatic level control (ALC) circuit is used to hold constant the rms signal-plus-noise level at the digital sampler. Since the system noise temperature varies as the antennas track, the ALC causes the gain to vary. The gain can be calibrated by a noise signal that is switched on and off at a frequency of a few Hertz and inserted into the front end. The level of the IF signal for both the on and off levels of the noise source is monitored at a detector just before the sampler. This provides continuous measurement of both the gain and the system temperature. The purpose of the ALC is to hold constant the mean level of the waveform being digitized relative to the quantization levels, which are fixed. The gain variation can be corrected for in the subsequent processing. In the BIMA interferometer the preferred mode of operation is different. The gain is adjusted at the beginning of an observation and the good temperature stability is relied upon to keep it constant. There is no definite opinion at this time as to which method would be best for the MMA. Since the ALC circuitry and the switched noise source are small items they should be included in the system, with the capability be turned on or off.

9.0 Remote Location of the Correlator

It has been suggested that if the array site is at a very high elevation, such as 16,000 ft., it might be advantageous to locate the correlator at a lower elevation, perhaps 20-30 km from the array, where the physiological effects of the reduced atmospheric pressure would be less stressful. The signals would run in optical fibers from the antennas to the correlator site, and the increased optical attenuation would be approximately 0.4 dB per km. (The effective attenuation of the detected electrical signal would be 0.8 dB/km). This might mean that more fibers would be required to carry the full bandwidth.

10.0 Conclusions

The design of the receiving system considered here is essentially straightforward, at least as far as operation in the interferometer mode (i.e. the cross correlation measurements) is concerned. Some questions remain regarding the total power measurements. These concern the problem of maintaining constant gain in the electronics chain from the front ends to the digital samplers or to a power detector located in the final IF stages. The use of beam switching reduces the time scale over which gain variations can be damaging to something less than a second, but we can only make a very rough estimate of the magnitude of variations on this time scale. Some further discussion on the amount of total power observing required is desirable. For example, should estimates of the total power requirement be based on reaching the system noise limit of sensitivity, or should allowance made for atmospheric effects?

10.1 Thermal Stability

The requirements for thermal stability of the MMA will be more stringent than those of most other instruments that have been developed by NRAO. There are two main reasons for this. First the frequencies are so high that serious phase changes can result from very small thermaly-induced changes in antenna structures, front ends, and reference transmission and multiplier components, as mentioned in section 4.2. Second, the total power observing requirement places severe constraints on amplitude stability, especially with the wide bandwidths possible for continuum observations, as discussed in section 7.0. It is therefore recommended that a special engineer or engineering group be devoted to thermal control of critical areas for the whole array system.

10.2 Development Projects

In a number of areas more than one way of approaching the requirements exists. In cases where the relative advantages are well understood we have identified the best choice. In a number of cases tests on a prototype system are required to verify conclusions and make further choices. These tests will be part of the development phase of the project and are listed below.

- Measurement of the frequency response of the transmission system (optical transmitter and receiver with about 2 km of fiber). Variations on a frequency scale comparable with natural spectral lines should be less than in the VLA waveguide, i.e. less than 0.1 dB peak-to-peak. Similar measurements including as much of the full analog signal path as possible should also be made.

- Measurement of the phase variation introduced in a frequency of about 2 GHz by transmission over 2 km of fiber. If a tunable reference frequency used, is the reference received at the antenna stable enough in phase when multiplied up to 360 GHz for the first LO? If not, what kind of filtering is required or is a fixed frequency necessary?

- Investigation of transmission of two reference frequencies over the same fiber using the same transmitter and receiver. What relationships between the frequencies result in intermodulation that would be a problem in the LO system?

- Comparison of the three methods suggested in section 4.3 for measuring the round-trip phase of a reference frequency over the fiber link. Which one gives the best separation of the signals traveling in the two directions? The tests should be made under conditions of varying temperature to simulate real conditions, and the accuracy of the technique determined.

- Testing of a prototype first LO system like that in Fig. 4. Is the phase noise low enough at the highest frequencies? Should the YIG oscillator be replaced by a higher frequency Gunn oscillator?

- Testing of a prototype IF filtering scheme, like the one in Fig. 3, for rejection of unwanted responses and the presence of LO harmonics or intermodulation products within the signal bands. How does this scheme compare with the use of a sideband-separating mixer with 90° phase switching, or LO offsets, to enhance the image rejcction?

- Measurement of gain variation with temperature and power supply voltage for front end and IF components. This will provide data for estimation of temperature stability, beam switching rate and other parameters required in total power observations.

- Studies of lag and FX correlator designs including the necessary custom chips to enable a choice between the two systems to be made.

References

Clark, B. G., Closure - Some Examples, VLA Electronics Memo. 171, April 1978.

Cornwell, T. J., Holdaway, M. A., and Uson, J. M., Radio-interferometric imaging of very large objects: implications for array design, Astron. Astrophys. 271, 697-713, 1993.

Frye, B., Forster, R., Lugten, J., Mundy, L., Plambeck, R., Thornton, D., and Welch, J., Gain and Phase Stability of Some components in the BIMA Array, MMA Memo. 131, July 1995.

Holdaway, M. A. and Rupen, M. P., Sensitivity of the MMA in Wide Field Imaging, MMA Memo. 128, June 1995.

Lilie, P., Why the "3 MHz" Ripple Moves With Time, VLA Test Memo. No. 190, Oct.1994.

S. Padin, D.P.Woody, J.A.Stern, H.C.LeDuc, R.Blundell, C.Y.E.Tong, and M.W.Pospieszalski, An Integrated SIS Mixer and HEMT IF Amplifier, Sixth Internat. Symp. on Space Terahertz Tech., Pasadena, CA, March 1995.

Perley, R. A., High Dynamic Range Imaging, in Synthesis Imaging in Radio Astronomy, R. A. Perley et.al. eds., Astron. Soc. Pac. Conference Series, 1989, see pp. 292, 311.

Thompson, A. R. and D'Addario. L. R., Frequency Response of a Synthesis Array: Performance Limitations and Design Tolerances, Radio Science, 17, 357-369, 1982.

Thompson, A. R., Moran, J. M., and Swenson, G. W. Jr., Interferometry and Synthesis in Radio Astronomy, Wiley, 1986 and Krieger, 1991, 1994, see p.151.

Van Gorkom, J. H. and Ekers, R. D., Calibration and Analysis of Spectral Line Image, in Synthesis Imaging in Radio Astronomy, R. A. Perley et.al. eds., Astron. Soc. Pac. Conference Series, 1989, see p. 343.

Van Gorkom, J.H., Bahcall, J.N., Jannuzi, B.T., and Schneider, D. P., A Very Large Array Search for Emission from HI Associated with Nearby Lyman α Absorbers, Astron. J., 106, 2213-2213, 1993.

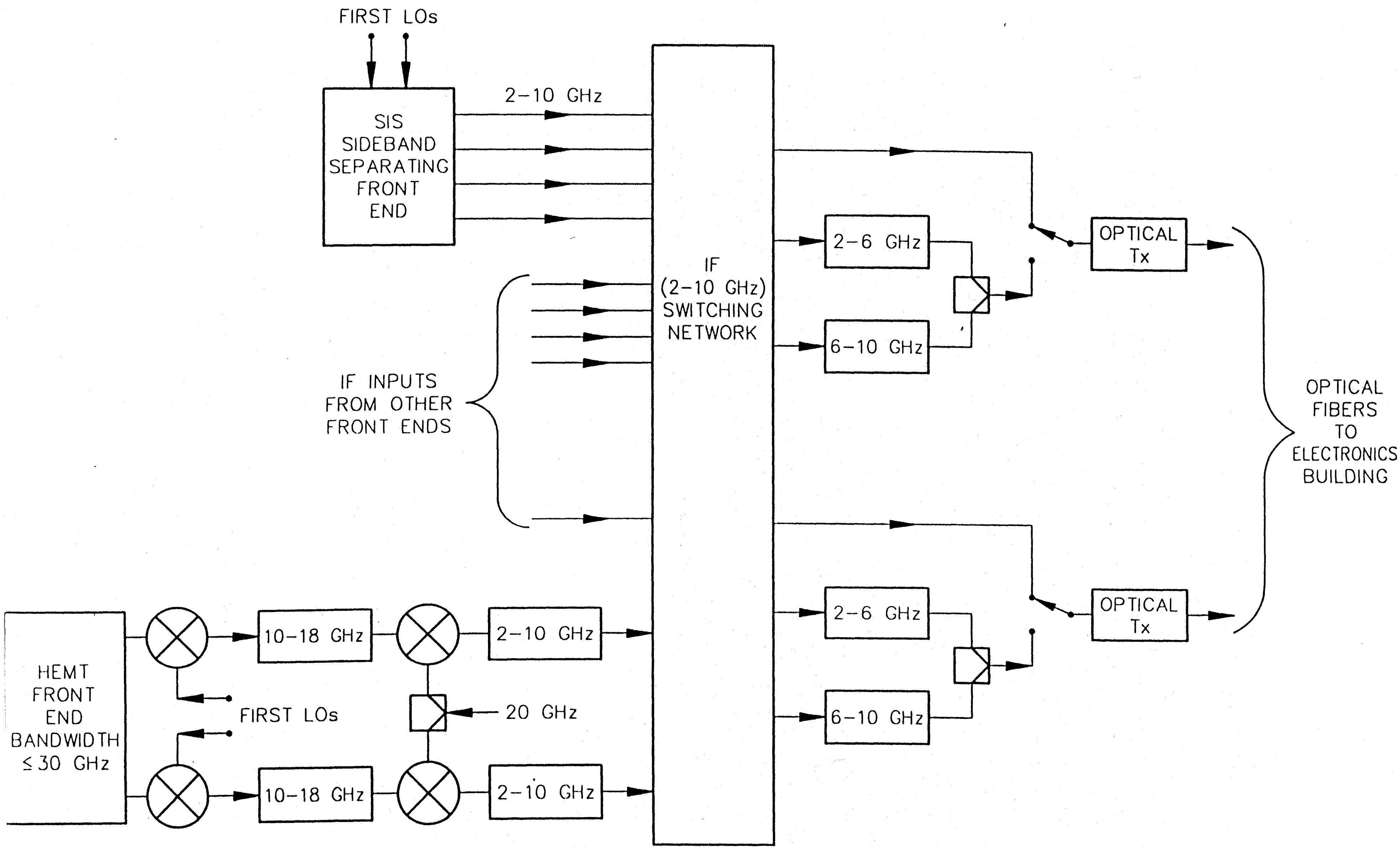


Figure 1. The IF system at an antenna for analog transmission over two fibers to the correlator building.

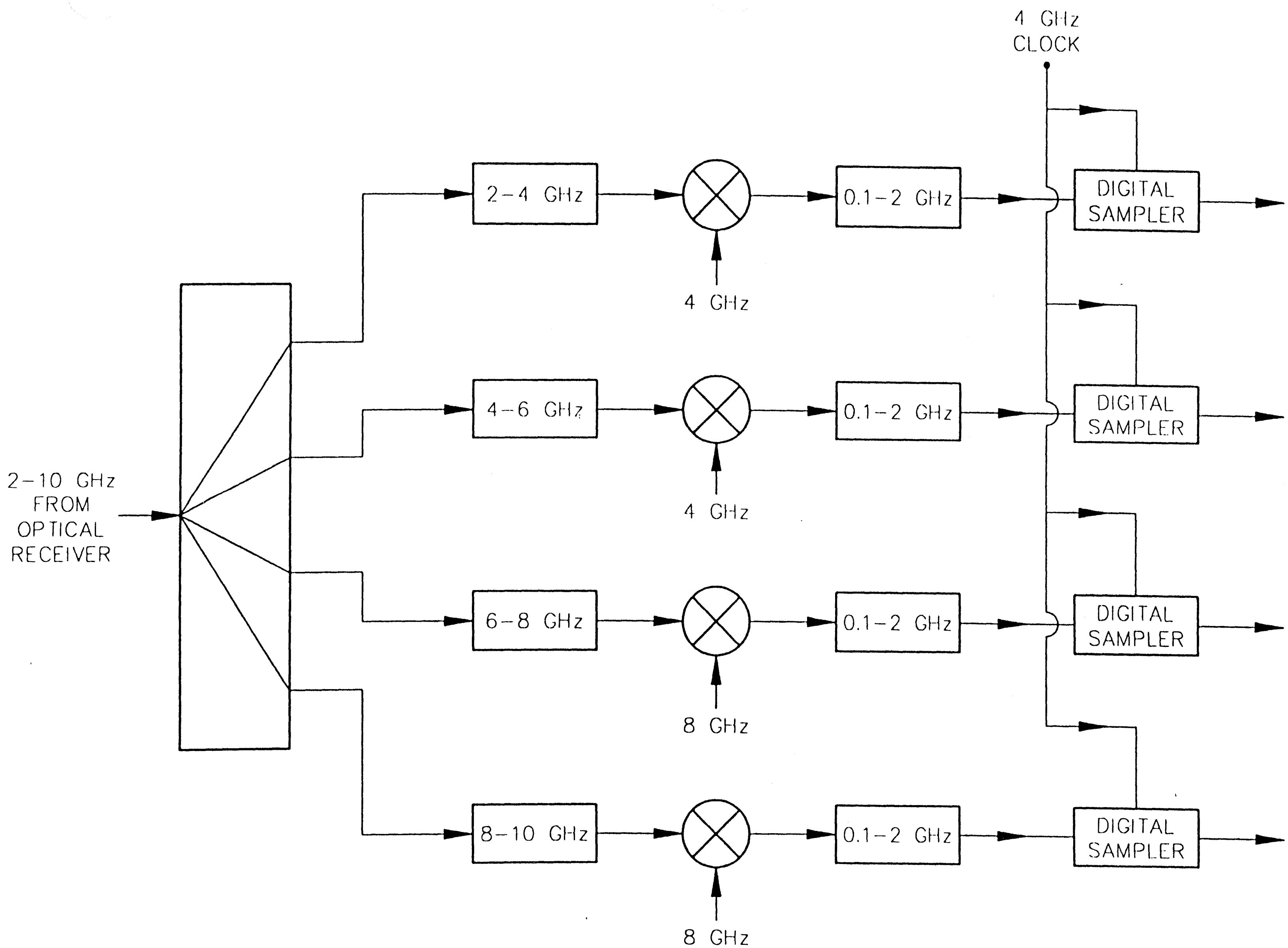
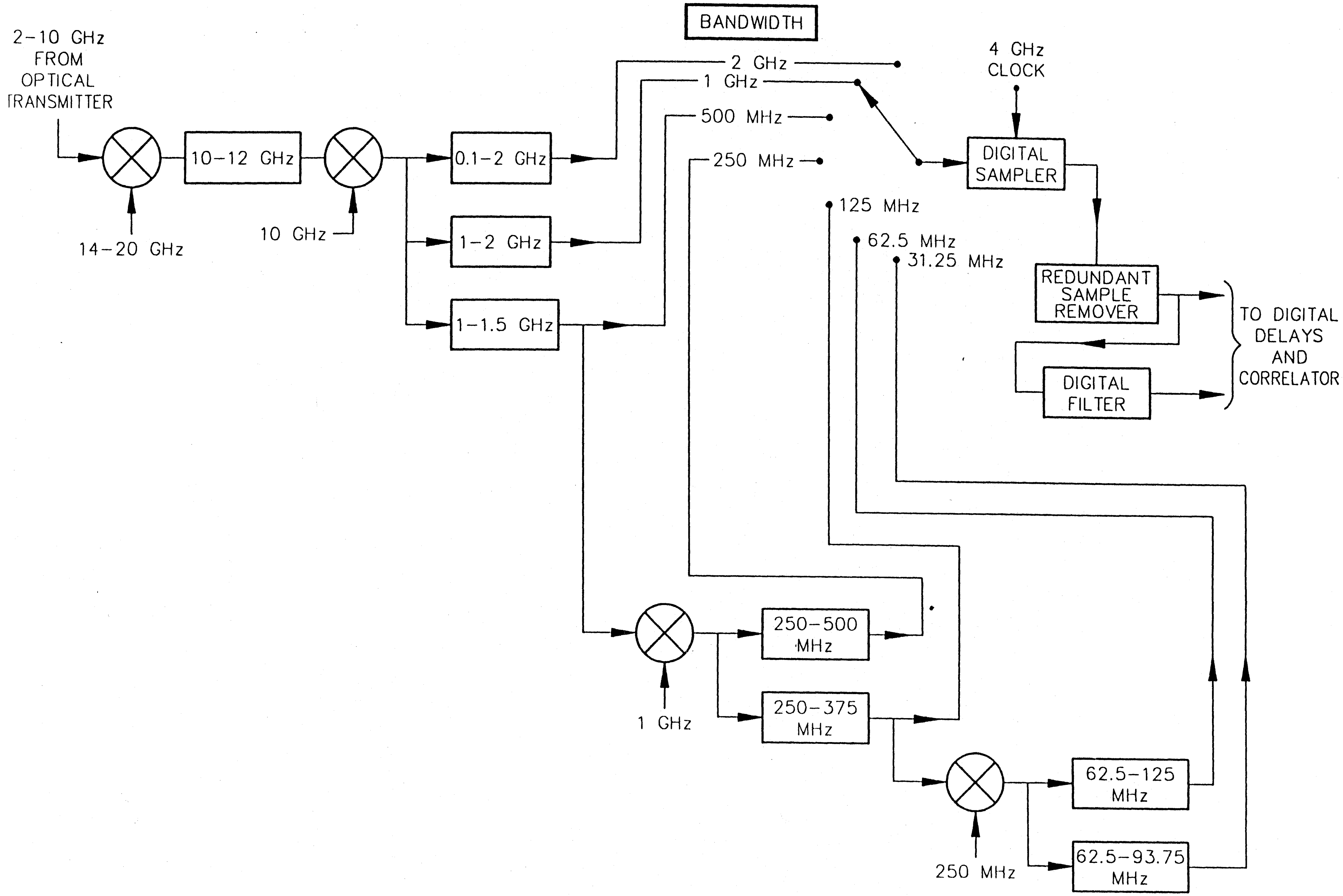


Figure 2. Filtering and digital sampling, for continuum observations, of the 8 GHz bandwidth signal from one fiber link.



igure 3. Filtering and digital sampling of one IF channel for spectral line observations. A minimum of four such units is required per antenna.

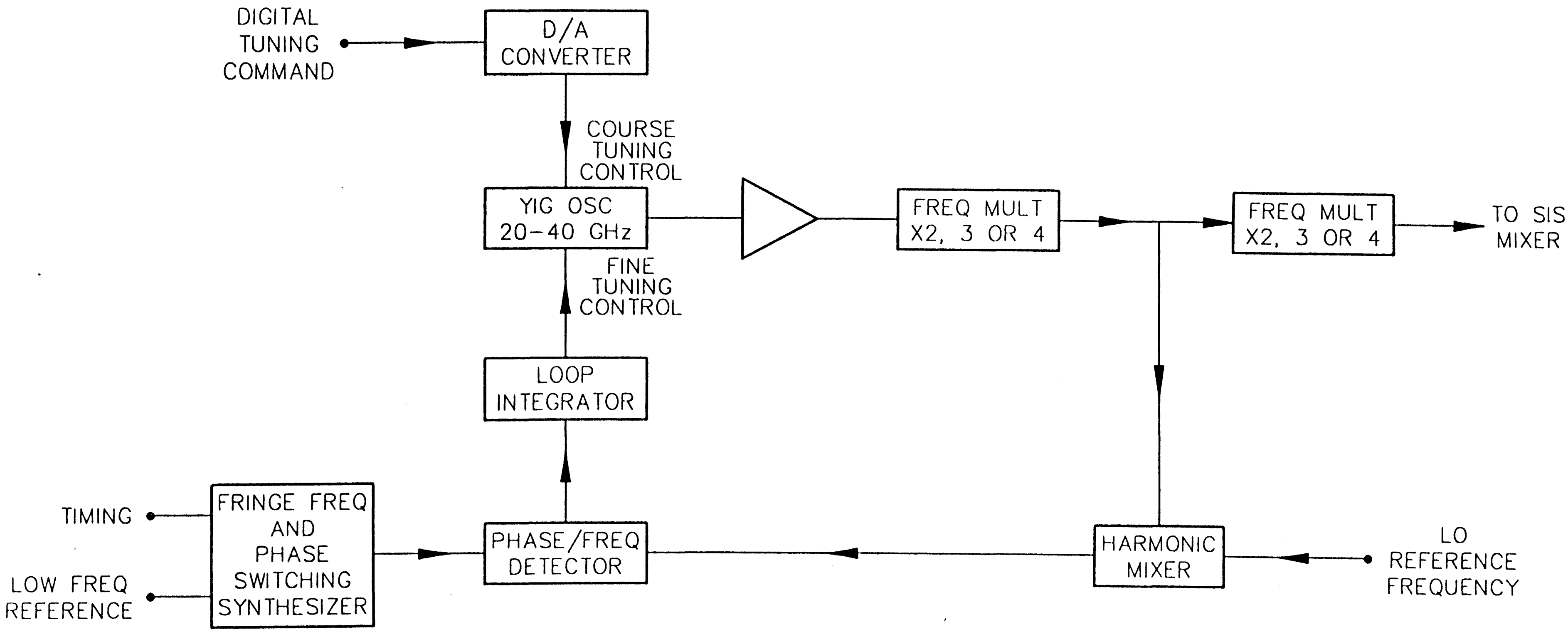


Figure 4. Phase-locked loop for generation of the first LO at an antenna, with fringe rotation and phase switching.

APPENDIX I

NATIONAL RADIO ASTRONOMY OBSERVATORY
Charlottesville, Virginia

May 5, 1995
revised: August 16, 1995

M E M O R A N D U M

To: D. Bagri, B. Clark, J. Carlstrom, L. D'Addario, D. Emerson, J. Romney, S. Padin, R. Sramek, D. Thornton, J. Welch

From: R. Escoffier

Subject: A Possible MMA Correlator Design

This memo describes the outline of a design for a MMA correlator. The lag design approach used here is not meant to be selected as the MMA standard but just to present a practical design to which future designs can be compared. The final decision on a MMA correlator architecture should be made later during the initial phases of an actual design project.

The correlator design is based on the MMA specifications listed below:

40 antennas
4 4-GHz samplers per antenna
3 KM maximum baseline
1024 lags per baseline at a 2 GHz bandwidth

The design below does not distinguish between continuum and spectral line observations. The samplers and correlators can be configured in factors of two from all samplers working at maximum bandwidths to a single sample per antenna working at the narrowest bandwidth.

A conservative design approach using 125 MHz interconnect technology is contemplated. Higher speed correlator chips and interconnect could be considered for cost effectiveness, but for now a well understood (essentially VLA correlator) technology is assumed. The design is based on a hypothetical correlator chip which is based, in turn, on the 1024-lag correlator chip used in the GBT spectrometer. It represents a conservative projection of what should be possible and affordable when the MMA is funded.

I. Block Diagram

A block diagram for the MMA correlator is seen in Figure 1. Four 4-GHz samplers are available for each antenna. This configuration is derived from the minimum system given in Dick Thompson's memo of April 18, 1995. A full

MMA CORRELATOR BLOCK DIAGRAM

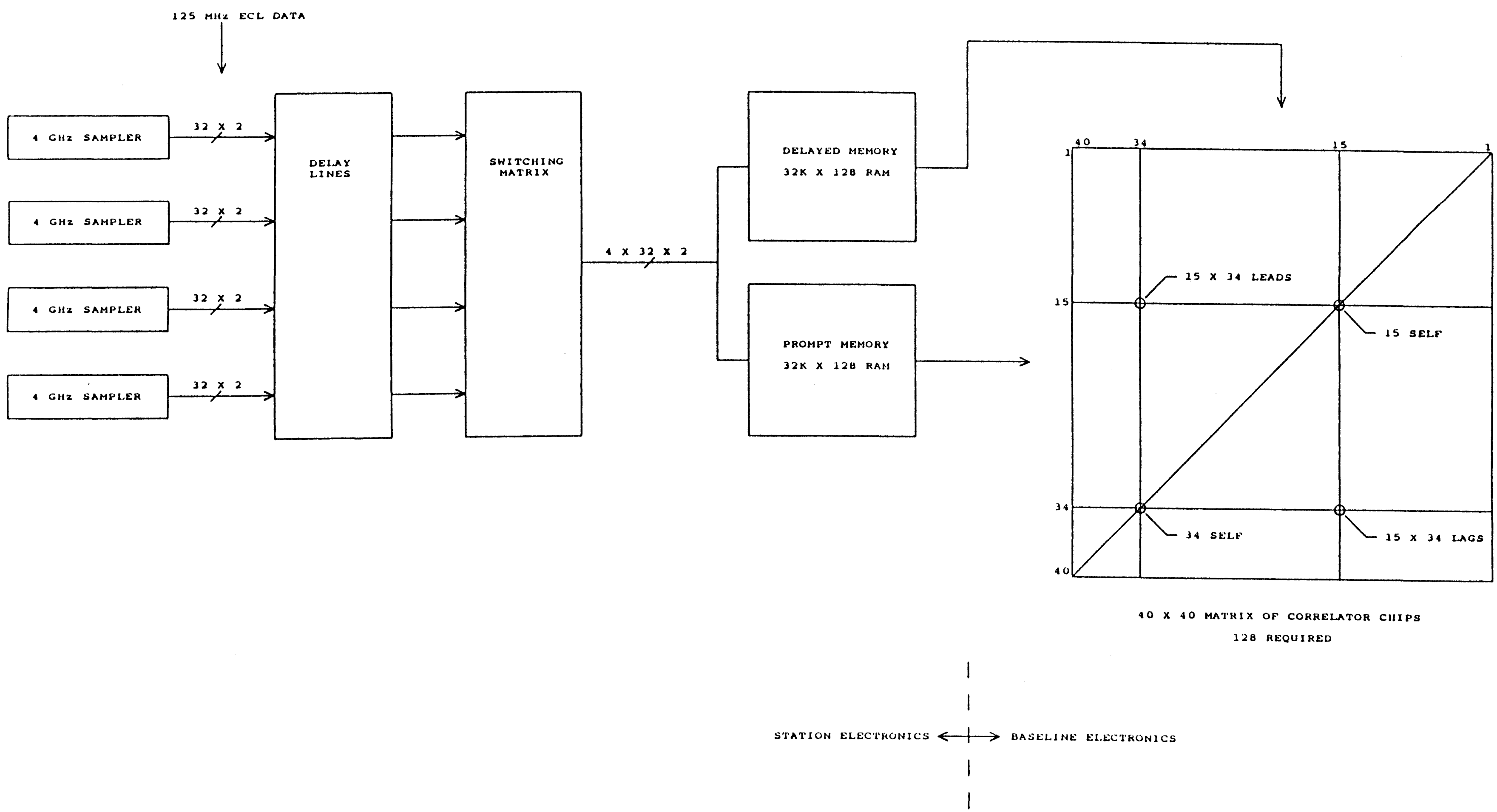


FIGURE 1

16-GHz bandwidth system, as described in this memo, would require eight 4-GHz samplers per antenna. Digital delay lines with up to 32-Ìsec of delay adjustment are provided for each sampler output.

A switching matrix is used to provide all of the mode versatility required. This includes reducing the number of active samplers from four to two or one per antenna and/or discarding samples to reduce the effective sample rate of a given sampler. As the aggregate data rate of the samplers goes down, the switching system will re-route the samples from active samplers to the correlator chips to optimize the number of lags for the observation being conducted.

The output of an active sampler is used to fill large RAMs in the memory system shown in Figure 1 so as to use the correlator chips more efficiently. The conventional technique used in correlators, where the correlator chips run at a lower clock rate than the samplers, uses a two-dimensional array of correlator chips to insure that every sample in the parallel output of one sampler is correlated with every sample in the parallel output of a second sampler. RAM memory is used here, however, to reduce the correlator requirements to a one-dimensional array of correlator chips.

The correlators seen in Figure 1 are formed into 40 by 40 arrays to correlate the outputs of the 40-station array. The 32-wide parallel outputs of the samplers (at 125 MHz) require an additional dimension to the correlator array as does the 4 parallel samplers. Thus, a total of 40 X 40 X 32 X 4 individual correlators (of some lag length) are required by this correlator.

Not shown in the block diagram of Figure 1 is the requirement for a long-term accumulator (LTA). The correlator chips themselves will provide short-term accumulation (from 1 to 16 msec). This relative long integration time provided by the correlator chips should allow the LTA to be made with high density (and inexpensive) dynamic RAMs. It is assumed that several integrations bins will be built into the LTA structure (for signal/reference/calibration, etc.).

II. Samplers

Four (or eight for a full 16-GHz system) 4-GHz samplers are available for each MMA antenna. By the time the MMA correlator is designed, it is assumed that there will be several approaches available for the design of this part of the correlator.

A phase lock loop can be used to phase shift the sample clock and adjust the exact sample time to a fraction of the sample period.

Either 3-level or 4-level samplers could be contemplated with the correlator chip and the sampler itself being the only part of the design significantly affected by this decision.

Integral to a 4-GHz sampler would be a 1-to-32 serial-to-parallel conversion stage allowing the sampler to use a 125 MHz output clock (actually, two such stages, one for each sampler bit, are required). The output of the sampler system for one antenna would hence be 4 X 32 X 2 signals with a 125 MHz clock rate. A given signal line from a sampler would carry a bit from every 32^{nd} sample.

III. Delay Lines

There will be 131,072 bits of RAM associated with the output of each 4-GHz sampler, yielding a delay range of 32 Ìsec (this is more than required, but is consistent with the size of fast RAMs). Since RAM addressing can only adjust the delay in steps of 32 samples, some additional logic will be required to obtain the final delay resolution of 1 bit.

The 131,072 bits would be provided in 16 1K X 8 RAMs (for each sampler bit). The entire delay requirement for one antenna will take 128 1K X 8 RAMs and associated logic and will probably fit on two identical PC cards of moderate size. The full 16-GHz system would have twice this number.

IV. Memory and Switching Matrix

The memory cards illustrated in the block diagram seen in Figure 1 will convert the 32-wide parallel sampler output (with each output carrying every 32^{nd} sample) into 32 parallel outputs of a different format. The samples from the 32 sampler outputs will be written into a large memory in time order and read from the RAM as 32 parallel outputs, each carrying a short burst of contiguous samples.

If the RAM is thought of as a circular buffer 1024 X 32 X 128 samples in circumference, each of the 32 (2-bit) outputs from a memory card would be assigned 1/32 of the total RAM. The 128-bit broadside input to the RAM buffer (obtained by splitting each of the 32 parallel sampler outputs into four parallel lines) are wired to store 128 consecutive samples into a given RAM address after a write pulse. Thus, the RAM buffer can be thought of as a linear time buffer containing 4,194,304 consecutive samples (originally taken at 4 GS/S) at any given time.

As stated above, each of the 32 outputs of the buffer is assigned 1/32 of the total buffer or 131,072 samples. At a clock rate of 125 MHz, the RAM can support a burst of contiguous samples from one output requiring about 1 msec to scan. In this 1 msec scan time, the entire RAM can be re-written at the 4 GHz sample rate. Hence, at the completion of each 131,072 sample scan, new samples are available for a subsequent scan. Thus, the correlator system will see short bursts of 131,072 contiguous samples originally taken at 4 GS/S but now slowed down to 125 MS/S from each of the 32 memory card outputs.

In this arrangement, all samples are used with the exception of 256 samples at the start of each burst which are required to fill the 256-bit lag generating shift register in the correlator before integration can begin. An additional small loss of sensitivity occurs since samples at the burst boundaries will not be correlated with samples in adjacent 1 msec segments.

For full versatility, two memories are required for each 4-GHz sampler. Each 40 antenna X 40 antenna correlator array has two dimensions and each axis requires one memory card per antenna (the two dimensions are driven by the prompt and delayed memory card outputs in Figure 1, the prompt signal represents the correlator input that drives all correlators in the 256-lag block and the delayed signal represents the signal that goes down the 256-bit shift register of the correlator).

When fewer than four samplers are being used, the switching matrix of Figure 1 can connect the remaining active sampler outputs to more than one memory card. (It would probably be possible to put multiplexor stages in the custom correlator chip to reduce the memory card requirement to only one per 4-GHz sampler.)

When a sample rate of less that 4 GHz is required, fewer that 32 inputs to the RAMs are required and the 32 outputs of the RAMs can be used to generate additional lags. Addressing in the "delayed memory" of Figure 1 can be offset to generate large lags allowing full digital versatility of the correlator with minimal number of switching stages.

For example, suppose only two samplers per antenna were active in a given observation. The correlator chips normally used by the two inactive samplers will be used to obtain twice the number of lags by having each active sampler drive its own memory cards plus an inactive sampler's memory cards. This possibility means that the correlator chips need not be cascaded together to produce more lags. As stated above, the delayed memory can, by offset RAM addressing, instantaneously generate the lag offset required by the higher lag correlator chips.

V. Correlators

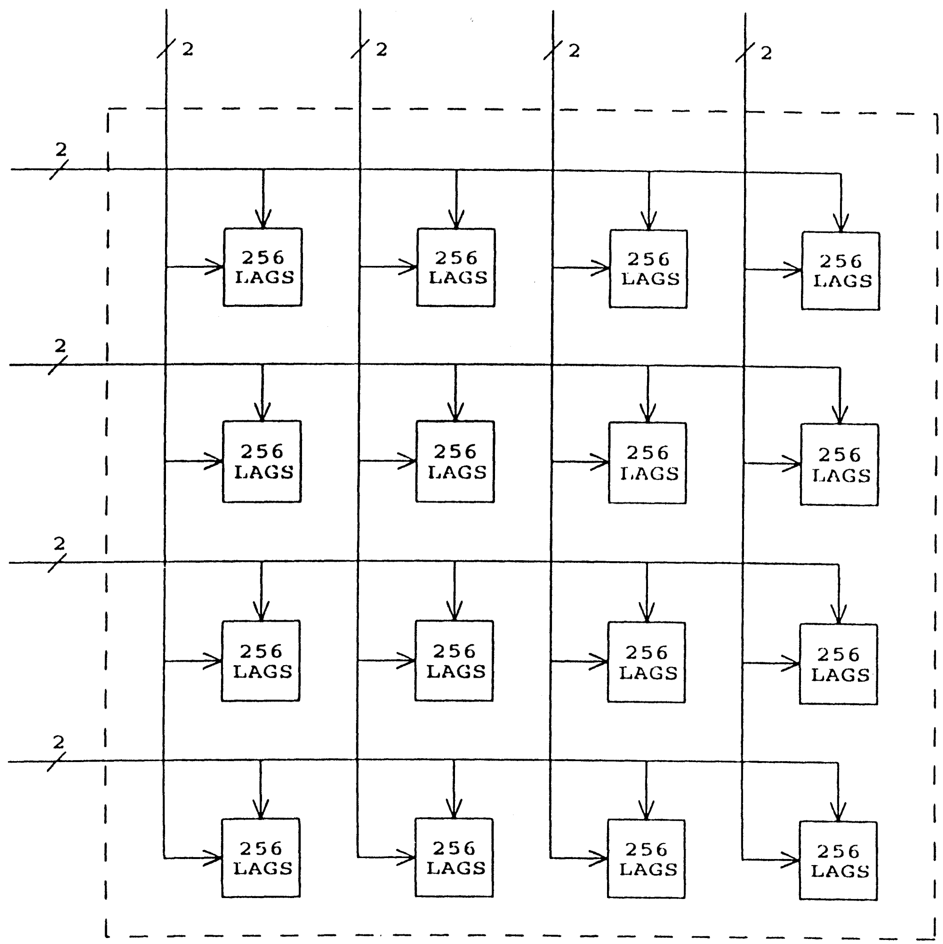
The correlator is designed around a proposed correlator chip. A block diagram of this chip is seen in Figure 2. The chip is proposed to be a 4 X 4 array of 256-lag correlators that operate at a clock rate of 125 MHz. A little bit of multiplexing on the chip would probably make the switching matrix easier to design, but this aspect of the design has not been pursued much at this point.

The ability to break the 256-lag correlators into two 128-lag correlators to support polarization observation will probably be necessary. Also, if the full 16-GHz system is selected, the correlator chip could be built as thirty-two 128-lag correlators. With this modification, the full bandwidth system could be built with the same number of correlators (but with twice the delay lines, memory cards and twice the 125-MHz cabling).

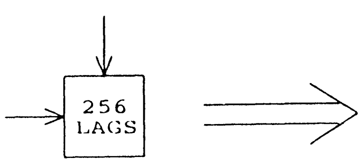
The total number of correlators required by this design is 40 X 40 X 32 X 4 or 204,800 256-lag correlators. By placing 16 such correlators on the chip in a small array, the number of chips required is reduced to a more practical number of 12,800 chips. Even with this number of chips, 200 to 400 correlator cards will be required for the MMA correlator.

The correlator chip seen in Figure 2 represents a factor of two increase in integration level from the 1024-lag correlator chip being used in the GBT spectrometer (assuming a 3-level by 3-level correlator). The GBT chips have 1024 3-level correlators, 1024 32-bit integrators and 1024 32-bit secondary

PROPOSED CORRELATOR CHIP



WHERE:



256-BIT LAG GENERATOR
256 3-LEVEL X 3-LEVEL MULTIPLIERS
256 12-BIT SHORT TERM INTEGRATORS
256 12-BIT SECONDARY STORAGE REGISTERS

FIGURE 2

storage registers for results readout. Cutting the short-term integrator to 12 or 16 bits, while increasing the total number of lags to 4096, results in an increase of the integration level of the chip by a factor of about two.

A higher speed correlator chip might be cost effective but would require a more expensive signal interconnect technology. One compromise might be to double the speed of the correlator chip but keep the data input rate at 125 MHz by putting 2-into-1 mux stages on the chips. This would halve the number of correlator chips required and would still allow use of a relatively easy interconnect technology.

VI. Long-Term Accumulation

A long-term accumulator design should be fairly straightforward. The one to several millisecond integration capacity of the correlator chips should allow high density and low cost DRAMs to be used here.

The LTA and the correlator switching networks can be designed for very rapid switching between modes. The fundamental memory cycle of 1 msec can be carried through to other parts of the system such that the system should have the ability to switch from full bandwidth continuum to spectral line, for example, many times a second. Additional integration/storage space can be put into the LTA to do essentially simultaneous wide band and narrow band observations.

VII. Performance

Straight factors of two trade-off between bandwidth and frequency resolution are made easy by the use of the memory cards. Because these cards can generate large lags by RAM addressing, the correlator arrays need not be interconnected. It might be advantageous to cascade the 256-lag correlator segments on the correlator chips with switching stages, but the correlator chips or matrices themselves need not be cascaded to increase the frequency resolution.

As the bandwidth is halved, the number of lags available for a given sampler doubles, and the frequency resolution improves by factors of four until the bandwidth (per sampler) goes below 62.5 MHz. After this point, factors of two improvement will occur unless recirculation is built into the correlator. The table below gives some of the performance parameters to be expected from this correlator design:

A) FOUR ACTIVE SAMPLERS PER ANTENNA (NO POLARIZATION CROSS PRODUCTS):

Total Bandwidth	Lags/IF	Frequency Resolution/IF
8 GHz	256	7.8125 MHz
4 GHz	512	1.953 MHz
2 GHz	1024	0.488 MHz
1 GHz	2048	122.070 KHz
500 MHz	4096	30.517 KHz
250 MHz	8192	7.629 KHz
125 MHz (oversampling)	8192	3.814 KHz
62.5 MHz (oversampling)	8192	1.907 KHz

B) FOUR ACTIVE SAMPLERS PER ANTENNA (WITH POLARIZATION CROSS PRODUCTS):

Total Bandwidth	Lags/Product	Frequency Resolution/IF
4 GHz	128	15.625 MHz
2 GHz	256	3.906 MHz
1 GHz	512	0.976 MHz
500 MHz	1024	244.140 KHz
250 MHz	2048	61.035 KHz
125 MHz	4096	15.258 KHz
62.5 MHz (oversampling)	4096	7.629 KHz
31.2 MHz (oversampling)	4096	3.814 KHz

C) TWO ACTIVE SAMPLERS PER ANTENNA (NO POLARIZATION CROSS PRODUCTS):

Total Bandwidth	Lags/IF	Frequency Resolution/IF
4 GHz	512	3.906 MHz
2 GHz	1024	0.976 MHz
1 GHz	2048	244.140 KHz
500 MHz	4096	61.035 KHz
250 MHz	8192	15.258 KHz
125 MHz	16384	3.814 KHz
62.5 MHz (oversampling)	16384	1.907 KHz
31.2 MHz (oversampling)	16384	0.953 KHz

D) TWO ACTIVE SAMPLERS PER ANTENNA (WITH POLARIZATION CROSS PRODUCTS):

Total Bandwidth	Lags/Product	Frequency Resolution/IF
2 GHz	256	7.8125 MHz
1 GHz	512	1.953 MHz
500 MHz	1024	0.488 MHz
250 MHz	2048	122.070 KHz
125 MHz	4096	30.517 KHz
62.5 MHz	4096	7.629 KHz
31.2 MHz (oversampling)	4096	3.814 KHz
15.6 MHz (oversampling)	4096	1.907 KHz

E) ONE ACTIVE SAMPLER PER ANTENNA:

Total Bandwidth	Lags/IF	Frequency Resolution/IF
2 GHz	1024	1.953 MHz
1 GHz	2048	0.488 MHz
500 MHz	4096	122.070 KHz
250 MHz	8192	30.517 KHz
125 MHz	16384	7.629 KHz
62.5 MHz	32768	1.907 KHz
31.2 MHz (oversampling)	32768	0.953 KHz
15.6 MHz (oversampling)	32768	0.476 KHz

In addition to the modes shown above, mixed modes (where one sampler samples a wide bandwidth and another sampler on the same antenna samples a narrow bandwidth) and subarrays will be easily accommodated by this design.

VIII. Estimated Size and Power Requirement

The (8-GHz bandwidth) system described above would require 160 4-GHz samplers and approximately 600 PC cards in the 6-U to 9-U EURO card size. This would require approximately four racks for the samplers and eight racks for the correlators. Power dissipation in the 100 to 200 KW range should be expected.

By far the most difficult design problem this correlator will present is in the signal cabling. One matrix of 40 X 40 correlators for a 4-GHz sampler will require 5120 125-MHz cables driving 51,200 loads. To this total a factor of 4 or 8 must be applied for the full 8- or 16-GHz system.

APPENDIX II

FX CORRELATOR FOR THE MMA: PRELIMINARY IDEAS

L. D'Addario
950905

1. INTRODUCTION

The FX correlator architecture has been implemented for several exisiting correlators because the number of multiplications and additions that must be performed per second is substantially lower than for an equivalent XF correlator, provided that the number of antennas and the number of spectral channels are large. This apparent advantage may be easily lost, however, because (1) the multiplications and additions are more difficult since they involve complex numbers of relatively high precision; and (2) the lack of symmetry produced by the separation of the machine into "F" and "X" sections reduces the efficiency of VLSI implementations. It is therefore difficult to generalize about which architecture is more cost-effective, so each application must be considered in some detail before the best choice for that case can be made. In addition to the question of how to achieve the lowest cost for fixed operating parameters, there is also a question of flexibility: synthesis telescope correlators (including the MMA's) are usually required to operate at a variety of bandwidths and spectral resolutions, and to support modes that allow trading of parameters, such as cross-polarization for spectral channels or baselines for bandwidth. Usually such flexibility can be more easily obtained with the XF architecture by virtue of its more regular structure.

A preliminary design for an XF architecture correlator has already been presented [1]. It seems worthwhile to explore the FX architecture for comparison, and this memo makes a start at doing so.

We take the approach of designing for the basic specifications rather than for any expansion thereof. That is, we will handle two baseband channels per antenna of 1 GHz bandwidth each; and we will achieve spectral resolution of 1 MHz at this bandwidth. This results in a correlator having the same performance as that of [1] when the latter is operated with the same input channelization (see Table C, line 2 of [1]). If an FX correlator with these specifications is substantially smaller (in chip count and other such measures) than the XF design, then it is worth pursuing further to see how tradeoffs, special modes, and expansions could be implemented; otherwise, further study is not justified.

2. DESIGN CONSIDERATIONS

The MMA requirements present some special difficulties. Unlike earlier FX correlators (especially the VLBA correlator), the input channel bandwidth cannot easily be matched to an achievable clock rate on a VLSI chip. We want to support channels at least 1 GHz wide, but practical clock rates are the order of 100 MHz. In principle this is no problem, since the necessary processing can be obtained from arrays of slower devices. The buffer memories needed to support this are in fact smaller than those needed for the XF design; only one DFT needs to be stored, whereas the XF needs to store one integration.

Another problem concerns wiring. The number of signals that must go from the antenna-based electronics to the baseline-based electronics is the same for both FX and XF (for the same total bandwidth), but for the FX case these require ~15 bits per sample to support the dynamic range at the output of the DFTs, whereas for the XF case they are still only 1 or 2 bits per sample. To keep the number of wires within reason, and also to limit the pin count on certain chips, we consider here the use of very high speed serial lines to connect the two major parts of the correlator.

Next there is the question of available VLSI technology for high speed signal processing. In the simplified, preliminary look taken here, it is assumed that all processing takes place at a clock rate of 125 MHz. This is the same as is assmumed for the XF chip [1], but here the operations that must be performed on each clock are sometimes more complicated. The performance measure of a chip is the product of its clock rate and the number of parallel operations per clock; the latter depends on the level of integration achievable, and will be one of the main uncertainties in estimating the cost effectiveness of the design.

3. ANTENNA BASED (FOURIER TRANSFORM) SECTION

See Figure 1. Each of the two channels is handled identically in separate hardware. Expansion to additional channels is straightforward. A switching matrix similar to that of the XF design could be added after the delay lines. (This might be useful to support alternative arrangements of the DFT processors, especially if additional digitizers were available. For further discussion, see section 6 below.)

To compute the real-time discrete Fourier transform (DFT) of a signal with 2 GHz sampling rate is quite a challenge. We will assume that pipelined DFT processors can be fabricated for operation at the 125 MHz clock rate, by which I mean that they accept 1 complex input value and produce 1 complex output value each clock cycle. Since the input is actually real, 2 samples are processed per clock (nominally the "real" and "imaginary" parts of the input value). Thus, 8 of these processors must operate in parallel to handle the full 2 Gsamp/s input.

To achieve the necessary throughput, each of the 8 DFT processors handles successive time segments of the input. For 1 MHz resolution, the DFT length must be 1024 (complex). The first 2048 samples are written to a buffer memory at full speed, and then are read into the first DFT processor 8 times more slowly. The next 2048 samples go to a second, similar buffer and then to the second processor. This continues until all 8 buffers have been written, and then begins again with the first buffer. On each 125 MHz clock, 32 b (16 samples) are written in parallel to the current buffer (the samples having been demuxed by 16 at the digitizers); and 4 b (2 samples) are read from every buffer into its own DFT processor. Further multiplexing/demultiplexing may be needed inside each buffer to achieve this speed.

At the output of each DFT, one complex result is produced per clock. For now, we take this to be encoded into a 16-bit word as (6,6,4) [i.e., 6 bits each for real and imaginary mantissas, and a common 4 bit exponent]. Each DFT has an output buffer into which these results are written; these 8 buffers mirror the input buffers, but they must be larger to accomodate the increased word length. On each clock, every DFT writes a 16 b result to its buffer, and a 128 b word (8 results) is read in parallel from the current buffer.

The length L=1024 DFT processors are the heart of this part of the correlator. I think that they can best be constructed using an architecture proposed by Canaris [2]. This does not use the Cooley-Tukey FFT, but rather is based on Goertzel filters; it has the strong advantage of not requiring storage of size L for the intermediate results between stages, and it has a symmetry that makes it well suited to VLSI implementation. We assume fixed-point complex arithmetic with 8-bit words (real and imaginary) internally, converted to (6,6,4) floating point at the output. The latter conversion does not reduce the word size, but some roundoff will be needed at the cross multiplier in order to prevent overflow in long integrations; conversion to floating point minimizes the significance of the roundoff loss. Since there are fewer DFTs than cross multipliers, it seems economical to do the conversion here.

The use of L=1024 complex transforms to compute the DFT of a length-2048 real sequence is not completely straightforward. It can be shown that the desired result is obtained only after some manipulation of the output; pairs of output points must be combined. The output buffer allows the data to be accessed in the required order, but another stage of computation is needed at the buffer output. This is neglected here, on the assumption that it can be accompished with a field-programmable gate array (2 or 4 per antenna) and will not require another custom VLSI chip.

The big question is how to arrange the DFT hardware onto chips. It is hoped that several of the L=1024 processors will fit onto one chip, perhaps as many as 8 of them. In that case, the pin count of the package becomes important, since it requires 32 input streams and 128 outputs. After adding power, clock, control, twiddle input, and test pins, the count might approach 200. This is marginally acceptable, but requires an expensive package. At this density, power dissipation per chip might also be a limitation.

At the output of this section, we now have 256 bit streams (counting both the A and B channels) of 125 Mb/s each that must go to the cross-correlation section; and we have similar data from each of the other 39 antennas. This makes for an unreasonably large number of cables, so I propose that we try to multiplex the output bits onto a smaller number of very-high-speed lines. If we could get up to 500 Mb/s, then only 64 wires per antenna would be needed (still a fairly large number for all 40 antennas). On the other hand, the electrical loss increases with frequency, and this may force the use of larger-dimension cables and connectors than would be needed at the lower speed; it thus could turn out that the large number of cables is more convenient after all. A careful study will be needed to find the best compromise.

At the moment I have very little idea of how much of this circuitry can fit on one chip. The 8x1024 DFT chip, if feasible, would have extraordinary performance: one L=1024 complex DFT every microsecond. This is ~100x faster than anything produced to date, as far as I know.

4. BASELINE BASED (CROSS CORRELATOR) SECTION.

Please see Figure 2. The high-speed multiplexing of the bitstreams from the antenna section still leaves 32 signals from each channel; each group of 4 signals carries the data from 1/8 of the spectral frequency channels from the DFTs (i.e, 128 of 1024 channels). We call each such group one "segment" of the spectrum.

An elementary multiplier-accumulator (MAC) accepts serial data from one frequency-channel segment of two antennas. The data appear sequentially by frequency channel so that all 128 channels are covered in just over 1 usec. The MAC first de-multiplexes each number and presents it in parallel to the complex multiplier. Each sample is on 4 wires at 500 Mb/s, so after demultiplexing we recover 16 parallel bits at 125 Mb/s in (6,6,4) format. The multiplier result is converted to (14,14,4) format and added to the partial sum read from the RAM; the new sum is then written back to the same RAM location. (A different number format or a longer RAM word might be needed to avoid loss of significance during an integration; this needs more careful study.) Two RAMs are provided to allow double buffering. At the end of an integration (approximately 1 msec, or 1k DFT spectra), the accumulator switches to the other RAM, allowing the first to be read out. A single-wire serial readout is sufficient at 4 Mb/s. During readout, the RAM contents are cleared so as to be ready for the next integration.

A total of 13,120 MACs is needed to cover the 780 baselines and 40 self correlations for 8 segments of 2 channels. These are arranged in half-matrix arrays of 820 MACs, as shown. Each array covers one segment of one channel, so 16 arrays are needed. Each output from the antenna electronics needs to go to only one array, minimizing the required wiring.

Again, a critical question is how many MACs can be placed on a VLSI chip? Each one requires 8 serial-to-parallel demultiplexers; one complex floating point (CFP) multiplier, one CFP adder, and 1024 bytes of fast RAM. The RAM must do a read and write in 8 nsec; if necessary, this can be accomplished by splitting it into slower RAMs that operate in parallel. The demultiplexers might better be placed on separate chips because of their higher clock speed, but then the pin count of the MAC chips becomes a more serious issue.

5. CHIP AND BOARD COUNTS

Depending on the level of integration achieved, a wide range of chip counts is possible. Tables 1 and 2 are an attempt to enumerate some possibilities. Note that many important things, including the samplers, delay lines, and long-term accumulator, are not included here.

TABLE 1: ANTENNA ELECTRONICS COUNTS

DFT1024s/chip	Total DFT chips	Cards	
8	80	10	4 ant/card, 8 chips/card
8	160	20	2 ant/card, 8 chips/card
4	240	20	2 ant/card, 12 chips/card
2	400	40	1 ant/card, 10 chips/card

TABLE 2: BASELINE ELECTRONICS COUNTS

MACs/chip	Chips/arry	Total MAC chips	Cards	
16	52	832	40	3 cards/array, 20 ch/cd
8	103	1,648	80	5 cards/array, 20 ch/cd
4	205	3,280	160	10 cards/array, 20 ch/cd
1	820	13,120	272	17 cards/array, 25 ch/cd

As is usual for an FX correlator with many antennas, the baseline section is likely to be dominant. Total VLSI chip count in these tables ranges from 912 to 13,520; this compares with 12,800 correlator chips in the XF design of [1]. It thus appears that under the most pessimistic assumptions on the FX, the costs of the two architectures may be about the same. But it seems possible that the FX may be cheaper by at least a factor of four. Even if accurate chip counts were available, this estimate is very rough; it does not account for development costs (including NRE for custom chips), nor does it calculate the costs of delay lines, buffer memories, long term accumualtor, and infrastructure. Remember that this assumes the same overall throughput (at 2 channels of 1 GHz each) and the same clock rate for both, but does not consider issues of flexibility or extremes of bandwidth or resolution (see next section).

6. RECONFIGURATION OPTIONS AND EXPANSION PATHS

6.1 More Baseband (Input) Channels

In the scheme of Figure 1, each antenna has available 16 DFT engines. These are arranged to handle 2 channels of 1.0 GHz bandwidth each. However, by re-arranging only the DFT buffers the same hardware could serve 1 channel of 2.0 GHz, 4 channels of 0.5 GHz, etc. For example, suppose we have 4 digitizers with sampling rates of 1.0 GHz. Then by splitting the input buffers so as to assign 4 DFTs to each digitizer, we can still keep up with the data flow (which is now using only 16 of the 32 input buffer signals per digitizer, or else is clocked at only 62.5 MHz instead of 125 MHz). The output buffers must be re-arranged slightly too; instead of reading 8 results each clock from one DFT, we read 4 results from the i-th DFT and 4 others from the (i+4)th DFT. In this way, the data from different channels always goes to different MACs for integration. Finally, each MAC must have twice as much RAM because it must handle twice as many spectral channels; and its readout rate or the minimum intergrating time must be doubled.

In all such cases, the total bandwidth remains 2.0 GHz. But we still have 1024 spectral channels for each input channel, so the frequency resolution improves in proportion to the number of input channels. For example:

	Input Channels	Bandwidth Each	Resolution	Mode [1]
	1	2.0 GHz	2 MHz	E.1
(Fig 1)->	2	1.0	1	C.2
	4	0.5	0.5	A.3
	8	0.25	0.25	

The last column gives the corresponding setup in the mode table of the XF correlator [1].

6.2 More Total Bandwidth

It is not possible to process more total bandwidth in this architecture without more hardware. However, it is straightforward to duplicate the system shown in Figures 1 and 2 as many times as desired, so that additional input channels of the same bandwidth are processed in parallel using independent hardware. Indeed, the two channels shown are handled entirely separately, and this could be done with any number of additional channels by replicating the circuitry.

When this is done, the spectral resolution per channel does not get worse; it stays the same. Unlike the XF architecture, there is no convenient way to trade resolution for bandwidth, since the resolution is fixed by the length of the DFTs. Whether the Goerzel filter DFT architecture [2] proposed here can be implemented in a variable-length form is something that should be investigated.

If the chip count for the 2 GHz version of this FX design is really only 1/4 that for the equivalent XF design, then the costs for 8 GHz total bandwidth would be comparable in the two architectures. With 1 GHz channels, the resolution of the FX would remain 1 MHz, while that of the XF would degrade to about 8 MHz.

7. CONCLUSIONS

I must emphasize that the chip count for this design, and hence its cost estimate, is not yet known. We have found only that there is a *possibility* that the chip count for certain specifications may be well below that of an XF correlator. Nevertheless, these very preliminary results are sufficient to conclude that the FX architecture cannot be dismissed, and that a more detailed study is warrented.

Considerable flexibility in input channelization is possible, and more spectral resolution is obtained as the number of input channels increases (for the same total bandwidth). Trading resolution for total bandwidth is more difficult than in an XF correlator (and is not possible in the rough design presented here); but if the cost saving is large enough, then the FX can be expanded to larger total bandwidths at a cost similar to that of the XF, and with no loss of spectral resolution.

ACKNOWLEDGEMENT

This memo benefited considerably from discussions with Ray Escoffier and from his comments on early drafts.

REFERENCES

[1] R. Escoffier, 1995, "A possible MMA correlator design." Memo to the MMA system design group, dated May 5, 1995.

[2] J. Canaris, 1993, "A VLSI architecture for the real time computation of discrete trigonometric transforms." J. of VLSI Signal Processing, vol 5, pp 95--104.

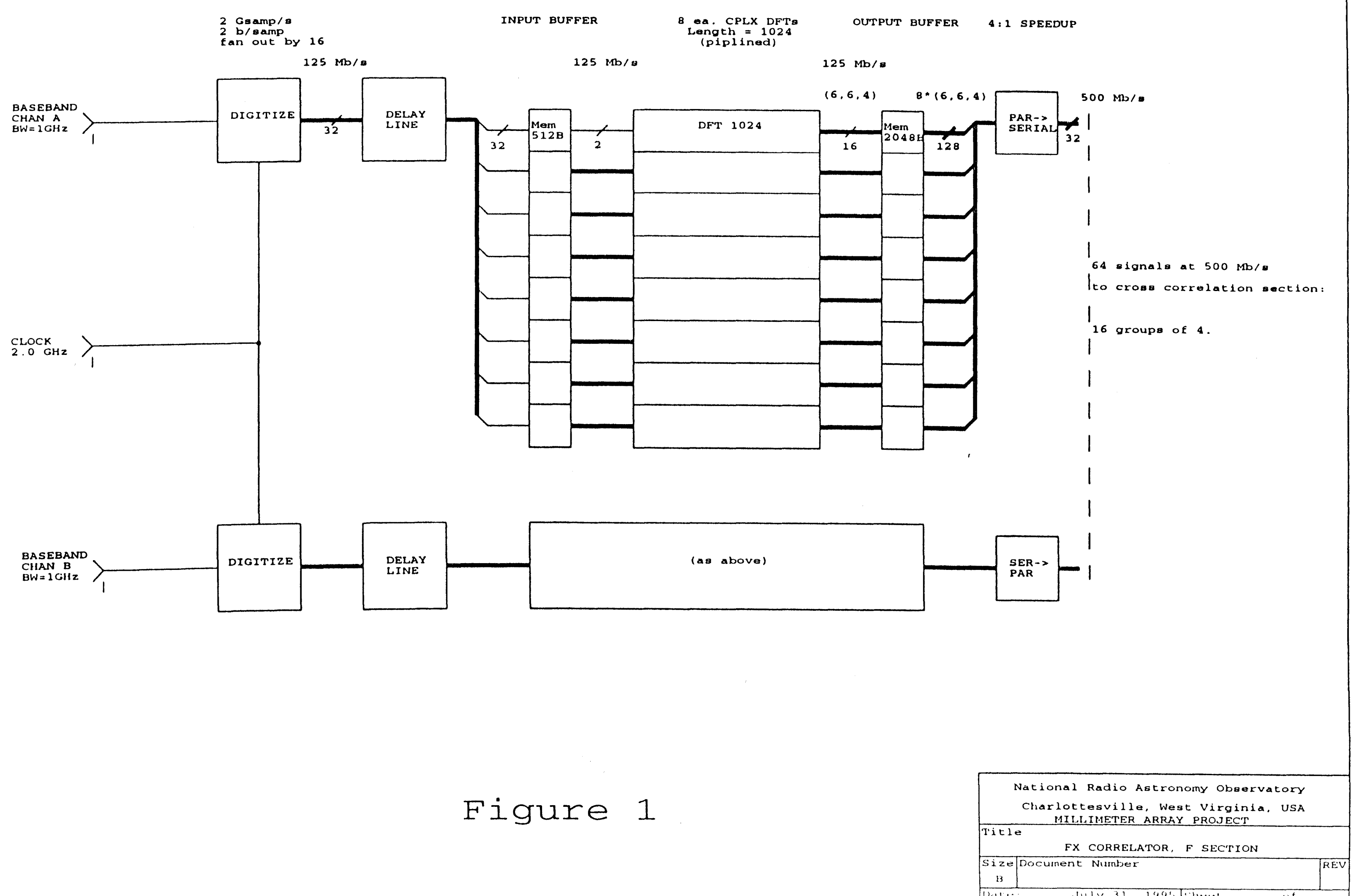
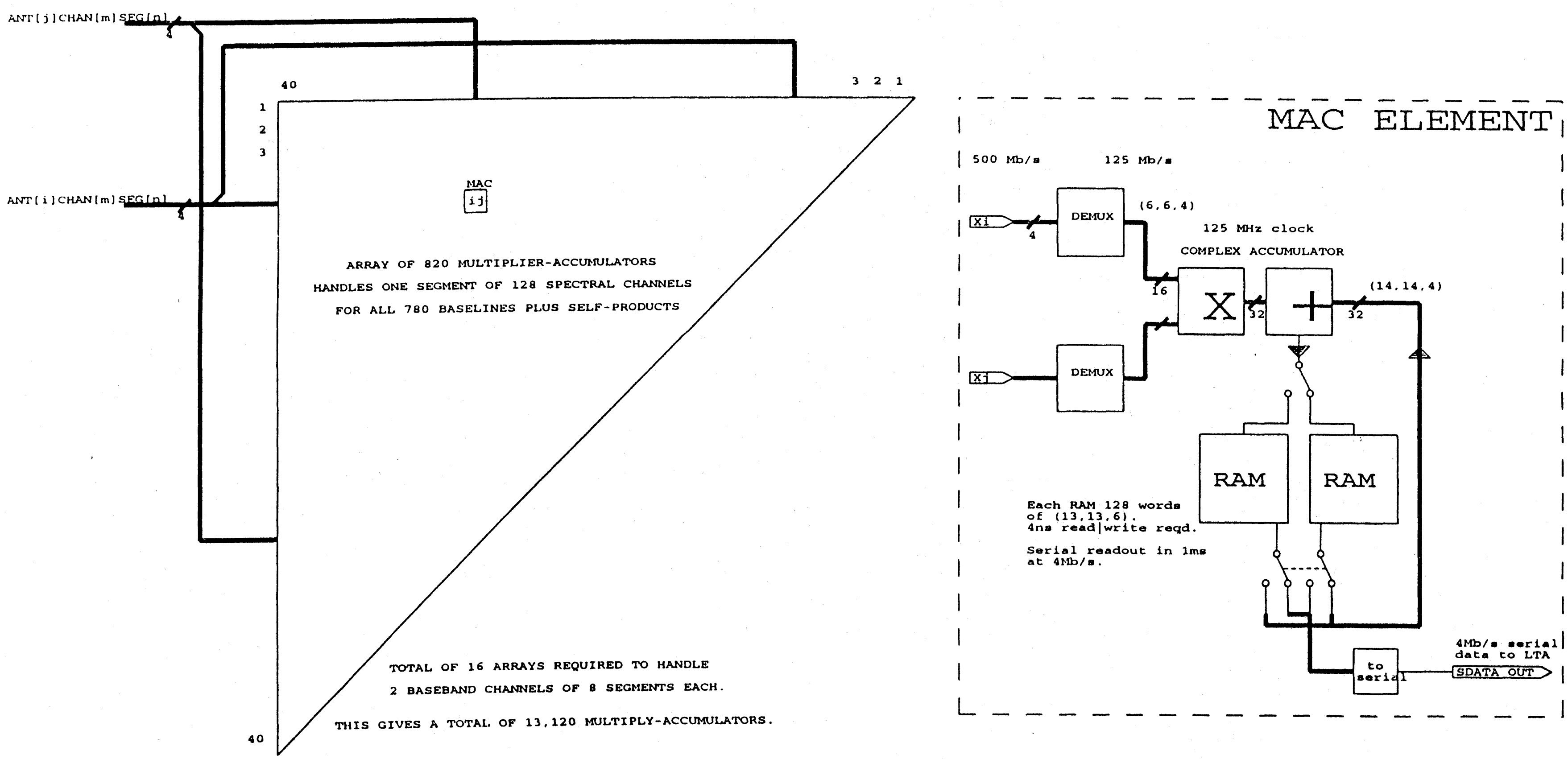

2 Gsamp/s
2 b/samp
fan out by 16
125 Mb/s
INPUT BUFFER
125 Mb/s
8 ea. CPLX DFTs
Length = 1024
(piplined)
OUTPUT BUFFER
125 Mb/s
4:1 SPEEDUP
BASEBAND
CHAN A
BW=1GHz
DIGITIZE
32
DELAY
LINE
32
Mem
512B
2
DFT 1024
(6,6,4)
16
Mem
2048B
8*(6,6,4)
128
PAR->
SERIAL
32
500 Mb/s
64 signals at 500 Mb/s
to cross correlation section:
16 groups of 4.
CLOCK
2.0 GHz
BASEBAND
CHAN B
BW=1GHz
DIGITIZE
DELAY
LINE
(as above)
SER->
PAR
National Radio Astronomy Observatory
Charlottesville, West Virginia, USA
MILLIMETER ARRAY PROJECT
Title
FX CORRELATOR, F SECTION
Size
B
Document Number
REV
Date: July 31, 1995 Sheet of

Figure 1



512 Mb/s INPUTS:

4x40 SIGNALS PER ARRAY
4x40x16 = 2,560 SIGNALS TOTAL

Figure 2

National Radio Astronomy Observatory Charlottesville, West Virginia, USA MILLIMETER ARRAY PROJECT	
Title	FX CORRELATOR, X SECTION
Size B	Document Number
Date: July 31, 1995	Sheet of