A Proposed WIDAR Correlator for the Expansion Very Large Array Project: Discussion of Capabilities, Implementation, and Signal Processing

NRC-EVLA Memo# 001

Brent Carlson, May 18, 2000

National Research Council of Canada,
Herzberg Institute of Astrophysics

ABSTRACT

A new architecture has been developed for the Expansion Very Large Array (EVLA) correlator project. This architecture is called WIDAR—an acronym for Wideband Interferometric Digital ARchitecture—and it uses novel digital filtering, anti-aliasing, and decorrelation methods to efficiently process wide bands with high spectral resolution and high spectral dynamic range. Additionally, the WIDAR technique permits fully digital $1/16^{th}$ sample geometric delay compensation. The correlator must correlate up to eight, 2 GHz basebands from each of up to 40 antennas. In the proposed system, 16384 spectral channels are available per baseline to be flexibly deployed across one or more 2 GHz basebands. In narrowband modes, up to ¼ *million* spectral channels can be deployed across one or more digitally filtered sub-bands. A full 4-bit system is envisioned that, along with decorrelation techniques, will enable greater than 10^5 spectral dynamic range even in the presence of powerful narrowband signals. This is important for the EVLA since it must, at some frequencies, operate in a severe interference environment. High spectral dynamic range and high spectral resolution allows for post-correlation interference excision where spectral channels containing interference are simply nulled out after correlation. This is a straightforward and effective method for dealing with interference in continuum observations. An upgrade path is available to allow the correlator to use adaptive interference cancellation techniques should such techniques eventually prove to be feasible. The architecture is well suited to providing fully digital phased-array capabilities where the VLA antennas will be phased together for high sensitivity VLBI. This document describes the correlator's capabilities and proposes an implementation of the correlator system for the EVLA. In the appendices it describes in some detail WIDAR signal processing and presents representative simulation results that clearly illustrate the performance capabilities of the correlator. This is a working document that is intended to stimulate further discussion and refinement.

Table of Contents

Table of Figures

1 Executive Summary

The Expansion Very Large Array (EVLA) project will significantly increase the bandwidth and number of antennas in the VLA. This will improve its sensitivity and dynamic range by at least an order of magnitude over the existing VLA. A significant part of the project is a new correlator that will have a large impact on the science output of the upgraded instrument. The new correlator must handle up to 16 GHz of bandwidth from each of up to 40 antennas and must operate in a severe interference environment. In an effort to meet these demanding requirements, a new correlator called **WIDAR** (Wideband Interferometric Digital ARchitecture) has been developed and studied in some detail. This correlator is designed to provide unprecedented signal processing capability to make maximum scientific use of the EVLA.

Basic Correlator Capabilities

The WIDAR correlator that will be described in this document will provide the following basic capabilities:

- In wideband modes, provide 16384 spectral channels across 16 GHz of total bandwidth (consisting of 8 x 2 GHz analog bands) on every baseline. Number of bands used and number of spectral channels per band is flexible.
- In narrowband modes (e.g. many programmable 8 MHz bands), provide up to 0.25 *million* spectral channels per baseline. Mixed wideband and narrowband modes are possible.
- Provide 1 Hz spectral resolution in narrowband radar mode.
- Simultaneous, fully digital, phased-VLA operation for VLBI.
- Spectral dynamic range $>10^5$ even in the presence of powerful narrowband signals.
- Each antenna is equipped with 2 GHz baseband systems and 4 GHz samplers. *All narrowband signals* are generated in the correlator with digital filters from these wide bands, eliminating the need for costly narrowband analog baseband systems.
- Multiple sub-array capability.

Interference Mitigation

A key requirement of the correlator is that it enable the EVLA to observe in severe RF interference environments. This interference is mostly narrowband in nature and can originate from terrestrial and satellite-based sources. The WIDAR correlator plans to use the following interference mitigation techniques.

- Provide high spectral resolution and spectral dynamic range $>10^5$ even in the presence of strong narrowband signals for post-correlation interference excision.
- Generate correlated output amplitudes that are immune to the modulating effects of time-variable interference.
- Provide real-time recognition of burst-like interference with subsequent correlator inhibit capability.
- Provide a reasonable upgrade path to allow installation of adaptive cancellation hardware should such techniques ultimately prove to be feasible.

Using digital FIR filters to "notch out" interference has also been considered, but it was found that excessively long filters are required to provide adequate performance. Post-correlation excision seems to provide better performance given the spectral resolution and spectral dynamic range provided by WIDAR.

WIDAR Signal Processing Overview

The WIDAR correlator relies heavily on digital filtering and techniques for decorrelation of aliased signals and sampler-generated narrowband interference harmonics. Digital filters can be quickly programmed (and thus reconfigured), have linear-phase characteristics, have calculable amplitude responses, and are immune to time-variable instabilities. Details and simulation results can be found in the appendices.

The correlator back-end is an XF[1] VLBI correlator and, combined with Local Oscillator techniques, can handle antenna arrays with baselines from zero to thousands of kilometers. The digital filters provide the front-end that allow the XF correlator back-end to be the most efficient possible—two correlator lags always produce one spectral output point. Since the dominant cost in a large XF correlator is the baseline correlator hardware (including the correlator chips), this efficiency ensures that a cost-effective solution is always obtained where high spectral resolution is required—even for very wide bandwidths and large antenna arrays.

Finally, a unique aspect of the WIDAR correlator design is that it can provide *fully digital* sub-sample ($<1/16^{th}$ sample) delay interpolation—eliminating the extra complication of circuitry that modifies the phase of the sampler clocks.

Cost and Power Requirements

A reasonably detailed cost estimate indicates that the 40-station correlator presented in this document can be constructed for $11 million U.S. This estimate includes some contingency but does not include the samplers, mains power system (UPS), or office space. A power estimate indicates that the correlator system will require about 100 kW from a power source with about a 135 kVA capacity.

[1] The XF architecture was chosen over 'FX' because it uses minimum bandwidth in data transmission from station-based to baseline-based hardware. This results in a significant savings in cost and complexity in a large, wideband correlator.

2 Simplified Correlator System Layout

A simplified layout of the correlator system[2] is shown in Figure 2-1.

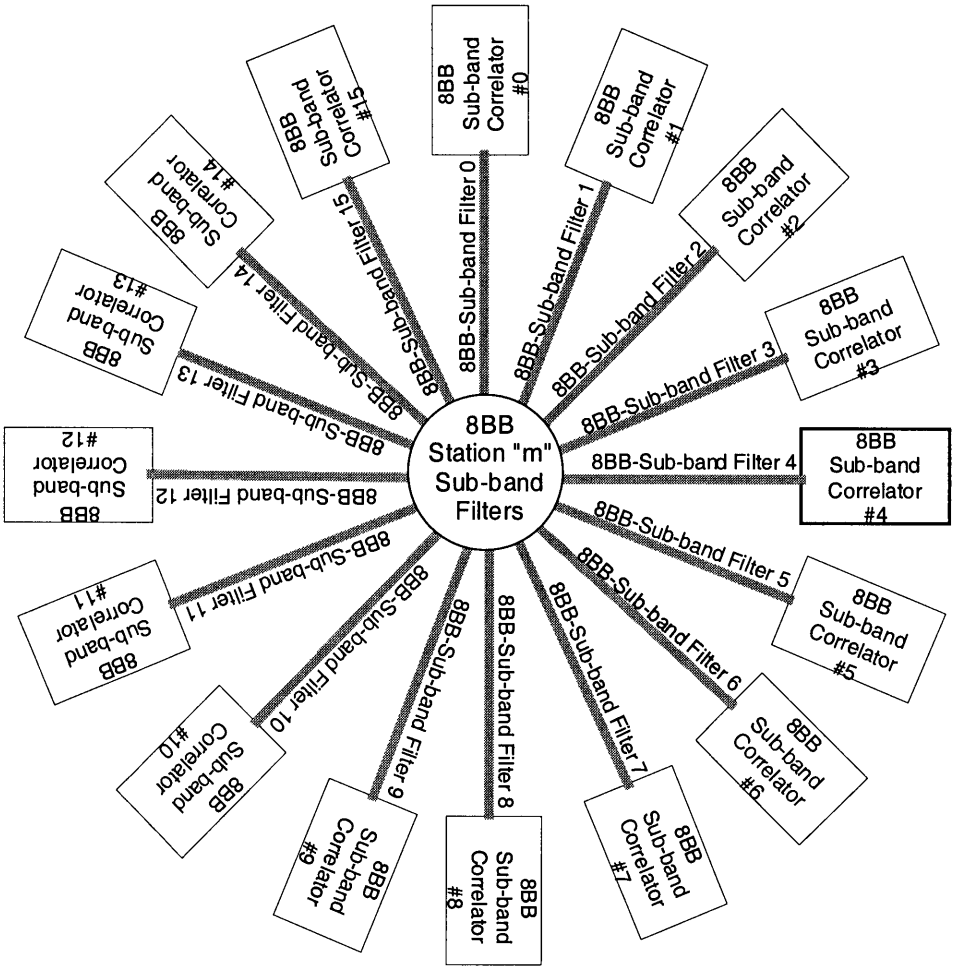


Figure 2-1 Simplified correlator system layout. The station electronics takes 8, 2 GHz basebands from each antenna and produces 16 sub-bands for each baseband. The output of each sub-band goes to an 8BB sub-band correlator where all of the correlations for that sub-band are performed.

In this layout, there are **M stations** (antennas) and each station has 8, **2 GHz basebands** (each baseband is defined as the output of a 4 Gs/sec sampler). These 8 basebands (8BB) can be arranged in polarization pairs up to 4 baseband pairs (4BBP). The electronics for each station in the correlator contains '**N' sub-band digital FIR filters** for each baseband. N is chosen to be 16 because resulting sub-band sample rates can be processed with existing and emerging digital technologies and there are not a significant number of sub-band boundaries. The output of a particular sub-band's filters *for all basebands* goes to an 8BB sub-band correlator. Thus, since N is 16, there are 16, 8BB sub-band

[2] For a discussion as to why this correlator layout was chosen as opposed to other possible layouts, please see Section 4.2.7.

correlators. Each sub-band correlator correlates all base*lines* for all base*bands* for a particular sub-band.

The primary purpose of the sub-band filters is to split the wide 2GHz band into narrower sub-bands operating at lower sample rates than can be handled by digital hardware. WIDAR uses techniques to seamlessly stitch together the individual sub-band correlations to reproduce the wide band spectrum.

In standard wideband modes, the correlator operates as shown in Figure 2-1 and as mentioned, seamlessly stitches the sub-bands together to yield the wideband spectrum. However, if less total bandwidth is required, then each sub-band correlator can be programmed to:

1. Correlate fewer basebands and increase the spectral resolution on the basebands that are being correlated.

2. Correlate fewer sub-bands. Increased spectral resolution is obtained by using each sub-band correlator to correlate one section of the overall lag chain on fewer sub-bands.

3. Decrease the sub-band bandwidth. Increased spectral resolution is obtained by lowering the sub-band sample rate while operating the correlator at the highest sample rate possible using **recirculation** memory. A squared increase (with decreasing sub-band bandwidth) in spectral resolution is obtained using this method.

4. A combination of one or more of the above. Each sub-band filter in each baseband is programmable in width and placement[3] within the 2 GHz baseband.

The following sections will elaborate on the basic layout just described. The very basic concept though, is that there are 16 sub-band correlators driven by sub-band filters from each station/antenna. Each sub-band correlator can be programmed to perform different functions depending on what bandwidth and spectral resolution are required.

[3] With a 'slot' and 'slot width' restriction explained in the next section.

3 Observing Modes

This section contains a description of correlator observing modes. The section starts out with a basic description of the capabilities of the correlator that observers should be cognizant of to make maximum use of the system. It then defines a nomenclature for describing observing modes and finally presents a number of tables of representative observing modes.

3.1 The Observer's View of the Correlator

3.1.1 Sub-arrays

The correlator produces the same correlation products on every baseline within a sub-array. There can be any number of sub-arrays—each sub-array can be thought of as a separate array with its own correlator much like a separate process, or window on a UNIX computer. Separate sub-arrays can have antennas that are in common, but all baselines within any sub-array must use the same observing mode or set of modes. It is acceptable for an antenna that is common to two sub-arrays to have (for example) two observing modes active—one mode applies to one sub-array and the other mode applies to the other sub-array. It is not possible to utilize correlator resources (spectral channels/correlator chips) assigned to two antennas that are (currently) part of mutually exclusive sub-arrays.

3.1.2 Antennas, Basebands, and Sub-bands

Each antenna can have up to 8, 2.048 GHz **basebands** (sampled at 4.096 Gsamples/sec[4] each). Each *baseband* can have its own delay/phase center on the sky and could have its own independently tunable Local Oscillator (if the actual antenna receivers are so equipped). The basebands can be arranged in up to 4 baseband (polarization) pairs.

Each baseband (and there are 8) of each antenna can be split with digital filters into up to 16 **sub-bands**. Each sub-band can be anywhere from $1/16^{th}$ of the baseband bandwidth to $1/256^{th}$ of the baseband bandwidth. In wideband modes, adjacent sub-bands can be seamlessly 'stitched' together to yield wider total bandwidths up to the baseband bandwidth of 2.048 GHz.

One additional 'special purpose' **radar mode** filter for *each* baseband can split the baseband into a sub-band with anywhere from $(1/16^{th})^2$ to $(1/256^{th})^2$ of the baseband bandwidth. Each special purpose filter—consisting of two cascaded filters—will normally be used only for the narrowband radar mode but could be used for other narrowband observing modes if desired. If this filter is active, then one of the normal sub-band filters for the baseband cannot be used.

[4] Chosen to be compatible with standard VLBI sample rates.

Each of the 16 sub-band filters (for each of 8 basebands) is individually programmable in bandwidth and placement within the baseband (within defined constraints). Each sub-band filter passband can only be within one of the integer 'slots' of the baseband that is the same as the bandwidth fraction of the filter. (Additionally, it is acceptable for more than one filter to occupy a given slot.) This is shown in Figure 3-1 for a sub-band bandwidth of 1/16 and 1/32 of the baseband. In the diagram there are 6 sub-band filters present—5 at 1/32 and 3 at 1/16—all of which could be within the same baseband. Slot numbering in the diagram is for 1/16 slots. The filter placements on the right of the figure are illegal.

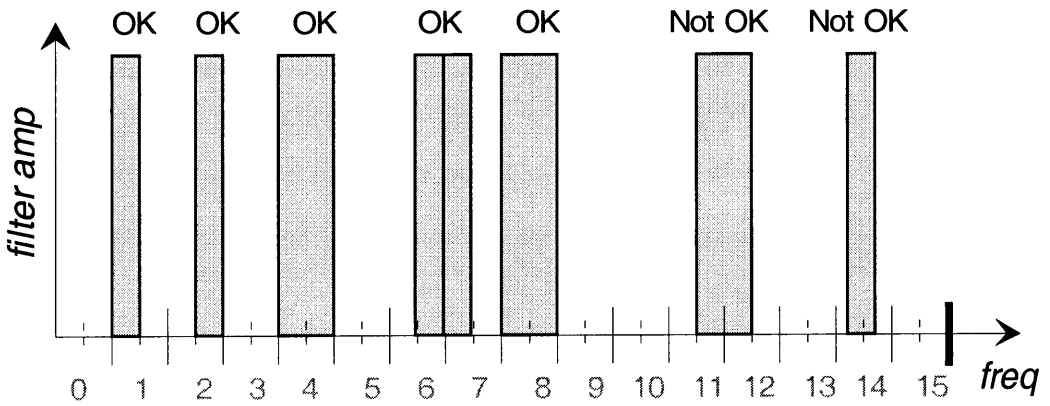


Figure 3-1 Example allowed and not allowed sub-band filter placements for 1/16 and 1/32 bandpass slots.

Example: If the filter is 1/16th of the baseband, then the passband of the filter must be within one of the 16 integer slots of the baseband.

Example: If the filter is 1/256th of the baseband, then the passband of the filter must be within one of the 256 integer slots of the baseband.

All correlated spectra up to and including the 'slot boundaries' of the filter passband are good data except that at the channel containing the slot boundary (a.k.a. the sub-band boundary) there is a $\sqrt{2}$ degradation in SNR (which quickly tapers off to negligible degradation within ~0.5 MHz of the edge of the sub-band in non-radar modes).

3.1.3 Baseline Correlations

The correlator effectively consists of ***16 sub-band correlators*** each of which processes *sub-bands* from up to 8 basebands (or up to 4 baseband pairs). Thus, each sub-band correlator could be viewed as a separately programmable correlator for each sub-array. For each baseline within *each* sub-band correlator there are effectively 2048 lags[5] (resulting in 1024 frequency points) that correlate the sub-band data from up to 8 of the basebands. If all polarization products are desired from all 4 baseband pairs, then 16 correlations must be performed—each one using 128 lags of each sub-band correlator. Additionally, in narrowband modes (e.g. 1/256th sub-band—8 MHz), there can be up to

[5] Which is actually within one correlator chip.

32768 lags per baseline in *each* sub-band correlator when '**recirculation**' is active. *The lags in each sub-band correlator can be used independently or in combination with lags from other sub-band correlators to produce an increasing number of spectral points.*

For one baseline within one sub-band correlator, each of the 8 basebands can be a different bandwidth (and therefore a different bit rate) and in fact in a different frequency slot in each baseband. Each of these single baseline sub-band correlators can perform 'recirculation' on only 4 basebands (or 2 baseband pairs)—a memory width limitation. It is illegal to mix recirculation and normal processing and it is illegal to mix different recirculation modes in the same sub-band correlator on the same baseline. Thus, if recirculation is active on one of these sub-band baseline correlators, then non-recirculation modes (or different recirculation modes) cannot be used at the same time on the same sub-band correlator. In this case, it is necessary to use a different sub-band correlator for the non-recirculation or different recirculation mode.

3.2 Observing Mode Nomenclature

The **Observing Mode** in the following tables indicates the number of basebands, number of polarizations, number of sub-bands, and the sub-band bandwidths that are used for that mode. The number of spectral channels that can be used for the particular mode is very flexible and is thus not part of the mode designation. The Observing Mode nomenclature is shown in Figure 3-2:

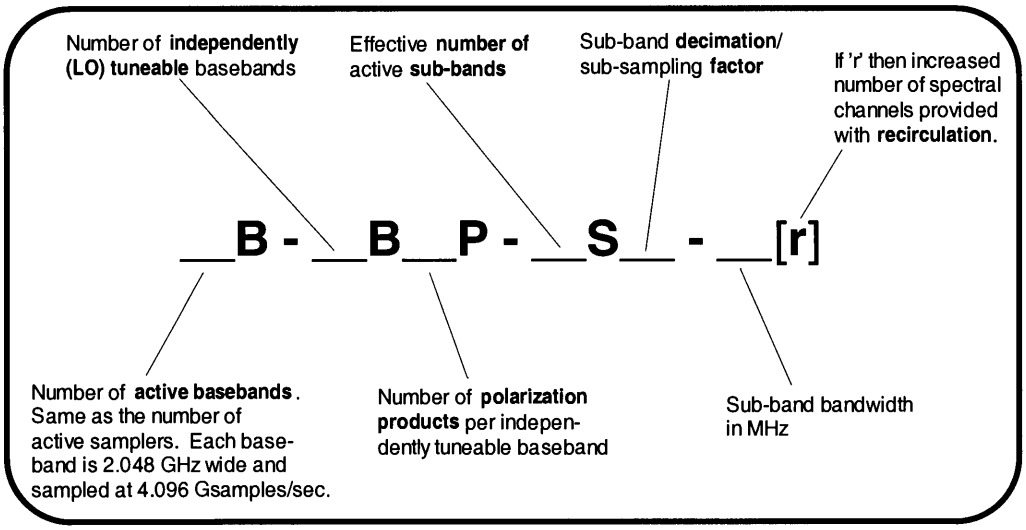


Figure 3-2 Observing Mode Nomenclature

Example: 8B-4B2P-16S16-128

In this example, there are 8 active basebands (**8B**), arranged as 4 baseband pairs (independently tuneable basebands), with 2 polarization products per baseband pair (**4B2P**). 16 sub-bands are used with a decimation factor of 16 (**16S16**). Each sub-band has a bandwidth of 128 MHz (**128**). This is the mode in the first row of Table 3-1.

Example: 1B-1B1P-1S256-8r

In this example there is 1 active baseband (**1B**), arranged as 1 baseband with an independently tuneable LO, with 1 polarization product (**1B1P**). 1 sub-band is active with decimation factor of 256 (**1S256**) and a sub-band bandwidth of 8 MHz (**8**). The '**r**' designator indicates that recirculation is active in the correlator to produce a maximum number of spectral channels. This is the mode in the first row of Table 3-2.

Example: 1B-1B1P-1S(64x128)-0.25r

In this example, there is 1 active baseband (**1B**), arranged as 1 baseband with an indepedently tuneable LO, with 1 polarization product (**1B1P**). **Narrowband radar mode** is active on one sub-band (**1S(64x128)**) where two FIR filters are cascaded, the first with a 1/64 bandpass and the second with a 1/128 bandpass. The sub-band bandwidth is 0.25 MHz and recirculation is active (**0.25r**). This is the mode in the second row of Table 3-6.

Example: 2B-1B2P-8S16-128

In this example, there are 2 active basebands (**2B**), arranged as 1 baseband pair with an independently tuneable LO, with 2 polarization products (**1B2P**). Only 8, 128 MHz sub-bands per baseband are used with a decimation of 16 (**8S16-128**). This illustrates what might be specified when observing at L-band for 1 GHz of contiguous bandwidth (although there is nothing in the mode that forces the sub-bands to be adjacent). This is the mode in the 12th row of Table 3-1.

3.3 Representative Observing Mode Tables

These tables define representative observing modes that the proposed correlator can provide. Each entry in the table uses the entire correlator capacity but what is not shown is that there is some flexibility in assignment or arrangement of the number of spectral channels allocated to a given cross-power result. The tables are not an exhaustive definition of what can be done with the correlator since many different modes can be combined to provide virtually an unlimited number ways of configuring the correlator.

A convenient way to use the tables is as follows:

Step 1: Determine the desired bandwidth per final cross-power result that one is interested in. This can be found in the tables by looking at column four—labeled "**BW per cross-power spectrum**".

Step 2: Decide what the desired spectral resolution is or how many frequency channels are required across the bandwidth of interest. This is found in columns six ("**Spectrl resolution**") and five ("**Max Spectral Channels per cross-power spectrum per bsln**") respectively.

Step 3: Looking to the left across the row, the maximum number of cross-power results that can be obtained *per baseline* is in the second column ("**# of cross-power**

spctra per bsln"). Find the number of cross-power results that are needed. Where entries from Steps 1, 2, and 3 in different rows are the same, it is necessary to look at the **Mode** to determine the desired number of baseband pairs and polarization products that are required.

Example: Step 1: We want a bandwidth per cross-power result of 2.048 GHz—look in column 4 to find the appropriate table. **Step 2:** We want 8192 frequency channels (spectral resolution of 250 kHz) per cross-power result—look in column 5 (or 6). **Step 3:** We want 2 cross-power results from 2 polarization products from 1 baseband pair—look in colums 2 and 1. Looking at Table 3-1, this is mode **2B-1B2P-16S16-128**

Mode	# of cross-power spctra per bsln	Total BW (GHz)	BW per cross-power spectrum (GHz)	Max Spectral Channels per cross-power spectrum per bsln	Spectrl res-olution (kHz)	Decima-tion Factor	Sub-band bandwidth (MHz)	Spectral channels per sub-band spectrum
8B-4B2P-16S16-128	8	16.384	2.048	2048	1000	16	128	128
8B-4B4P-16S16-128	16	16.384	2.048	1024	2000	16	128	64
8B-8B1P-16S16-128	8	16.384	2.048	2048	1000	16	128	128
4B-2B2P-16S16-128	4	8.192	2.048	4096	500	16	128	256
4B-2B4P-16S16-128	8	8.192	2.048	2048	1000	16	128	128
4B-4B1P-16S16-128	4	8.192	2.048	4096	500	16	128	256
2B-1B2P-16S16-128	2	4.096	2.048	8192	250	16	128	512
2B-1B4P-16S16-128	4	4.096	2.048	4096	500	16	128	256
2B-2B1P-16S16-128	2	4.096	2.048	8192	250	16	128	512
1B-1B1P-16S16-128	1	2.048	2.048	16384	125	16	128	1024
2B-2B1P-8S16-128	2	2.048	1.024	8192	125	16	128	1024
2B-1B2P-8S16-128	2	2.048	1.024	8192	125	16	128	1024
2B-1B4P-8S16-128	4	2.048	1.024	4096	250	16	128	512

Table 3-1 Representative Wide-band Observing Modes

Mode	# of cross-power spctra per bsln	Total BW (MHz)	BW per cross-power spectrum (MHz)	Max Spectral Channels per cross-power spectrum per bsln	Spectrl res-olution (kHz)	Decim Factor	Sub-band bandwidth (MHz)	Spectral channels per sub-band spectrum
1B-1B1P-1S256-8r	1	8	8	262144	0.035	256	8	262144
1B-1B1P-2S256-8r	2	16	8	131072	0.061	256	8	131072
1B-1B1P-4S256-8r	4	32	8	65536	0.122	256	8	65536
1B-1B1P-8S256-8r	8	64	8	32768	0.244	256	8	32768
1B-1B1P-16S256-8r	16	128	8	16384	0.488	256	8	16384
2B-2B1P-1S256-8r	2	16	8	131072	0.061	256	8	131072
2B-1B2P-1S256-8r	2	16	8	131072	0.061	256	8	131072
2B-1B4P-1S256-8r	4	32	8	65536	0.122	256	8	65536
2B-2B1P-2S256-8r	4	32	8	65536	0.122	256	8	65536
2B-1B2P-2S256-8r	4	32	8	65536	0.122	256	8	65536
2B-1B4P-2S256-8r	8	64	8	32768	0.244	256	8	32768
2B-2B1P-4S256-8r	8	64	8	32768	0.244	256	8	32768
2B-1B2P-4S256-8r	8	64	8	32768	0.244	256	8	32768
2B-1B4P-4S256-8r	16	128	8	16384	0.488	256	8	16384
2B-2B1P-8S256-8r	16	128	8	16384	0.488	256	8	16384
2B-1B2P-8S256-8r	16	128	8	16384	0.488	256	8	16384
2B-1B4P-8S256-8r	32	256	8	8192	0.977	256	8	8192
2B-2B1P-16S256-8r	32	256	8	8192	0.977	256	8	8192
2B-1B2P-16S256-8r	32	256	8	8192	0.977	256	8	8192
2B-1B4P-16S256-8r	64	512	8	4096	1.953	256	8	4096

Table 3-2 Representative 8 MHz Narrow-band Observing Modes

Mode	# of cross-power spctra per bsln	Total BW (MHz)	BW per cross-power spectrum (MHz)	Max Spectral Channels per cross-power spectrum per bsln	Spectrl res-olution (kHz)	Decim Factor	Sub-band bandwidth (MHz)	Spectral channels per sub-band spectrum
4B-4B1P-1S256-8r	4	32	8	65536	0.122	256	8	65536
4B-2B2P-1S256-8r	4	32	8	65536	0.122	256	8	65536
4B-2B4P-1S256-8r	8	64	8	32768	0.244	256	8	32768
4B-4B1P-2S256-8r	8	64	8	32768	0.244	256	8	32768
4B-2B2P-2S256-8r	8	64	8	32768	0.244	256	8	32768
4B-2B4P-2S256-8r	16	128	8	16384	0.488	256	8	16384
4B-4B1P-4S256-8r	16	128	8	16384	0.488	256	8	16384
4B-2B2P-4S256-8r	16	128	8	16384	0.488	256	8	16384
4B-2B4P-4S256-8r	32	256	8	8192	0.977	256	8	8192
6B-xxxP-xS256-8r	**not* avail.**	**not avail.**	**not avail.**	**not avail.**	**not avail.**			
8B-xxxP-xS256-8r	**not avail.**	**not avail.**	**not avail.**	**not avail.**	**not avail.**			
8B-8B1P-1S256-8	8	64	8	2048	3.91	256	8	2048
8B-4B2P-1S256-8	8	64	8	2048	3.91	256	8	2048
8B-4B4P-1S256-8	16	128	8	1024	7.81	256	8	1024
8B-8B1P-2S256-8	16	128	8	1024	7.81	256	8	1024
8B-4B2P-2S256-8	16	128	8	1024	7.81	256	8	1024
8B-4B4P-2S256-8	32	256	8	512	15.6	256	8	512

Table 3-3 Representative 8 MHz Narrow-band Observing Modes (Cont'd)

* A maximum of 4 basebands (2 baseband pairs) can be used in recirculation mode (**r**).

Mode	# of cross-power spctra per bsln	Total BW (MHz)	BW per cross-power spectrum (MHz)	Max Spectral Channels per cross-power spectrum per bsln	Spectrl res-olution (kHz)	Decim Factor	Sub-band bandwidth (MHz)	Spectral channels per sub-band spectrum
1B-1B1P-1S128-16r	1	16	16	131072	0.122	128	16	131072
1B-1B1P-2S128-16r	2	32	16	65536	0.244	128	16	65536
1B-1B1P-4S128-16r	4	64	16	32768	0.488	128	16	32768
1B-1B1P-8S128-16r	8	128	16	16384	0.977	128	16	16384
1B-1B1P-16S128-16r	16	256	16	8192	1.953	128	16	8192
2B-2B1P-1S128-16r	2	32	16	65536	0.244	128	16	65536
2B-1B2P-1S128-16r	2	32	16	65536	0.244	128	16	65536
2B-1B4P-1S128-16r	4	64	16	32768	0.488	128	16	32768
1B-1B1P-1S128-16	1	16	16	16384	0.977	128	16	16384
8B-4B4P-1S128-16	16	256	16	1024	15.6	128	16	1024
1B-1B1P-1S64-32r	1	32	32	65536	0.488	64	32	65536
1B-1B1P-2S64-32r	2	64	32	32768	0.977	64	32	32768
1B-1B1P-4S64-32r	4	128	32	16384	1.953	64	32	16384
1B-1B1P-8S64-32r	8	256	32	8192	3.91	64	32	8192
1B-1B1P-16S64-32r	16	512	32	4096	7.81	64	32	4096
2B-2B1P-1S64-32r	2	64	32	32768	0.977	64	32	32768
2B-1B4P-16S64-32r	64	2048	32	1024	31.2	64	32	1024
8B-4B4P-1S64-32	16	512	32	1024	31.2	64	32	1024
1B-1B1P-1S32-64r	1	64	64	32768	1.95	32	64	32768
4B-2B4P-1S32-64r	8	512	64	4096	15.6	32	64	4096
4B-2B4P-16S32-64r	128	8192	64	256	250	32	64	256
8B-4B4P-1S32-64	16	1024	64	1024	62.5	32	64	1024

Table 3-4 Representative 16, 32, and 64 MHz Narrow-band Observing Modes

Modes	# of cross-power spctra per bsln	Total BW (MHz)	BW per cross-power spectrum (MHz)	Example Max Spectral Channels per cross-power spectrum per bsln	Spectrl res-olution (kHz)	Decim Factor	Sub-band bandwidth (MHz)	Spectral channels per sub-band spectrum
4B-4B1P-8S16-128	4	4096	1024	2048	500	16	128	256
+								
4B-4B1P-8S256-8r	32	256	8	4096	1.95	256	8	4096
4B-2B2P-8S16-128	4	4096	1024	2048	500	16	128	256
+								
4B-2B2P-8S256-8r	32	256	8	4096	1.95	256	8	4096
2B-1B2P-8S16-128	2	2048	1024	4096	250	16	128	512
+								
2B-1B2P-8S256-8r	16	128	8	8192	0.977	256	8	8192
2B-1B2P-8S16-128	2	2048	1024	4096	250	16	128	512
+								
2B-1B2P-1S256-8r	2	16	8	65536	0.122	256	8	65536
2B-2B1P-8S128-16r	16	256	16	4096	3.91	128	16	4096
+								
2B-2B1P-2S256-8r	4	32	8	32768	0.244	256	8	32768
2B-2B1P-8S128-16r	16	256	16	4096	3.91	128	16	4096
+								
1B-1B1P-1S256-8r	1	8	8	131072	0.061	256	8	131072
1B-1B1P-1S16-128	1	128	128	1024	125	16	128	1024
+								
4B-4B1P-15S256-8r	60	480	8	4096	1.95	256	8	4096
1B-1B1P-1S16-128	1	128	128	1024	125	16	128	1024
+								
2B-2B1P-7S256-8r	14	112	8	8192	0.977	256	8	8192
+								
1B-1B1P-1S256-8r	1	8	8	131072	0.061	256	8	131072

Table 3-5 Representative Mixed-bandwidth Observing Modes

Mode	# of cross-power spctra per bsln	Total BW (kHz)	BW per cross-power spectrum (kHz)	Max Spectral Channels per cross-power spectrum per bsln	Spectrl res-olution (Hz)	Decim Factor	Sub-band bandwidth (kHz)	Spectral channels per sub-band spectrum
1B-1B1P-1S(64x64)-0.5r	1	500	500	262144*	1.9	4096	500	1048576
1B-1B1P-1S(64x128)-0.25r	1	250	250	262144	0.95	8192	250	1048576
1B-1B1P-1S(256x256)-0.03125r	1	31.25	31.25	262144	0.12	65536	31.25	1048576
1B-1B1P-1S(256x256)-0.03125	1	31.25	31.25	16384	1.9	65536	31.25	16384
1B-1B1P-2S(32x64)-1r	2	2000	1000	262144	3.8	2048	1000	262144
2B-1B2P-1S(32x32)-2r	2	4000	2000	262144	7.6	1024	2000	262144
2B-1B2P-1S(16x32)-4r	2	8000	4000	262144	15.3	512	4000	262144
2B-1B2P-1S(16x16)-8r	2	16000	8000	131072	61	256	8000	131072
4B-2B2P-1S(32x32)-2r	4	8000	2000	262144	7.6	1024	2000	262144

Table 3-6 Representative Narrow-band RADAR Observing Modes

* Number of spectral channels limited by size of double-buffered recirculation memory. If necessary, triple buffering could increase this by several factors but may be cost prohibitive.

4 Proposed Hardware Architecture

This section contains a description of the proposed hardware architecture of the WIDAR correlator for the EVLA. It contains figures and explanations of hardware that, it is believed, can be implemented with technology available (or soon to be available) at this writing. Further refinement may be required, but the descriptions given could be used as a starting point to allow serious hardware design and implementation to begin.

4.1 Architecture Overview

Figure 4-1 is an overview of the proposed architecture showing all signal and data paths, circuit boards, and systems. Figure 4-2 shows the phasing subsystem. A description of each follows.

In Figure 4-1, the digital signals start out as sampled data at the antennas. The sampled data is transmitted to a central site via a Fiber-Optic Transmission System (FOTS). From there it goes to an Optical Patch Panel or switch that allows antennas at different sites to be connected to the correlator depending on the antenna array configuration. Fiber optic receivers on the Station Board convert the data to electrical signals where digital delay compensation and sub-band filtering is performed. Each Station Board processes 2 x 2 GHz basebands and so the output of 4 Station Boards from each antenna is rearranged by the Sub-band Distributor Backplane into 16 sub-band cables each containing data from 8 basebands. Each of the 16 cables from the Sub-band Distributor Backplane (for each antenna) then goes to a sub-band correlator rack where it is fanned-out to multiple Baseline Boards with the Station Data Fanout Board. The Station Data Fanout Board is necessary because each Baseline Board does only a fraction of the baselines required. For a 40-station correlator, a fanout of 5 is required. (Aside: if the correlator were to be expanded to include more stations and baselines, then daisy-chained Station Data Fanout Boards would feed more correlator racks containing more Baseline Boards. In this case no re-design would be required—just replication of existing hardware.)

Data from 18 stations (or, perhaps 16) enters the Baseline Board via the Baseline Entry Backplane. This backplane is simply a vehicle for getting the data on the Baseline Board via blind-mate connectors. The Baseline Board correlates data from 9 "X" antennas and 9 "Y" antennas for a total of 81 baselines—one 9x9 parallelogram of the entire baseline matrix. If this board is too big, then this can be scaled back to say, 8 "X" antennas and 8 "Y" antennas for a total of 64 baselines (or 8x8 parallelogram). Indeed, the 8 x 8 configuration may be optimal for a 40-station correlator.

The backend control/readout computers are proposed to be PCs running Linux with standard Ultra-SCSI2 (or 3) interfaces, large disks and tape archive/backup devices. USCSI2/3 interfaces on the Station Boards and Baseline Boards allow very high performance data transfer (40 to 80 Mbytes/sec per USCSI2 bus) and a simple upgrade path to inexpensive, more powerful computing as it becomes available in the future. A separate 'Beowulf' cluster of PCs (which is not part of the correlator per se) can be used for back-end real-time data processing and imaging.

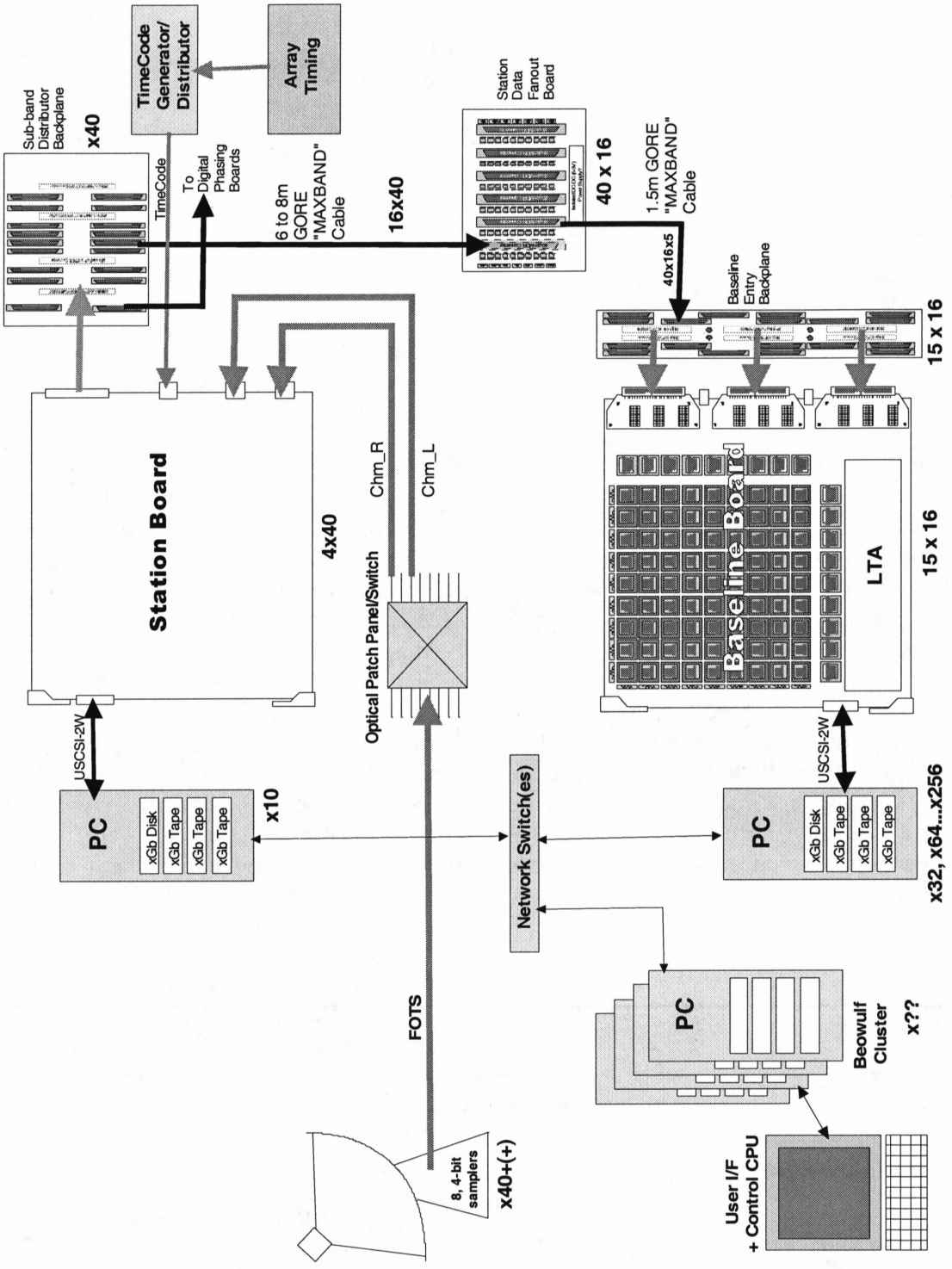


Figure 4-1 Proposed WIDAR EVLA Correlator overall hardware architecture showing the number of instances required for each module for the full correlator.

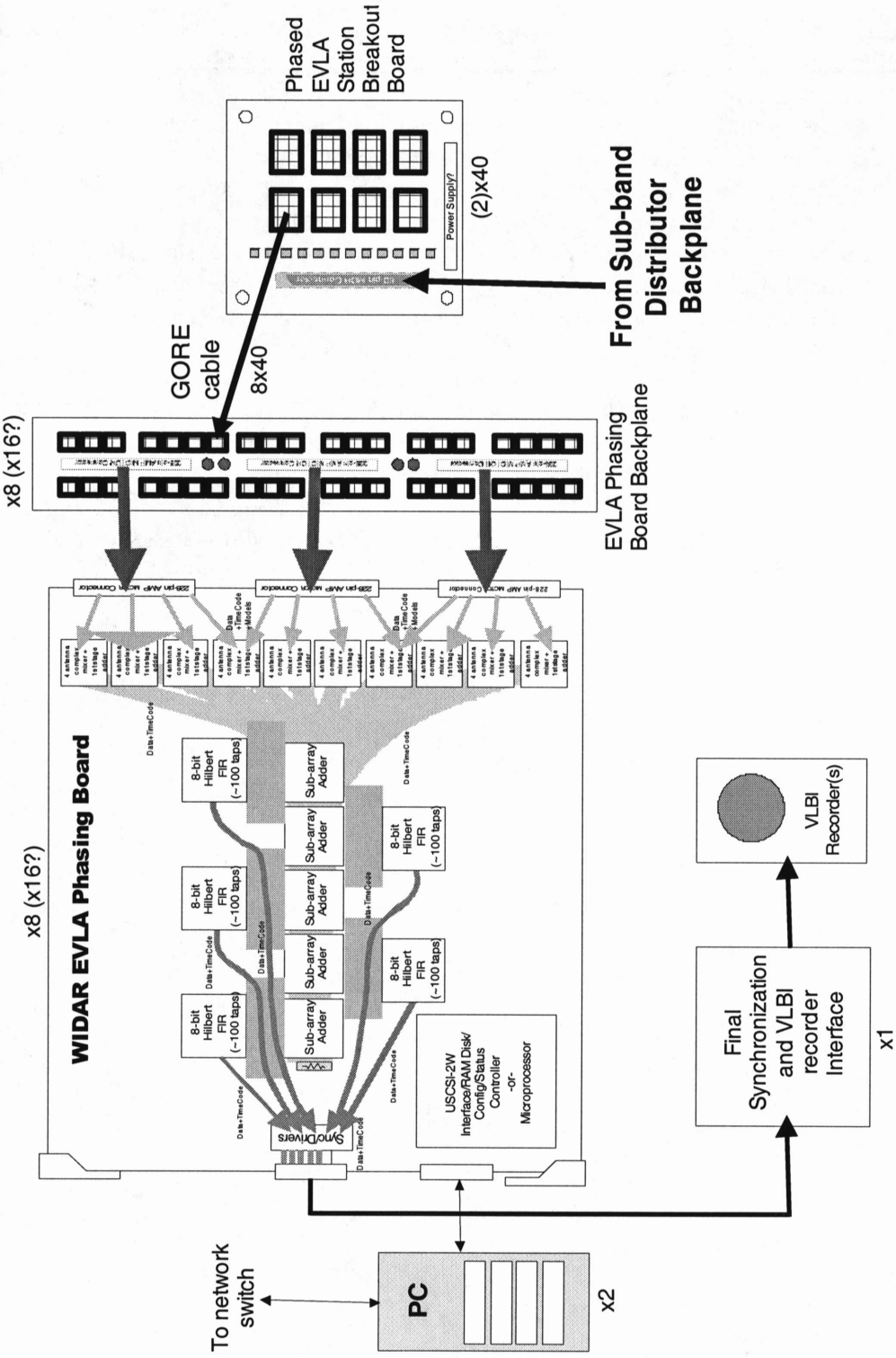


Figure 4-2 Proposed WIDAR digital phasing subsystem. Each phasing board produces phased output for multiple sub-arrays and one sub-band of one baseband.

The digital phasing subsystem shown in Figure 4-2 operates as follows. Special outputs from the Sub-band Distributor Backplane contain multiple sub-bands from 1 or 2 basebands. These are shown in Figure 4-1 as going "To Digital Phasing Boards" and on Figure 4-2 as coming "From Sub-band Distributor Backplane". A cable (or more than one cable—but we will consider just one for now) from each station goes into a "Phased EVLA Station Breakout Board" for each station where the multi-sub-band data gets broken out into individual sub-band data streams. Each of these sub-band data streams from each station goes into an "EVLA Phasing Board Backplane". Thus, the EVLA Phasing Board Backplane has data for one sub-band, from one baseband, for all 40 antennas coming into it. The EVLA Phasing Board plugs into the backplane and can thus phase all antennas or selected sub-arrays of antennas on one sub-band.

Phased outputs from multiple EVLA Phasing Boards go into the "Final Synchronization and VLBI recorder Interface"[6]. This is where final sub-band-to-sub-band synchronization and synchronization to VLBI timing occurs before data is finally transmitted to one or more VLBI recorders. This final step assumes **that EVLA array timing is frequency locked to VLBI timing** (i.e. a hydrogen maser). If this is not the case, then a D/A and A/D step is required—something that should be avoided for maximum signal integrity.

[6] It is not yet clear whether this final piece of hardware is part of the NRC correlator effort or not—it is an issue that must be decided.

4.2 Individual Hardware Module Descriptions

The following subsections provide a reasonably detailed description of each hardware module required in the system.

4.2.1 Station Board

The Station Board shown in Figure 4-3 performs all of the station-based processing for one baseband pair from one antenna. Thus, there are 4 Station Boards required for each antenna, and for a 40-station correlator, 160 boards will be required. The Station Board contains the fiber optic receiver module/modules, geometric delay compensation, sub-band FIR filters, wideband autocorrelators, and generation of delay models, phase models, and timing required for the downstream correlator boards.

The data enters the board via the dual-channel Fiber Optic Receiver Module (or, 2 single-channel modules). This module demultiplexes the data down from the multi-Gbps fiber to 2 outputs, each of which is 4Gs/s data time-demultiplexed by 16 with 4 bits per sample for a total of 128 bit streams at 256 Mbits/sec each. The interface from the module to the motherboard is defined in such a way as to allow any receiver/data generator to be plugged into it. Once the correlator (or part of it) is installed, old VLA antennas with 50 MHz bandwidth per baseband can be integrated into the new correlator by using a converter module instead of the fiber optic receiver module. Nominally, this converter module would have a 50 MHz analog baseband in, sampled at 128 or 256 Ms/s. The data is then translated to 4 Gs/s using zero-insertion interpolation that causes replicas of the sampled signal to appear across the 2 GHz band. The output is then presented in the same fashion to the motherboard as is done by the fiber module (i.e. demultiplex by 16, at 4 bits per sample). Processing by the rest of the board proceeds as usual except that only 50 MHz of the 2 GHz band is unique. If the 50 MHz band is already sampled (i.e. we don't have access to the analog baseband signal), then the digital data can still be zero-insertion translated, but the correlator clock rates must be adjusted accordingly). It must be decided whether this Old Antenna Converter Module (OACM) is an NRC or NRAO responsibility.

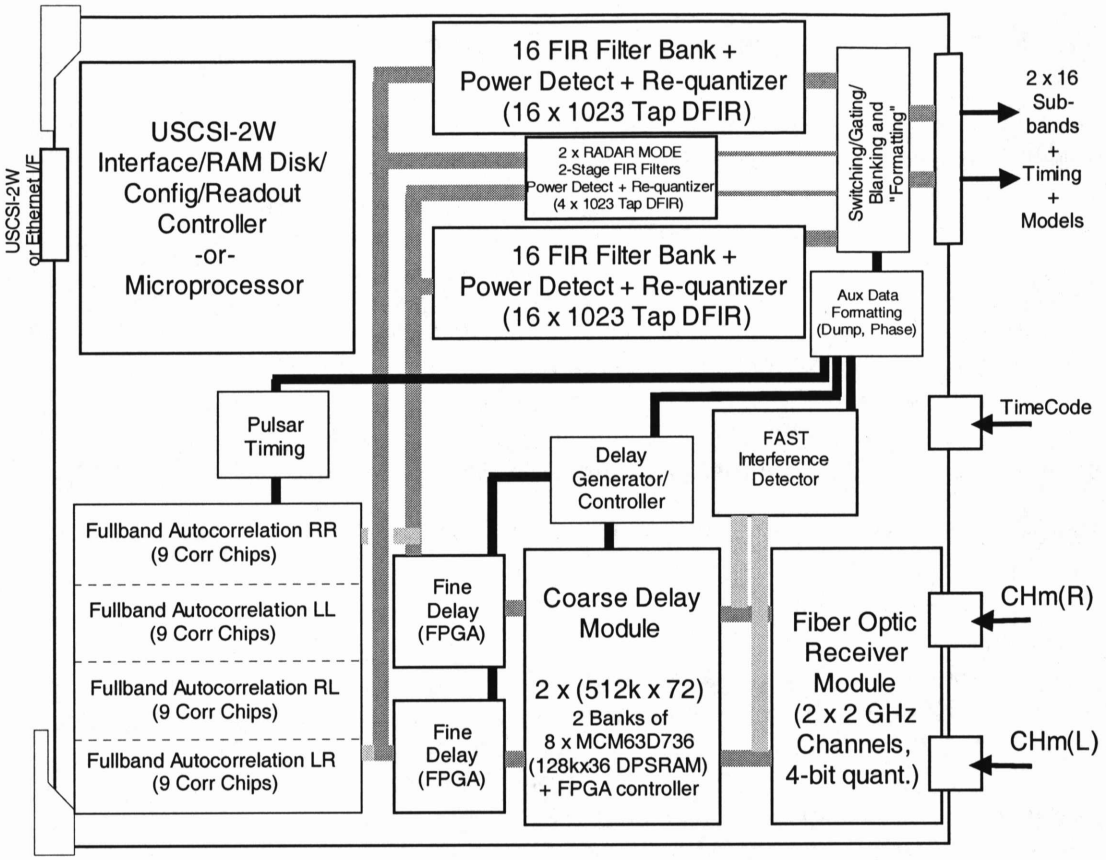


Figure 4-3 Station Board block diagram. The data from the antenna enters into the Fiber Optic Receiver Module where it comes out in N=16 time-demultiplexed format. From there it goes into the coarse delay module and fine delay that performs delay compensation to +/- 0.5 samples at the 4 Gs/s rate. FIR filters then generate sub-bands at up to 256 Ms/s each.

The 128 bit streams of data (plus two timing synchronization streams—one for each baseband) or 2 x 64 streams, go into the Coarse Delay Module. It is proposed that this module is a daughter board as well so the memory (and therefore baseline length) can be upgraded if longer baseline antennas are added to the array. For +/-300 km baselines, this module must provide (512k x 64) x 2 memory. As indicated in the figure, this can be done with 16 DPSRAMs (Dual-port synchronous static RAMs), each of which is 128k x 36 operating at 128 Ms/s. This module simply changes delay to within 1 sample precision at 256 Ms/s—equivalent to within 16 samples of precision at 4 Gs/s. Final *integer* delay compensation to +/- 0.5 samples at 4Gs/s is performed with Fine Delay FPGAs on the motherboard. These can be on the motherboard since their function will not change if more or less coarse delay memory is available. These devices are switches that re-route time-demultiplexed data streams and insert or remove 1 sample of delay on each stream depending on what the delay is. The Coarse Delay Module and Fine Delays are in these locations so that all delay changes at all three levels (coarse, fine, very fine[7]—with phase rotator offsets in the correlator) happen in precise time

[7] Very fine delay provides effective compensation to 1/16 of a sample of delay and is describe in the Appendix on WIDAR signal processing.

synchronization. For example, a coarse delay change immediately effects a fine delay change which, after the delay through the FIR filters, effects a very fine delay change. Additionally, the delay for each R and L baseband is independently programmable and therefore can have a different delay/phase-center on the sky.

Data out of the Fine Delay now goes into the bank of DFIR (Demultiplexed FIR) filters. These filters are designed to operate on time-demultiplexed data with decimation $2^k \cdot N$ (recall that N is the demultiplexing factor; k is an integer). There are N=16 filters for each baseband. The outputs of each filter are single 4-bit sampled data streams at 256 Ms/s (i.e. sub-band outputs). Two special purpose very narrowband filters are also available for Radar Mode and for very narrow sub-bands. These filters are implemented as 2-stage DFIRs and allow bands as narrow as 30 kHz to be produced. It could be argued that each of the sub-band DFIRs should be 2-stage DFIRs since this affords the maximum in flexibility. However, the 2-stage DFIR requires a second 4-bit quantization step and a "re-demultiplexer"—still, it may be worth considering designing the DFIR chip to have 2-stage capability. A block diagram of the DFIR architecture is shown in Figure 4-4.

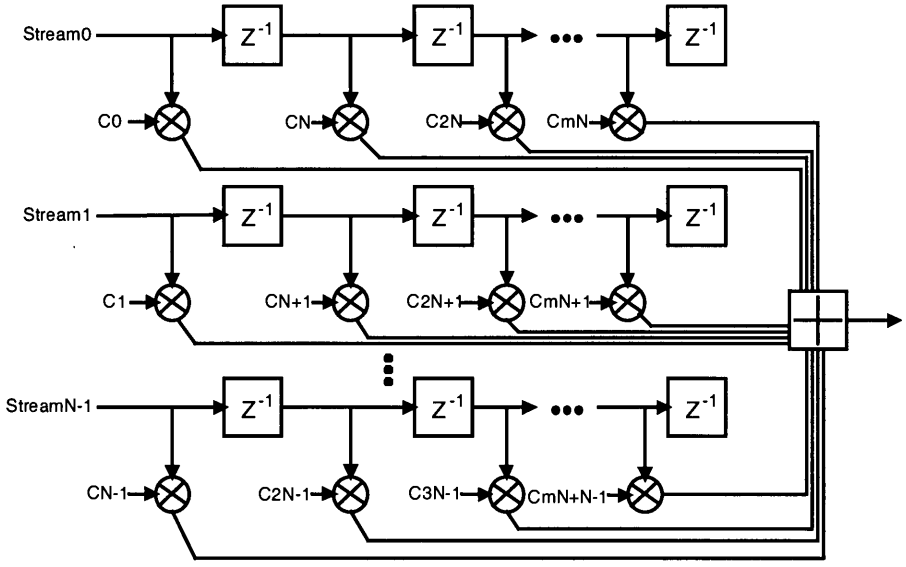


Figure 4-4 DFIR (Demultiplexed FIR) chip architecture. Bandpass filtering *and* decimation results in an efficient parallel architecture. Note that a DFIR chip could be cascaded in width and length for larger N and/or more taps.

After filtering, the sub-band data streams (which are individual sampled sub-bands with no demultiplexing) go into a switch[8] and gating/blanking module which allows one or more sub-band's data to be replicated into other sub-bands' transmission channels. This allows flexibility in correlator lag assignment. Also, the output of each sub-band can be gated/blanked with pulsar gates or the interference detector.

[8] Switching can also be performed by programming sub-band filters the same—effectively replicating the same sub-band signal in different transmission channels. However, physical switching may yield a more precise, unambiguous solution.

One weakness of the WIDAR design is that once the sub-bands are produced, it is not possible to obtain the autocorrelation spectrum of the wideband signal[9]. Luckily, it is not too difficult to perform this function on the time-demultiplexed data since the data is all on the Station Board. To achieve this, every demultiplexed data stream must be cross-correlated with every other stream—including itself. For N=16, this requires 136 (calculated from $\frac{N(N+1)}{2}$) cross-correlations per correlation product. This can be done with 9 correlator chips, each one performing 16 cross-correlations. Thus, for all of the 4 polarization auto-products, a total of 36 correlator chips (same as the correlator chips used on the Baseline Board) are required. It is proposed that this function be placed on a daughter module so that the motherboard is not too large.

The purpose of the "FAST Interference Detector" is to detect bursts of interference and very quickly gate the data off after it comes out of the sub-band FIR filters. This is placed before the Coarse Delay module so that enough time is available to do this before the interference shows up out of the filters. Nominally, 512k deep memory allows 1 millisecond of delay—but on longer baselines, this delay could drop significantly. (If at least 1 millisecond of delay (allowing 1 millisecond of integration time) is needed, then the Coarse Delay RAM will have to be made larger.) Conceptually, the interference detector is an autocorrelation spectrometer[10] attached to at least one data stream. Thus, some very high-speed real-time spectrum analysis (or lag analysis?) would have to be done to determine if interference is present and react accordingly. This could be very difficult and it is probably a good idea to put this circuitry on a daughter board. If burst interference timing is known well enough in advance, then gating/blanking could simply be done with an on-board programmable timer.

The purpose of the "Aux Data Formatting" function is to get additional timing and synchronization signals ready to be transported with the data to the downstream correlator boards. This includes the signals DUMPTRIG, DELAYMOD, PHASEMOD, and TIMECODE—described in a following section. In order to simplify the implementation of the "Sub-band Distributor Backplane" that the Station Boards plug into, each Station Board (or perhaps just one 'Master' Station Board) must generate the delay and phase models for all basebands and sub-bands. This is simply a "register stuffing" operation that should not incur any significant overhead.

The "...Config/Readout Controller..." (daughter) module will provide an external CPU with a way of programming the board (FIR filters, models etc.) and reading out autocorrelation coefficients, status information etc. It is proposed that this interface be an Ultra-SCSI2 (or 3) interface with Dual-port RAM. A controller (simple program control FPGA) on the Station Board side of the RAM will configure the board as commanded through the RAM interface, and readout coefficients and status and deposit them in the RAM interface. A controller on the SCSI side of the interface will enable data transfer

[9] This isn't entirely true, it is possible to obtain the auto-power spectrum of a baseband as long as it is replicated with a slightly different frequency shift, and is sampled separately. A cross-correlation can then be performed with the anti-aliasing methods active. The ability to do this depends on the agility of the analog system.

[10] It is reasonable to assume that a total power (lag 0 autocorrelator) detector would not be useful since the effects on total power detection after quantization would be significantly reduced by the quantizer's AGC.

to/from a master on the SCSI bus. The advantage of using this standard interface is that the board now just looks like a RAM-disk 'device' to the attached CPU (PC) and it is simple to upgrade the CPU or replace a failed CPU with a COTS (Commercial Off-The-Shelf) box. If this is not acceptable, then a CPU could be put on this board with a 100 Mbit (or Gbit) Ethernet interface or a "Front Panel Data Port" type interface currently in use in some NRAO equipment. It is highly desirable for maintenance purposes to use the same interface (and the same daughter board) for the Station Board, the Baseline Board, and the Phasing Board. Ultra-SCSI 2/3 offers greater than 40 Mbyte/sec data transfer rate—effectively eliminating correlator induced bottlenecks to data flow (especially out of the Baseline Board).

Finally, to enable correlator system testing without using data from antennas, it is proposed that a test vector generation memory be put on-board in front of the Coarse Delay Module similar to what is done in [1]. This memory can be loaded with test vectors that could be anything from simple repeating patterns to simulated sky data complete with frequency shift and delay. Operationally, this capability will allow all of the downstream hardware to be tested on a regular basis thus guaranteeing the integrity of the correlator system. In principle, this test vector memory could just be the Coarse Delay Module memory, but for performance and simplicity it may be better to use a separate memory with the capability of switching it into the data path going into the Coarse Delay Module.

4.2.2 Sub-band Distributor Backplane

A layout diagram of the "Sub-band Distributor Backplane" is shown in Figure 4-5. The purpose of this backplane is to rearrange the data so that there are 16 output cables, each one containing all baseband data, models, and synchronization from one sub-band. Each output then goes to one sub-band correlator. As discussed in a later section, arranging the correlator in this fashion permits a great deal of flexibility for trading off the number of basebands, number of sub-bands, and spectral resolution per cross-correlation product without the need for a switch in front of Station Boards. This backplane also rearranges the data for transmission to the Phasing Subsystem. In this case, individual cables contain multiple sub-bands from 1 baseband pair. This is done to allow one Station Board (and therefore one baseband pair) to be configured for phased-VLA operation while the other 3 Station Boards can be configured for normal interferometer correlator operation. This allows a great deal of flexibility when operating the VLA in simultaneous phased and interferometer mode.

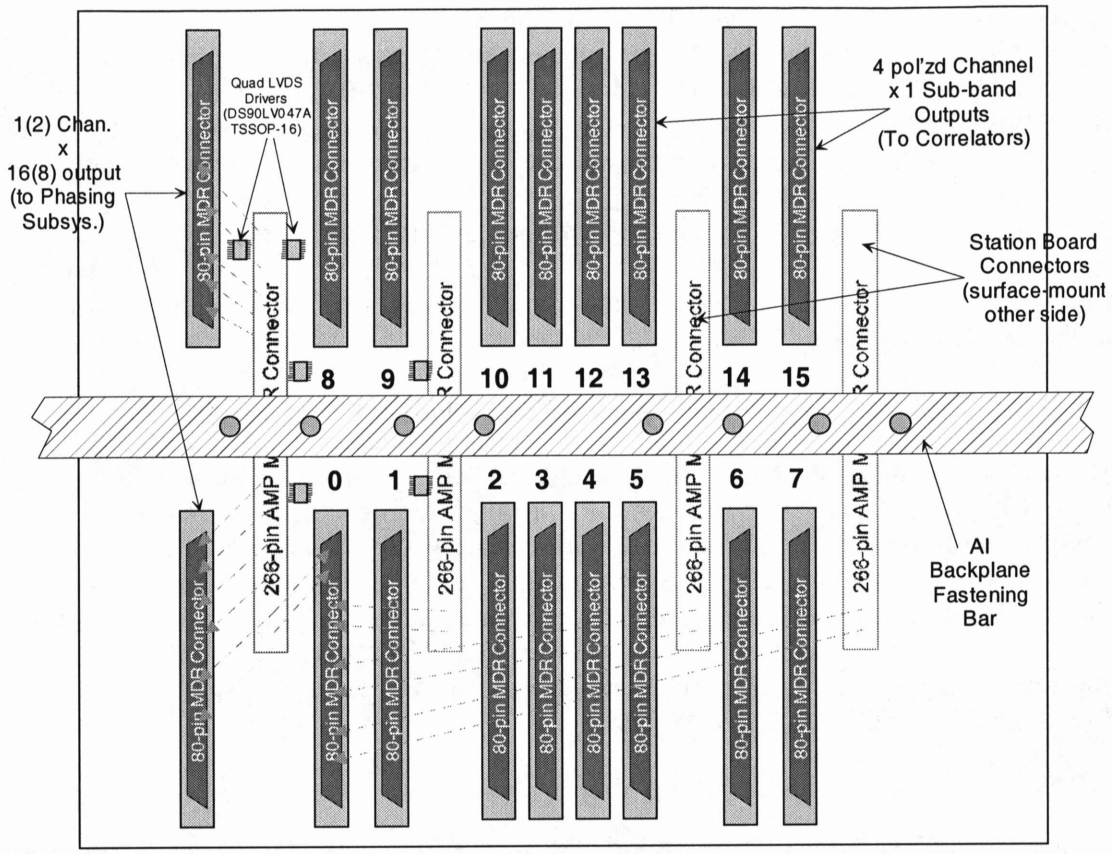


Figure 4-5 The Station Sub-band Distributor Backplane. This backplane re-routes the data so that each of 16 cables going to the correlator contain all of the data, timing, models, and synchronization for one sub-band. Cables going to the Phasing Subsystem are arranged to carry multiple sub-bands from 1 (or 2) baseband pairs.

4.2.3 Fiber Optic Entry Backplane

Figure 4-3 shows fiber optic cables and a "TimeCode" entering the Station Board. Blind-mate access to these by the Station Board will require another backplane not shown in the figure. This backplane may be integrated with the Sub-band Distributor Backplane for a monolithic solution or it may be a separate circuit board to allow more flexibility when connected to antennas that do not yet use fiber. (i.e. connecting to old VLA antennas as previously mentioned or "foreign" antennas that may require different connectors and a different module in place of the Fiber Optic Receiver Module.)

4.2.4 Station Data Fanout Board

In nominal wideband correlator modes, each of 16 sub-band correlators cross-correlates the data for one sub-band from all baseband pairs. A sub-band correlator contains several correlator boards each of which correlates a "parallelogram" matrix of up to 81 baselines. A 45-station baseline matrix is shown in Figure 4-6, showing parallelograms with sides of length 9. This means that each correlator board must have data from 9 'X' stations and 9 'Y' stations for a total of 18 stations. This arrangement also requires that data from 1

station go to 5 correlator boards. The purpose of the Station Data Fanout Board is to provide this fanout capability. The Station Data Fanout Board is shown in Figure 4-7.

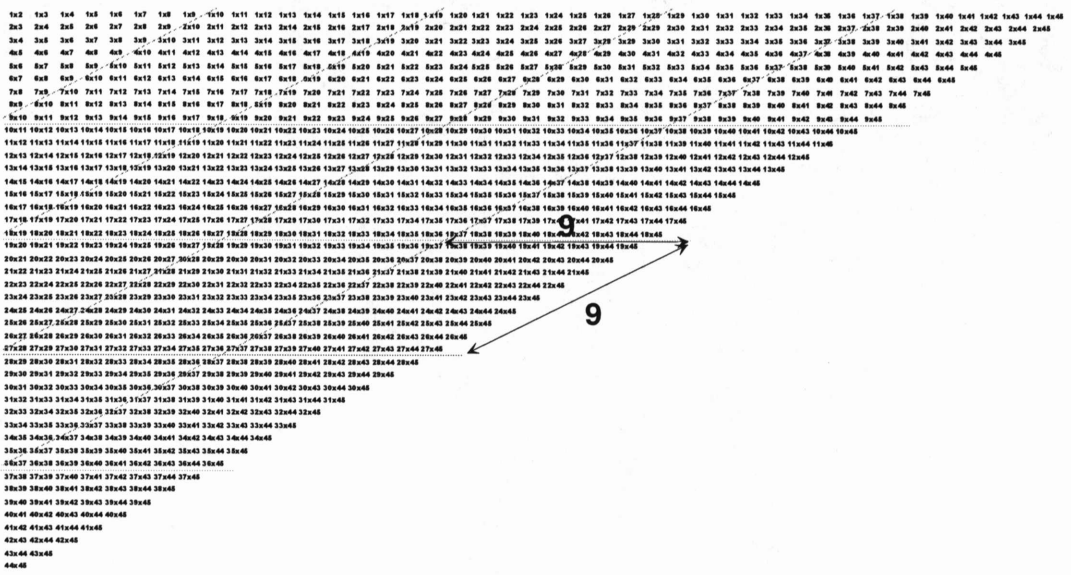


Figure 4-6 Example 45-station baseline matrix. Each correlator board does 81 baselines, requiring input from 9 'X' stations and 9 'Y' stations for a total of 18 stations. 15 correlator boards are required to do all baselines.

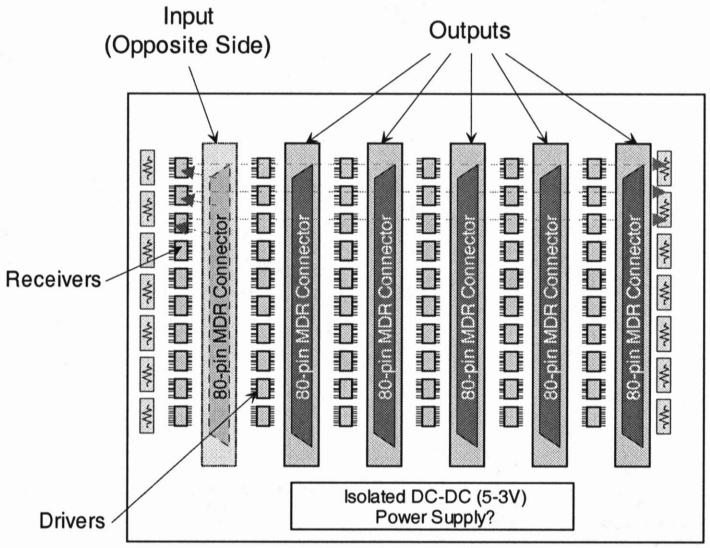


Figure 4-7 Station Data Fanout Board. This board fans out data from one station (all basebands, one sub-band) to up to 5 correlator (Baseline) boards.

A 40 station VLA correlator would probably correlate 64 baselines on each Baseline Board, requiring 16 station input (8 'X' and 8 'Y') with the same factor of 5 fanout requirement.

4.2.5 Baseline Entry Backplane

This backplane is shown in Figure 4-8. Its purpose is to route data from 18 (or more likely, 16) stations into the Baseline Board.

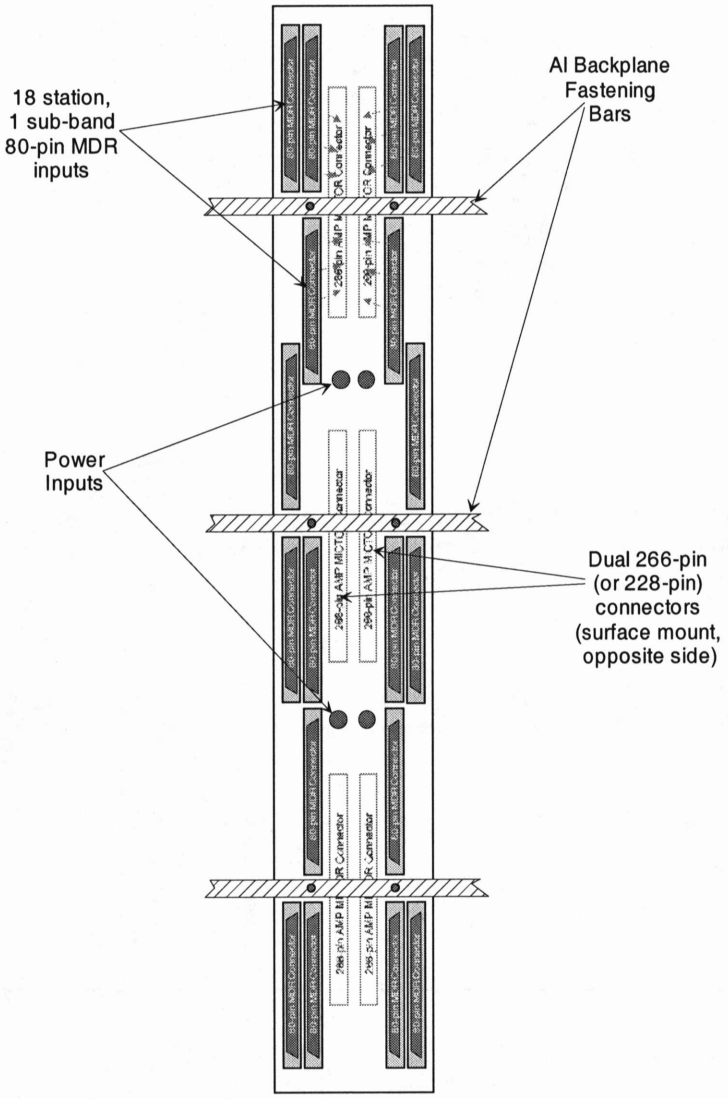


Figure 4-8 Baseline Entry Backplane. This backplane routes the data from up to 18 stations (all baseband pairs, one sub-band) to a Baseline Board. This board contains no drivers or active circuitry.

The combination of the Station Data Fanout Board and the Baseline Entry Backplane ensures that data from the Station Board to the Baseline Board is routed in a point-to-point fashion mostly over impedance controlled cable. This is important for signal integrity at 256 MHz clock rates.

4.2.6 Baseline Board

This board is where up to 81 cross-correlations are performed. The Baseline Board is shown in Figure 4-9.

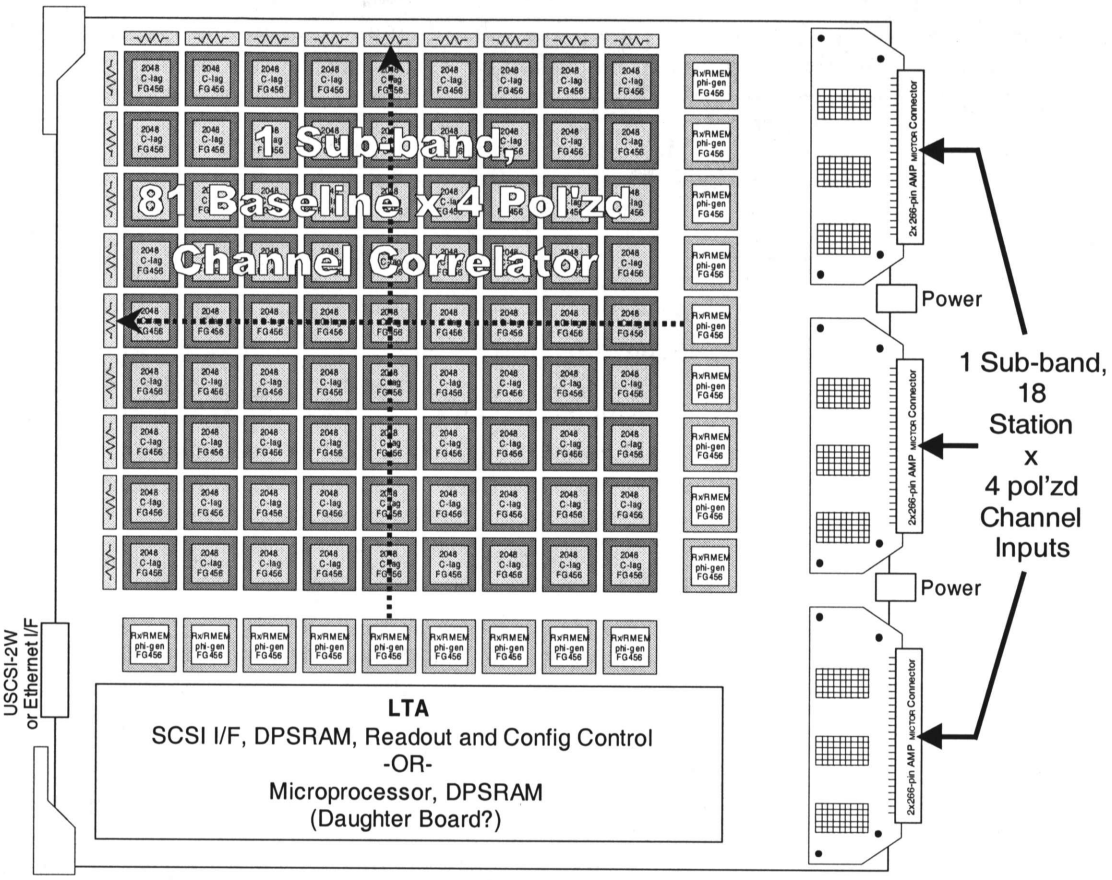


Figure 4-9 Baseline Board layout/block diagram. The board shown does 81 baselines with 18 station inputs although for a 40-station correlator it may be better to do 64 baselines with 16 station inputs. Data from each station goes to a receiver chip complete with recirculation memory, a delay-to-phase lookup table, and drivers to be able to drive correlator chips on a tapped transmission line as shown.

Data from the Baseline Entry Backplane gets into the Baseline board with 6 AMP MICTOR 266-pin controlled impedance connectors. These connectors are straddle surface-mount to the edge of the PCB, and so an entry daughter board is required to get 6 connectors on one Baseline Board. The daughter boards gain entry to the motherboard via commercially available ball surface-mount connectors as indicated. Ideally, it would be better to eliminate the daughter board entry but the pin density requirements are such that this daughter board will most likely be required. With 4-bit sampled data, there are 32 data and 5 clock and control differential lines per station. For 18 stations, this comes to 666 differential lines or 1332 contacts—requiring just over 5, 266-pin connectors. If only 16 stations go into a board then 5, 266-pin connectors are required. If some spectral dynamic range is sacrificed with 3-bit requantization then (without including a data valid

line) 4, 266-pin connectors are required—eliminating the need for a straddle-mount daughter board solution.

Unfortunately, the insertion force required for these particular connectors is prohibitive. If 6, 266-pin MICTOR connectors are used, then the maximum insertion force is 400 lbs—clearly not feasible! Fortunately, there are other manufacturers of high-density connectors. Hypertronics Corp. makes a 303 contact[11], straddle surface-mount connector with a 1.5 oz insertion/extraction force per contact that should be suitable. If 4-bit data is used, then for 18 stations, 5 connectors are required with a daughter board. For 16 stations, 592 differential signals are needed requiring 4 connectors. At 4.567 inches per connector, 18.3 inches of PCB edge is required—just able to fit within a 12U, 19 inch height circuit board without a daughter board and with room for power entry. If 4 of these connectors are used, the insertion/extraction force is 113 lbs—a bit more than the 81 lb insertion force for a 12U card with 4 x 96 pin DIN 41612 connectors. However, 113 lbs of insertion force does not seem prohibitive.

From the above discussion, it would seem that the Baseline Board should be designed for 64 baselines with 16 station inputs to eliminate daughter boards, extra cost, and prohibitive insertion forces. Detailed design will have to seriously consider these issues but at this point it seems feasible for the board to be designed as discussed.

The "Rx/RMEM phi-gen Block"

Each station's data gets received and synchronized on the Baseline Board by the "Rx/RMEM phi-gen" block in Figure 4-9. This block contains an FPGA, double-buffered recirculation memory (**RMEM**), and a delay-to-phase RAM lookup table for very fine delay tracking. The FPGA receives and re-synchronizes station data, timing, and models *to its own* clock that is transmitted with the data—there is no need or attempt to synchronize to any other station's clock at this time.

The FPGA contains up to 8 phase generators that take the phase models (PHASEMOD) and generate 4-bit phase for each sample for the sub-band of each active baseband (see Figure 4-10 for the proposed cross-correlator architecture). Each phase generator is a 1st order, 32-bit digital frequency synthesizer commonly used in VLBI correlators. For very fine delay tracking, the FPGA takes delay for each of up to 8 basebands and converts it into a phase that gets added to the phase accumulator in the synthesizer before quantizing the 32-bit phase to 4 bits. Since delay for each baseband is sent serially in one stream with the data, this operation is performed with a lookup table in a single 2k x 8-bit RAM. The contents of the lookup table depend on which sub-band slot within the particular baseband is active.

The recirculation memory (RMEM) consists of two 2M[12] x 32 DDR SDRAMs (Double Data Rate Synchronous Dynamic RAMs). DDR SDRAMs are proposed to be used

[11] Hypertronics also makes a 372-contact connector with 1 oz insertion force per contact—although it is not known if this connector is available at the time of this writing.
[12] Longer RAMs will most likely be available when the correlator is implemented. The 32-bit width restricts use of RMEM to 4 basebands.

because they offer the largest memory, highest speed, and lowest cost compared to synchronous static RAMs. A large memory is necessary to provide for a large number of spectral channels at low bandwidths with minimum SNR loss and so the correlator chip accumulator readout rate is reasonable—2M RAM requires the accumulators to be read out every 8 milliseconds. These RAMS are now becoming available for core memory in new-generation high-speed PCs. DDR SDRAMs have some latency for the first "CAS" access but can operate at the 256 MHz clock rate[13] for subsequent accesses for a "RAS" page (up to 2k words long). This means that when using the RMEM, every 2k words or so, about 3 or 4 samples will have to be flagged invalid. Two RAM devices are required so that only single port RAM is necessary and so that while one RAM is being written to, the other RAM is read out and fed to the correlator chips multiple times (with a different start address each time to synthesize different delays). This memory will also serve as "precomputation delay" memory when correlator chips are effectively chained together for higher spectral resolution at full-speed (i.e. 128 MHz sub-band bandwidths). It should be noted that it is planned that both data *and phase* are written into the RAM—this simplifies the design and eliminates the need for complicated "rewind" control of the phase generators.

The use of RMEM does incur an additional SNR loss that is a result of the finite length of the RAM. This SNR loss is according to the following equation[14]:

$$SNRloss = \frac{tot.\#of\ freq\ pnts}{\sqrt{2} \cdot RAM\ size}$$

For 2M RAM and 262k frequency points, this works out to an SNR loss of about 9%.

Finally, this block transmits clock, data, phase, synchronization pulses, and dump triggers to the correlator chips in its column or row. Currently, it is planned that the correlator chips tap off a terminated transmission line. The BGA/FGA package for the correlator chip, the simple routing of the transmission line, and the number of chips being driven makes this seem feasible. However, it may be necessary to have each correlator chip receive and re-transmit the signal if further study shows that the tapped transmission line is not reliable (although this may require too many chip I/O lines).

Correlator Chip

Each correlator chip correlates the data from one sub-band for up to 4 baseband pairs on one baseline. Final data/clock synchronization for the X and Y stations will occur on the correlator chip itself: one station will be deemed to be the master, and the other station's data will be synchronized to it and its clock. Since the station receiver FPGA has already performed stream-by-stream synchronization, this final synchronization step should be simple with the use of an additional frame clock synchronized to the 1PPS (so that only

[13] Actually, the clock rate is 128 MHz, but with accesses on both the rising and falling edges of the clock—hence "DDR".

[14] It is possible to use triple-buffering to reduce this loss so that it is proportional to the number of frequency points synthesized by a particular RMEM and correlator chip but it may be difficult to implement and may be cost-prohibitive.

clock phasing to synchronize the frame clocks is necessary). Since there are 4 baseband pairs, the correlator chip must perform up to 16 separate cross-correlations. A proposed architecture for each cross-correlator (not including required data switches etc in the correlator chip itself) is shown in Figure 4-10 and Figure 4-11.

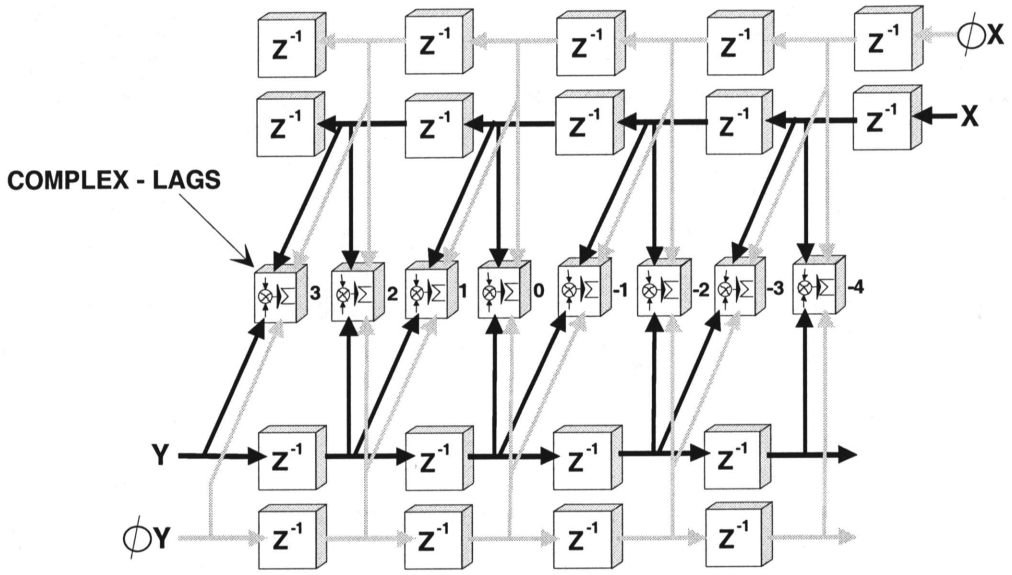


Figure 4-10 Proposed cross-correlator architecture. The phase is carried with data and final fringe rotation is performed at each lag. Since the phase rates are small compared to the sample rates, the phase does not contribute to the power dissipation of the device. If, as in other VLBI XF correlator chip architectures, the Y-station in-phase and quadrature data is generated with a fringe rotator at one end of the delay line, additional power will be generated by the chip as it carries the now quickly changing complex data along the delay line.

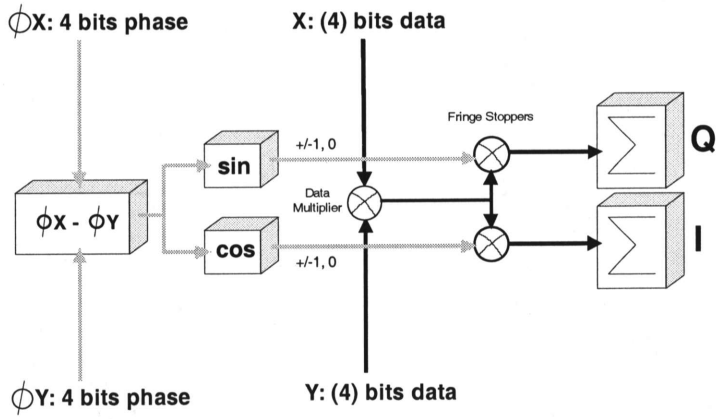


Figure 4-11 Correlator chip lag. 4-bit quantized phase is differenced to form the baseline phase. This goes into sin and cos lookup tables whose outputs are +/-1 and 0 for 3-level fringe rotation or +/-2, +/-1, and 0 for 5-level fringe rotation. 5-level fringe rotation achieves a 6 to 8 dB improvement in spectral dynamic range in the presence of strong narrowband signals compared to 3-level fringe rotation. This design has an advantage in that only one 4-bit multiplier is required—which is then further qualified by a shift and/or sign change in the fringe stopper. This could have a significant impact on the feasibility of a 4-bit correlator chip. A full 4-bit/15-level system with 5-level fringe rotation incurs a total sensitivity loss of about 1 – (0.985*0.985*0.975) = 5.5%

The correlator chip also double-buffers the accumulators so that correlation can occur while readout is in progress and each of the 16 cross-correlations should be able to be read out independently (so each baseband can have its own integration time). Additionally the chip contains circuitry to allow internal chaining of 128 complex-lag correlators for higher spectral resolution on fewer basebands. Figure 4-10 does not show support for oversampling (two delays per lag instead of one). This is an option that can be included at the expense of increased chip complexity. Finally, the correlator chip will always operate at the 256 MHz clock rate. Data in each cross-correlator (and therefore each baseband or baseband pair) can be at different, related rates and this is facilitated with the use of data enable clocks that enable the shift registers at the appropriate times. Running the correlator chip at the maximum clock rate ensures that power dissipation—and therefore power requirements—remain approximately constant and simplifies clock generation and distribution (i.e. there is only one clock and there is no need to worry about sub-rate clock phase etc.).

Some preliminary work on a correlator chip design has been done by the low-power VLSI design group (Margala et al.) at the University of Alberta. The chip is still in the design stage and they expect to fabricate a prototype version[15] with 16 to 64 lags in 0.18 μm CMOS in August 2000. The design is based on the lag architecture shown in Figure 4-10. This preliminary investigation has yielded the following results:

- A 4-bit "Braun-Wooley" multiplier uses $1/10^{th}$ to $1/8^{th}$ of the total power of the chip and about $1/5^{th}$ of the total number of transistors in the lag (not including the delay-line shift registors). The 4-bit multiplier uses about 400 transistors. Reducing it to a 3-bit multiplier would reduce it to 225 transistors.

- Each lag—including 22-bit ripple-counter accumulators with 7-bit pre-accumulators for a total of 29 bits—will take about 2200 transistors. This does not include the readout/storage registers that may be implemented in distributed RAM blocks for efficiency. Based on this estimate, a 2048-lag correlator chip would be about a 4 to 5 million transistor device.

- Based on this work, they estimate that a 2048-lag correlator chip, with a 250/256 MHz clock, fabricated in 0.18 μm CMOS with a ~1.6 V power supply will dissipate about 1.5 W.

From the above results, it is clear that a 2048-lag correlator chip operating at 250 MHz is within the ballpark of feasibility. In the worst case, even if the power requirements exceeded the estimate by a factor of 2, it would still be reasonable to build a 1024-lag chip resulting in a factor of 2 decrease in spectral resolution.

[15] This prototype will be fabricated using the CMC (Canadian Microelectronics Center) facilities and is only meant to test the feasibility and power requirements of the architecture—not act as a prototype for the real EVLA correlator chip itself.

Chaining Correlator Chips

To provide very high spectral resolution in wideband modes, it is possible to chain correlator chips in columns or, indeed, chain all correlator chips on a Baseline Board. If this chaining is performed, it would reduce the number of baselines that can be processed and the baselines that are available would be fixed by the routing of the data to the Baseline Boards. This may be a useful mode (not included in the mode tables at the beginning of this document) for wideband searches of narrow lines on a reduced number of baselines. At some point, a decision needs to be made whether this option should be included in the design or not.

Support for VLBI

The correlator chip of Figure 4-10 contains all of the elements of a VLBI chip. Some essential ingredients are missing however to allow the chip and the Baseline Board to be used in a VLBI correlator. A fine/vernier delay and phase modifier must be generated on a baseline basis to achieve required +/-0.25 sample effective baseline-based delay tracking accuracy [1]. This requires that the delay models go to the correlator chip and that the chip calculates the baseline-based delay. It is expected that this capability will add little to the complexity of the chip. It must be decided if the chip, and therefore the Baseline Board itself must be capable of VLBI so that the appropriate logic can be included.

Readout, LTA and Processor Interface

As with the Station Board, it is proposed that the Baseline Board have an Ultra-SCSI2/3 interface to the outside world. The LTA (Long Term Accumulator) consists of Dual-Port Static RAM (DPSRAM)[16]. An FPGA controller[17] on the correlator chip side of the DPSRAM reads out data and stores it in DPSRAM for retrieval over the SCSI bus. The controller also reads a configuration area in the DPSRAM and configures correlator chips, Rx/RMEM FPGAs, and delay-to-phase lookup table RAMs as configurations get requested through the DPSRAM interface. Dumping of the correlator chips is entirely triggered by the DUMPTRIG signal generated on the Station Board and carried with the data, so there is no need for "hard" real-time control of the board with a microprocessor. As readout data appears in DPSRAM, a PC connected to the SCSI bus reads out the data and frees up memory. A controller on the SCSI side of the DPSRAM provides SCSI access to the DPSRAM.

4.2.7 Alternative Correlation Configuration

An alternative arrangement to the configuration discussed in the previous sections would be to have each correlator board only correlate data for one sub-band and one baseband pair. In this case, many more baselines must be correlated on one correlator board,

[16] The 'DPSRAM' will most likely be implemented as a bank of high-speed SDRAM SIMMs (used in PCs) with interleaved accesses from the correlator chip side and SCSI side since DPSRAM does not provide enough memory capacity for this application.
[17] Or maybe more than one depending on the required readout rate.

although it would seem infeasible to correlate all baselines for one baseband pair and one sub-band on one board. In fact doing so, with 128 lags per correlation product, all polarization products, and a 2048 lag correlator chip would require 195 correlator chips on one board—clearly infeasible. It would be better to use 4 correlator boards and a correlator chip doing 4 baselines (for a total of 16 correlator boards for all baseband pairs). Here, the 40-station baseline matrix is split into 4 equal triangles—each triangle a 20-station correlation matrix.

This configuration would require a distribution backplane taking the place of the Station Data Fanout Board and the Baseline Entry Backplane that performs the fanout of the data to the 4 correlator boards. (This distribution backplane would be fed with data from all antennas but only one baseband pair—requiring all 40 cables plugging into it or a breakout board for each station providing access to the desired baseband pair.) In this case, each correlator board requires access to up to 40 stations—each one requiring baseline-based recirculation memory[18], a controller, delay-to-phase lookup table etc. A slight advantage here is that only 40 x (2 x 4 + 5) = 520 (differential) data lines (assuming 5 clock, model, and control lines and assuming 4-bit samples) is required to go into each correlator board. However, routing of the data on a correlator board cannot be done as shown in Figure 4-9 (Baseline Board) and this may pose serious problems at a 256 MHz clock rate. Also, if this option is chosen, the correlator would not be expandable beyond 40 stations since the 4 x 4 board arrangement would be specifically tailored to 40 stations.

The primary advantage of this "quadrant" architecture is that the correlator could be built and delivered in quadrants. Each quadrant would correlate all baselines and sub-bands for 1 baseband pair. Thus, the EVLA could get all antennas "on the air" with 1 baseband pair with only a quarter of the correlator hardware. This is not a particular advantage for the EVLA since old antennas will be upgraded to new antennas over a period of a few years. If the correlator is delivered (even one sub-band at a time over a period of time) before all antennas are upgraded, then it is more important that data from old antennas be able to be processed by the new correlator with new antennas. A mechanism that could achieve this is described in the section on the Station Board.

The primary disadvantage of the "quadrant architecture" is that flexibility in the correlator is reduced. In the proposed architecture, it is possible to easily tradeoff baseband pairs for spectral resolution since within a sub-band correlator, all basebands go to every correlator chip. Providing this tradeoff is simply a matter of programming the correlator chip. To provide the same functionality with the alternative quadrant architecture would require a large, high-speed switch before or after the 4 Station Boards (for each station)—something that could be difficult and costly to construct.

Nevertheless, it may be worthwhile studying this alternative configuration in more detail than presented or considered here to determine if it is better than the configuration being proposed.

[18] Meaning that the data must be delayed differently for each station depending on whether it is the 'X' or 'Y' antenna in the correlation.

4.2.8 Phased EVLA Station Breakout Board

This board is part of the Phasing Subsystem shown in Figure 4-2. Its purpose is to breakout data from a phasing cable coming from the Station Sub-band Distributor Backplane so that each individual sub-band of one baseband (a single sampled data stream) can be routed to a Phasing Board. This board is shown in Figure 4-12. Nominally, 2 boards like this will be used per station—one for each baseband of a pair. More boards (with more outputs from the Station Sub-band Distributor Backplane) can be used if more basebands are required. Fewer sub-bands of more baseband pairs can be accomodated—its all a matter of the Sub-band Distributor Backplane.

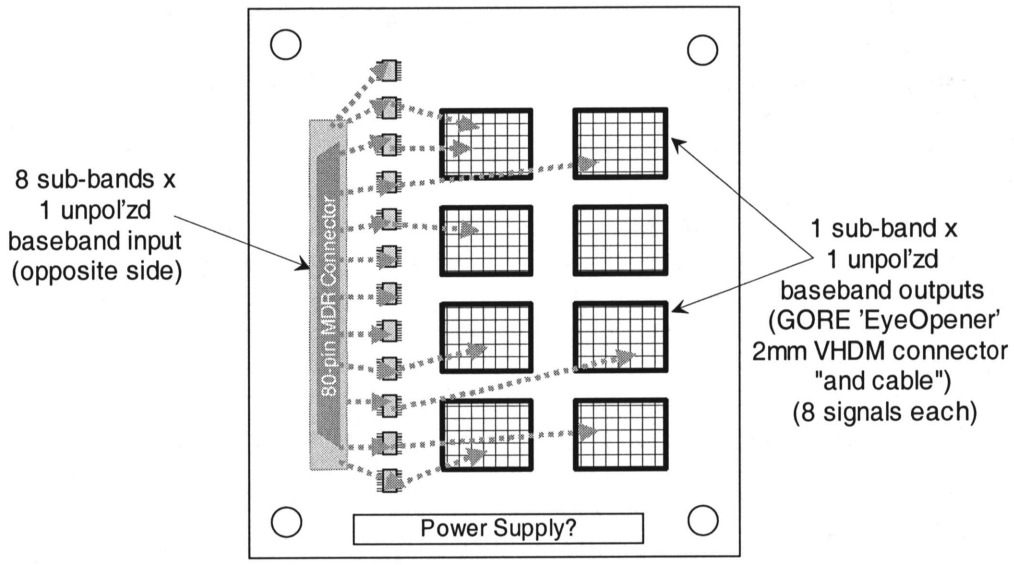


Figure 4-12 Phased EVLA Station Breakout Board. This board breaks out multi-sub-band data for one baseband to individual connectors which can then each be routed to one Phasing Board.

4.2.9 EVLA Phasing Board Backplane

This backplane is required to get data from 40 stations (one sub-band of one baseband or one sample stream) onto one Phasing Board. This can be done as shown in Figure 4-13. In the figure, there are 40 connectors coming from the Station Breakout Boards and gaining entry to the Phasing Board with three 266-pin AMP MICTOR connectors (or Hypertronics connectors...).

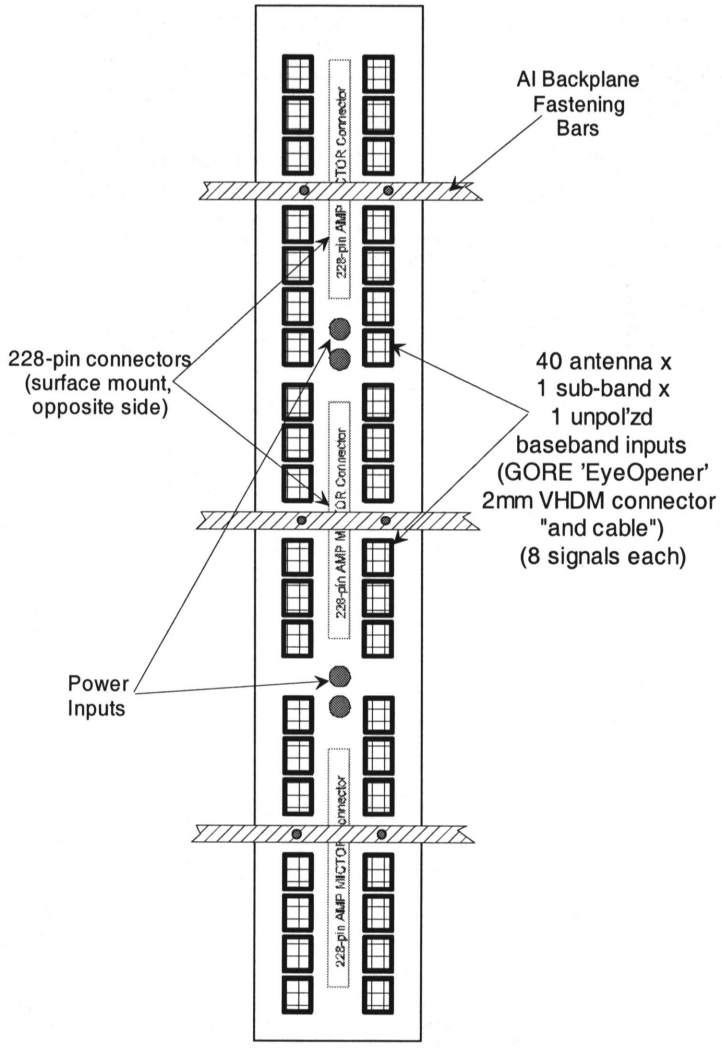


Figure 4-13 The EVLA Phasing Board Backplane. This board facilitates entry of data from 40 stations—1 sub-band of one baseband (one sample stream) from each station—into a Phasing Board.

4.2.10 EVLA Phasing Board

This board, with a proposed pseudo-layout/block diagram shown in Figure 4-14, is where data from up to 40 stations are phased together to produce an output to a VLBI recorder (Figure 4-2). The board actually allows multiple sub-array outputs to be generated—the maximum number of sub-arrays that can be generated is a design parameter and will be fixed once the board is implemented.

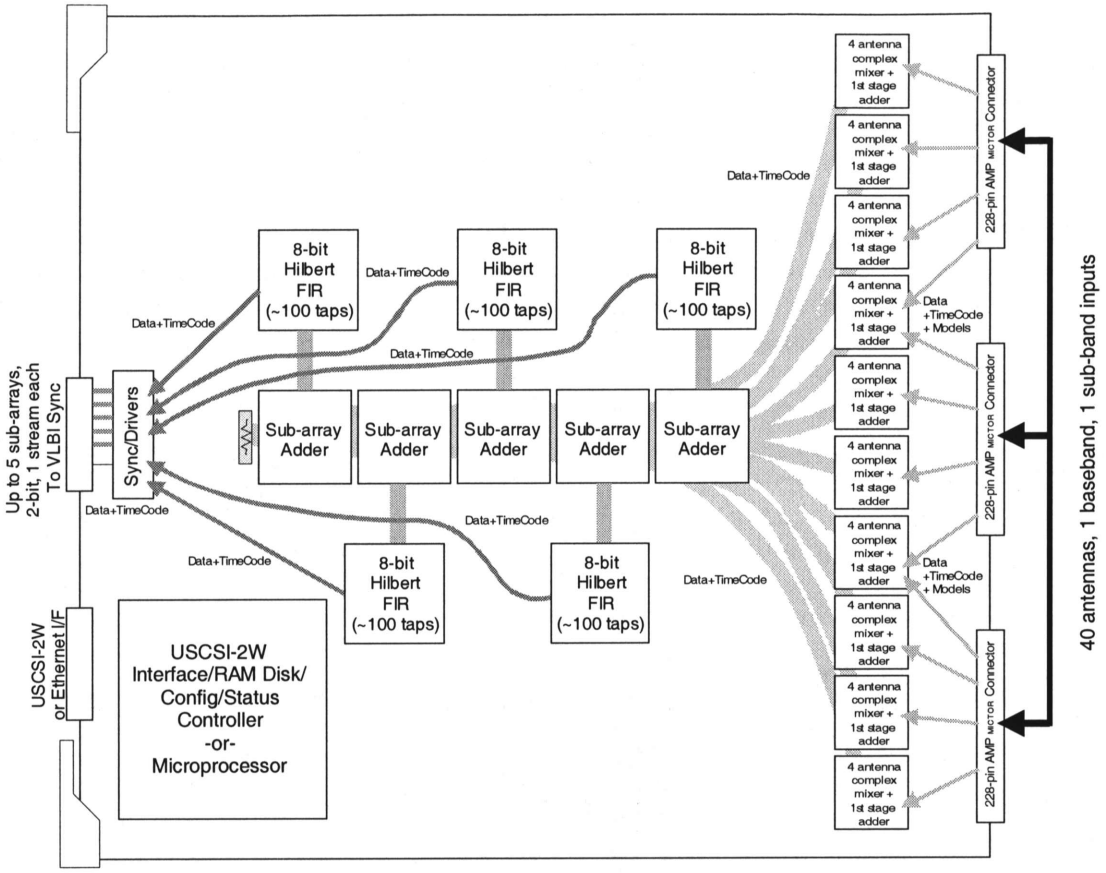


Figure 4-14 The EVLA Phasing Board. Data from 1 sub-band of 1 baseband from up to 40 antennas/stations are phased together in one or more sub-arrays. The board is completely digital, and if the VLBI clock is locked to the VLA clock, does not require a D/A and A/D step.

Data enters the board and goes into 10, "1st stage mixers and adders"—each of which takes data from a fixed 4 stations/antennas[19]. It is desirable to use a fixed assignment of stations into the first stage adders to eliminate the need for a massive (and probably impossible to implement) switch which would be required to allow complete flexibility. However, reasonable flexibility is gained since it will be possible to enable or disable phasing of *each* station within each group of 4. For example, if stations 1, 8, 13, and 23 are to be phased together (using 4, 1st Stage Adders), then other antennas in 1st Stage Adders 1, 2, 4, and 6 cannot be used for phasing.

The 1st Stage Adders perform a complex mix and add of 4 stations to the phase-center of the array or sub-array. This action stops fringes, removes any (array *differential*) frequency shifts, and performs very fine delay correction (delay-dependent phase offset) as determined by the delay and phase models from the Station Board(s). This action therefore restricts the phase-center of the phased-array/sub-array to be the same as the phase-center of the interferometer array/sub-array (for the particular baseband—since

[19] This number is nominal and can be changed to something else that may be better. Increasing this number decreases flexibility, while decreasing it increases the size of the downstream Sub-array Adder.

each baseband can have a different delay/phase center on the sky). The output of each 1st Stage Adder is a complex (nominally) 8-bit number (i.e. 8-bits for the in-phase component and 8-bits for the quadrature component) which gets routed to all downstream Sub-array Adders.

Each Sub-array Adder (5 are shown and there can be as many sub-arrays as there are sub-array adders) adds the complex data from selected 1st Stage Adders to produce a phased complex output. The quadrature component of this complex data goes into a -90^{o} phase shifter implemented with a ~100 tap Hilbert transformer FIR filter. This Hilbert transformer should be able to be implemented in a reasonable size FPGA since every other tap coefficient is zero, and, except for a sign change, the tap coefficients are symmetrical about the center tap. The in-phase data is delayed appropriately and finally added to the phase shifted quadrature data to produce a final (real) sampled phased output.

The phased outputs are synchronized (using their TIMECODEs carried with them), requantized, and transmitted to a final synchronization/VLBI interface box as shown in the Phasing Subsystem diagram of Figure 4-2.

The Phasing Board is proposed to have the same processor interface as the Station Board and the Baseline Board. This interface is used to allow a PC to set the configuration of the board and read out any status information.

4.3 Correlator Rack Layout

Figure 4-15 and Figure 4-16 are side and front views of possible layouts for the racks containing the Baseline Boards. These racks will have the highest-density intra-rack cabling due to the fanout of the data from the Station Data Fanout Board. In the figures, the blue "ducting" is where the cable runs would potentially be. Station Board racks would have a similar layout only the intra-rack cabling is not as dense. Figure 4-17 is an artist's concept of what the entire correlator might look like.

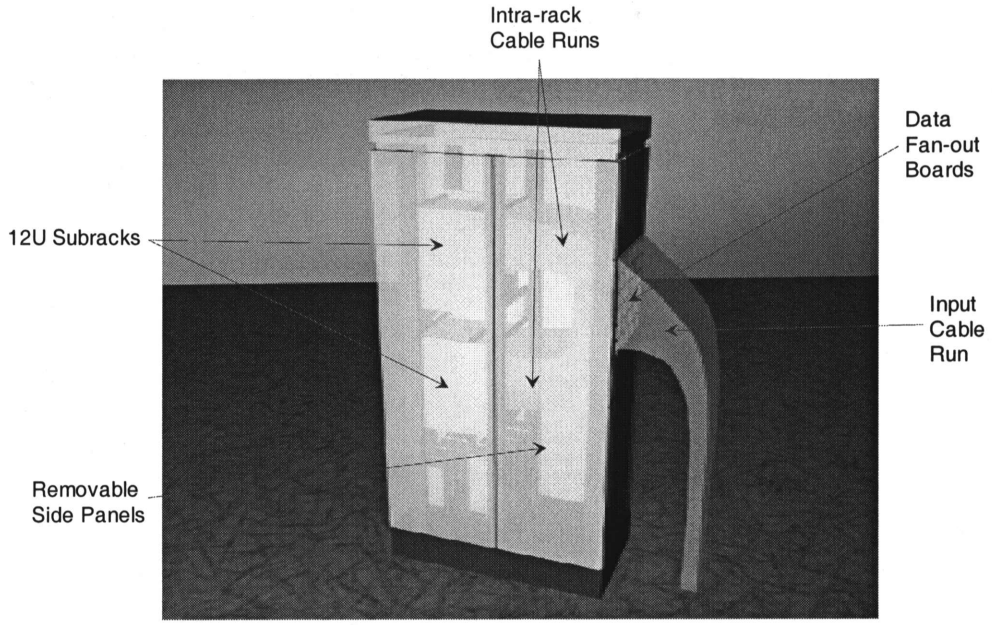


Figure 4-15 Side-view of possible layout of a sub-band correlator rack containing Baseline Boards. The blue "ducting" are cable runs for the sub-band data cables coming from the Station Boards (Sub-band Distributor Backplane). Each sub-rack contains up to 8 Baseline Boards. There will be 16 such baseline racks and 10 similar station racks in the correlator system.

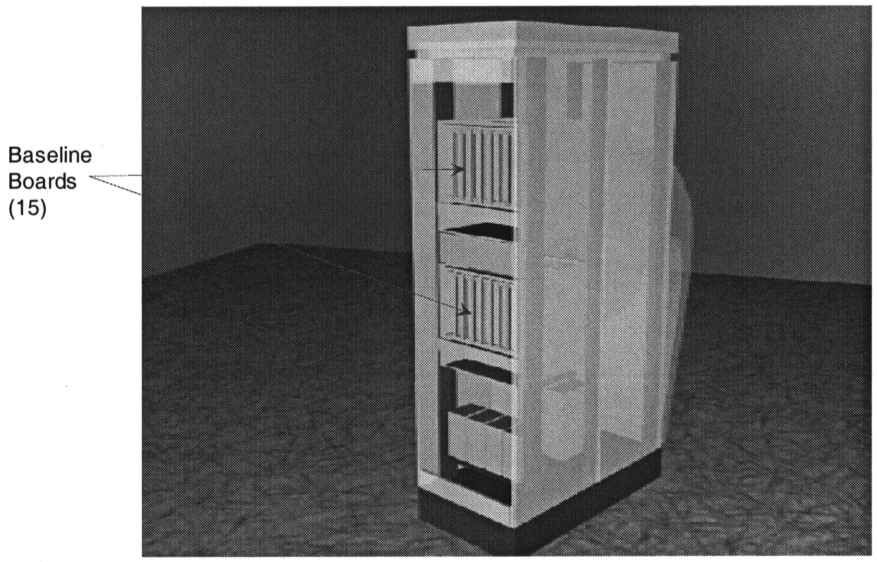


Figure 4-16 Front-view of possible layout of a sub-band correlator rack containing Baseline Boards.

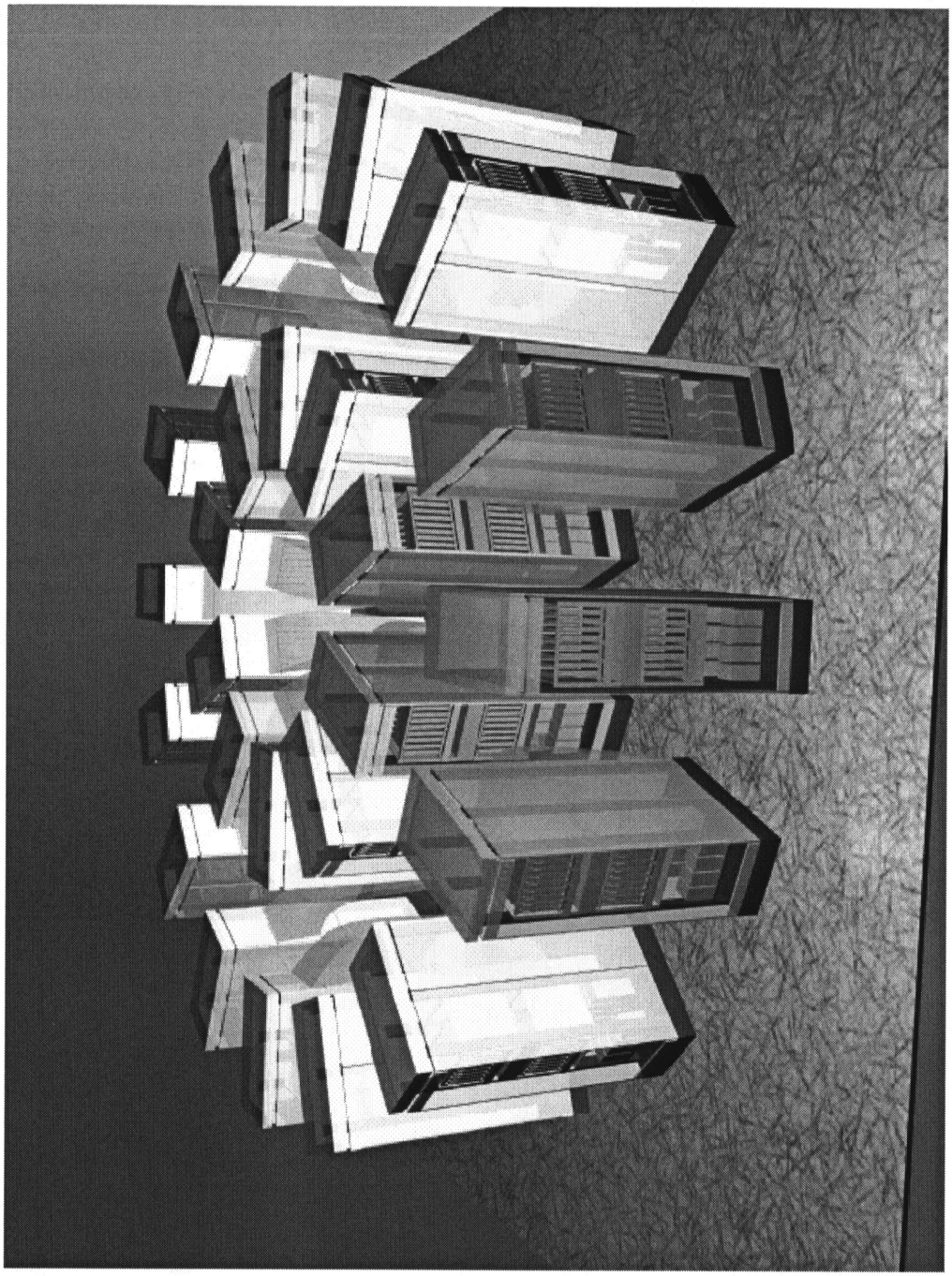


Figure 4-17 Artist's concept of what the full correlator system might look like. The inner racks are station racks and the outer ones are baseline racks. This arrangement may be useful to keep station-to-baseline cabling reasonably phase-matched and of minimum length for best signal fidelity.

4.4 Data, Model, and Timing Formats

This section describes proposed methods for distributing timing, models, and other synchronization information throughout the correlator. The basic approach is to generate timing and synchronization information in key places and have it propagate throughout the system along with the data. This is an effort to simplify system synchronization, eliminate potential ambiguities, and always be able to determine system integrity.

4.4.1 Array/Correlator Timing Overview

A simplified overall timing diagram for the EVLA including the correlator and phasing subsystem is shown in Figure 4-18. A central TIMECODE and Clock get distributed to the antennas and the correlator. Local Oscillators, sampler clocks and all other system timing get synchronized to this master time although the figure does not presume to know anything about Local Oscillator and timing delivery to the antennas. TIMECODE and Clock get fed to the correlator via the Station Boards—there is one direct input and one input that travels along with the sampled data from the antennas. The Station Boards distribute the TIMECODE and Clock to the Baseline Boards where the correlated data gets saved with TIMECODE timestamps.

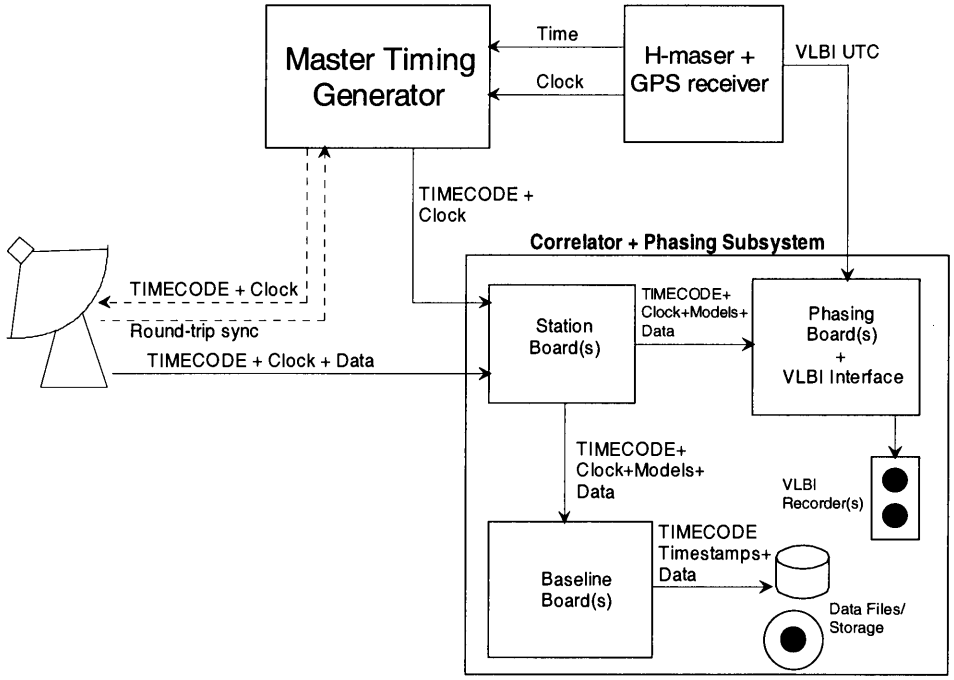


Figure 4-18 Simplified EVLA timing diagram. The dotted lines are conceptual and beyond the scope of this document. The correlator is fed with a TIMECODE signal and 256 MHz clock that is locked to the sampler clocks and (probably) ultimately a Hydrogen Maser time standard. TIMECODE gets distributed to Baseline Boards and Phasing Boards along with the data from the Station Boards.

4.4.2 Sub-band Cable Layout

The outputs from the Station Sub-band Distributor Backplane (Figure 4-5) go to 16 separate sub-band correlators containing Baseline Boards. The data and signals for one sub-band are in one cable as shown in Figure 4-19.

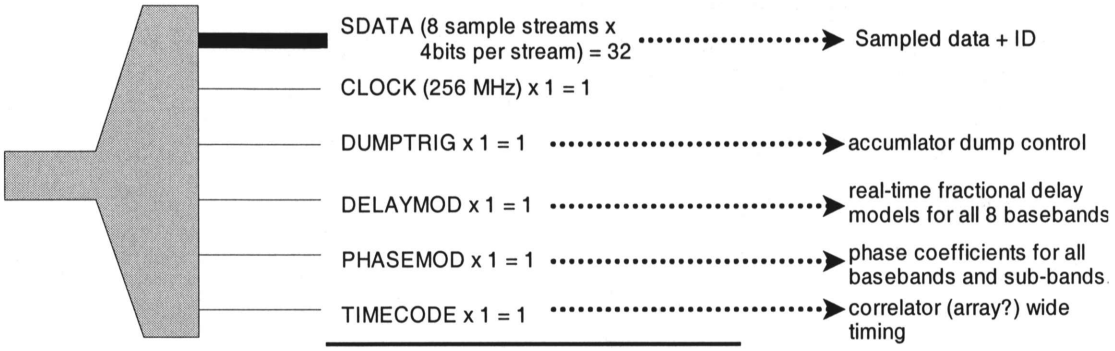
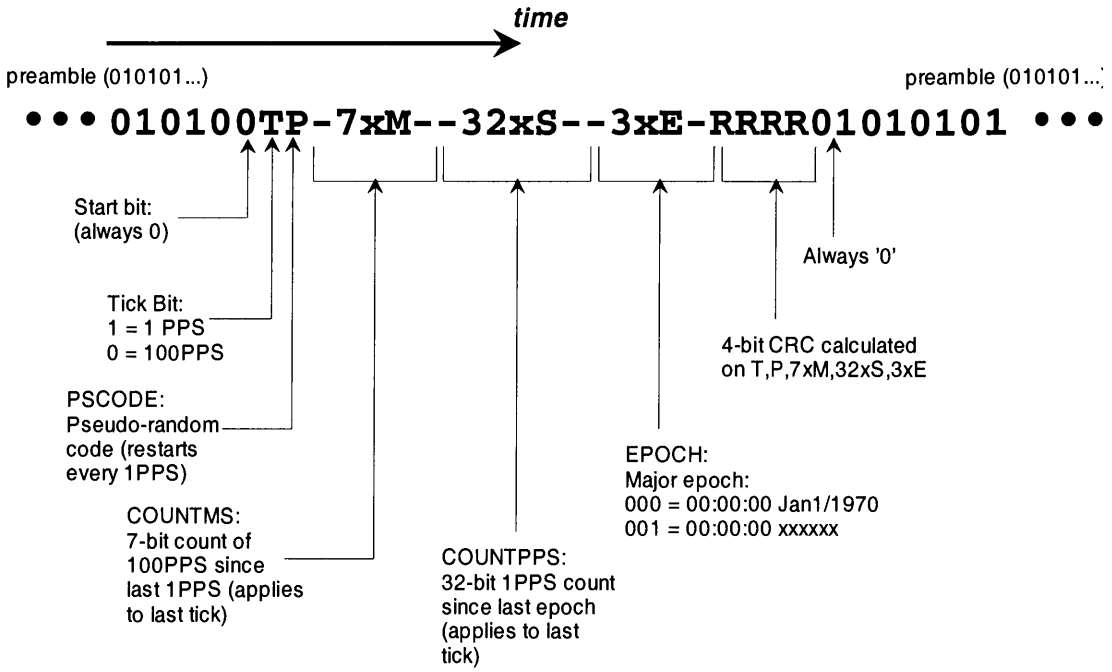


Figure 4-19 Sub-band cable output. All signals are (differential) LVDS and fit on one commercially available GORE cable. The data and clock rates are always at a constant frequency of 256 MHz.

The clock is at a constant rate of 256 MHz and all other signals are at the same bit rate. Synchronization of each signal to its clock is done in the receiver FPGA on the Baseline Board. This allows some timing skew between signals on the same cable and between different stations—however, the timing skew should be kept small since the synchronization circuitry will only be able to absorb small (~1 sample) differential delays. Final synchronization of X and Y data is performed on the correlator chip itself—using either the X or Y station clock as the master. A framing clock sent with the data from the receiver FPGAs to the correlator chip will simplify correlator chip circuitry required to the do the synchronization (i.e. the correlator chip will just have to synchronize to the framing clocks rather than any embedded information).

4.4.3 TIMECODE Format

TIMECODE is a single data stream that, along with a 256 MHz clock, contains all of the timing information for the correlator. A proposed format for TIMECODE is shown in Figure 4-20 below.



TIMECODE Format

Figure 4-20 Proposed TIMECODE format. This signal contains all absolute timing information required by the correlator. TIMECODE gets generated at one location and propagates through the correlator (and, perhaps, the entire array) with the data. All actions within the correlator use this as a reference, and all output data gets timestamped with timestamps using TIMECODE as a reference. TIMECODE requires an additional 256 MHz clock signal. 1PPS is a tick that occurs once per second, and 100PPS is a tick that occurs 100 times per second.

4.4.4 SDATA Format

The SDATA ("Sampled DATA") is the sampled and FIR filtered data coming from the Station Board. It is proposed that this be 4 bits wide and use the format shown in Figure 4-21.

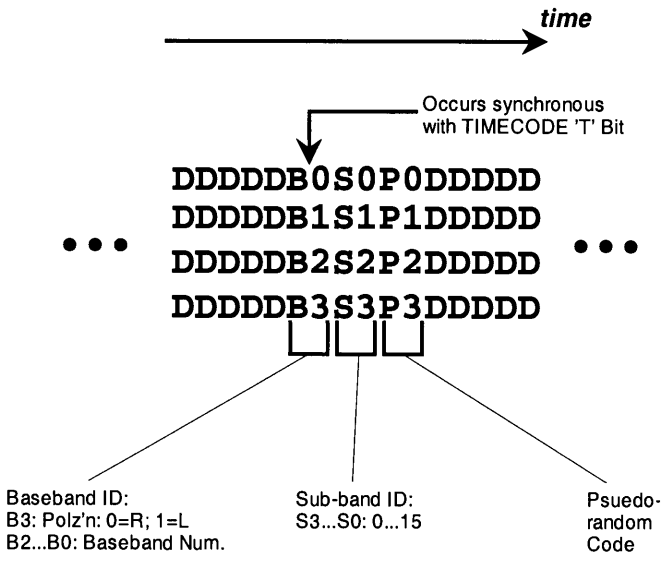


Figure 4-21 Proposed SDATA (Sampled DATA) format. Identifiers embedded in the data are synchronized to TIMECODE. The data is (nominally) 2's complement notation from –7 to +7 with the unused –8 state being used for flagging invalid or gated data. The embedded data will be set to the –8 state before going to the correlator chip so that no false correlation of embedded information occurs.

4.4.5 DUMPTRIG Format

DUMPTRIG is used to signal dumping of correlator chip accumulators. It is generated on Station Boards and thus allows for station rather than baseline-based dump specification. This is advantageous in that baselines that have stations that are further from the array center can be dumped faster. However, this is not the most efficient method since short baselines that are far from the array center will be dumped faster than they need to be. Nevertheless, the DUMPTRIG mechanism allows the specification and timing of dumping to be centralized to the Station Boards—potentially simplifying the system, eliminating possible time synchronization problems, and eliminating the need for an intelligent processor on the Baseline Board to keep track of time, trigger dumps, etc.

A possible format for DUMPTRIG is shown in Figure 4-22. DUMPTRIG is not synchronized to TIMECODE so it is possible for data to be dumped synchronous to any arbitrary event such as pulsar timing or other external timers. In order to ensure that low-level timing synchronization occurs (i.e. remove or determine time skews so that the exact sample when dumping occurs is known), it is necessary to run tests with the dump trigger synchronized to TIMECODE.

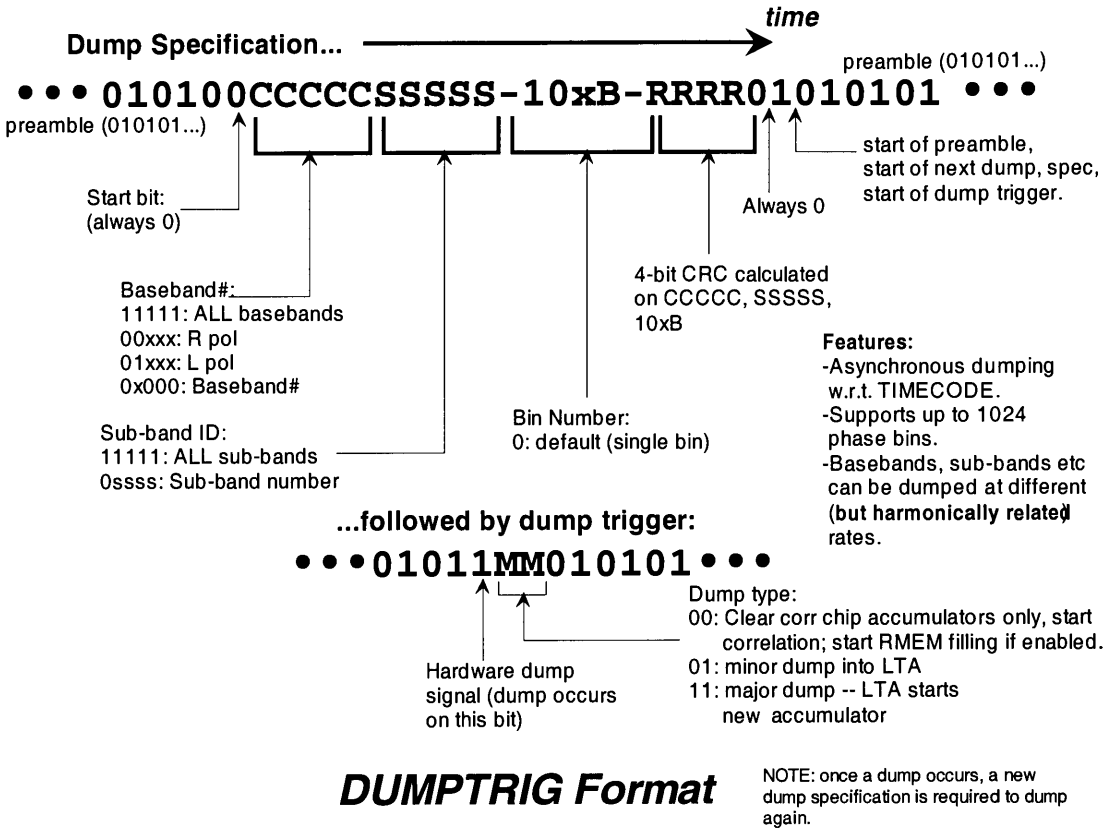


Figure 4-22 Possible DUMPTRIG format. With this format, dumps that are to be performed are set up and then executed with a "dump trigger". Dumping can occur synchronized to pulsar timing or external events. All of the dump control signals are embedded in the signal including capabilities for phase binning and long-term accumulation. Dumping of each sub-band within each of the 8 basebands can be different but it is necessary for dump timing within a given station to be harmonically related. When data gets dumped, it is timestamped with TIMECODE timestamps.

When data from the correlator chip gets dumped, it gets timestamped with TIMECODE and an offset that is a count of samples since the last 100PPS. This offset will most likely be calculated in the station receiver FPGA on the Baseline Board and is necessary for accurate timestamping of data.

DUMPTRIG Interaction with Recirculation Memory (RMEM)

There are some restrictions on the integration times that are available when recirculation modes are active. Recirculation uses two memory buffers and works by filling up a buffer (on a station basis with each station receiver FPGA on the Baseline Board) and, when full, recirculating the buffered data through the lag chain at high speed with different delay offsets each time (meanwhile the other buffer is filling) to synthesize a longer lag chain. Each time the buffer is burst through the correlator chip, the chip must be read out and saved in the LTA for that unique chunk of the lag chain. Thus, it is necessary to impose the restriction that the minimum integration time for any one cross-correlation product is the time it takes to fill up the buffer memory at the sub-band sample rate. A mitigating factor here is that the receiver FPGA can be configured to use

only a part of the memory available. This is an SNR issue (see previously defined equation) determined by the maximum lag chain that is to be synthesized and the sub-band sample rate (not the burst/recirculation rate into the correlator chip which is always larger up to 256 Ms/s). The minimum integration time is according to the following equation:

$$T_{\min} = \frac{mem\ size}{subband\ sample\ rate}$$

For example, with a sub-band bandwidth of 1/256th of 2.048 GHz (sample rate of 16 Ms/s), and a memory size of 2M (2000000) words, the minimum integration time T_{min} is 0.125 seconds. Because of the large buffer memory of 2Mwords that is being used, a lag chain up to 524288 long (using 2048 lags per sub-band correlator x 16 sub-band correlators) can be synthesized without significant loss of SNR. If only 1000000 words of memory are used, the integration time can be 0.0625 seconds, but the maximum synthesized lag chain should be reduced to 262144 lags to avoid significant SNR loss.

A *proposed*[20] mechanism for handling recirculation with the DUMPTRIG signal is as follows:

1. 1st buffer memory starts filling when the "Dump type" (MM in the dump trigger) is 00. Correlator accumulators are cleared.

2. A new dump trigger comes along with either a '01' or '11' code. Correlator accumulators are cleared. Start filling 2nd buffer memory. This trigger says what to do with the data that will subsequently be burst through the correlator from the first buffer.

3. Subsequent dump triggers alternate the filling of RMEM buffers 1 or 2 and burst correlation out of the alternate buffer. Unlike non-recirculation operation, the dump trigger indicates what to do with the data that is now sitting in a buffer waiting to be correlated rather than what to do with the data sitting in a correlator chip (as is the case for non-recirculation).

Note that there will also be interaction with phase binning. The 10-bit bin number will indicate which phase bin the data will go into. Also, if some synch-loss occurs (e.g. cable disconnect) then resynchronization will be established on the next DUMPTRIG even though some of the existing data will be bad. The main purpose for step 1. is to meet the requirement for "correlation on demand" with no loss of data.

Correlator Chip Data Storage in the LTA

It is planned to design the Baseline Board so that all dumping and timestamping of data is driven by signals coming from the Station Board (TIMECODE, DUMPTRIG). The data will be read out of the correlator chips and stored in (effectively dual-port) RAM using a

[20] A more tightly-coupled mechanism could be used where extra fields in DUMPTRIG tell the receiver FPGA on the Baseline Board *exactly* what to do.

relatively simple but very high performance controller implemented in an FPGA. A large RAM buffer is required so it is proposed to use Synchronous DRAM (SDRAM) SIMMs[21] that find use in mass-produced PCs. A few of these SIMMs, for a few hundred dollars, should provide at least 512 Mword capacity (8 Mwords per correlator chip). This arrangement will eliminate the need for a microprocessor to read out correlator chip data but more importantly should enable very high performance operation particularly when many phase bins and very high dump rates are needed.

It is useful to consider a simple mechanism that would allow the readout controller to read correlator chip data and store it in appropriate memory locations with support for phase binning and recirculation. The proposed mechanism is to write the row or column address (e.g. row address if 'X' and column address if 'Y') into the correlator chips along its row or column where data should be stored in the LTA. When the controller reads out the data, it uses the row and column address to determine which memory page in the LTA to store the data. The final address will be a function of the row and column address, the baseband correlation, the phase bin, the recirculation offset, and the current location of available buffer space in the memory page allocated to the particular correlator chip. Phase bin, recirculation offset, TIMECODE, and TIMECODE offset would also be written into the correlator chip so that the readout controller has all of the information necessary to store the data. Additionally, for simplicity it is probably best to always treat the correlator chip as having 16 individual correlated outputs that get finally rearranged or chained together to form the final lag results by a back-end processor.

4.4.6 DELAYMOD Format

The DELAYMOD signal contains quantized real-time delay models for each of the 8 basebands. For the EVLA, these are used as addresses into a lookup table that generates the phase offset needed for very fine delay tracking (see the Appendix). The actual lookup table coefficients are a function of the delay, the sub-band bandwidth, and sub-band slot. The lookup table will be accessed by the station receiver FPGA on the Baseline Board and will nominally be a 2k x 8-bit RAM—8, 8-bit address pages with a final 8-bit phase output. The 8-bit phase output of the RAM gets added to the phase accumulator for the particular baseband before being finally quantized to 4 bits. The proposed DELAYMOD format is shown in Figure 4-23.

[21] With interleaved accesses from the correlator chip side and the SCSI interface side so that they effectively look like dual-port memories—although semaphores will have to be used to provide synchronization of data.

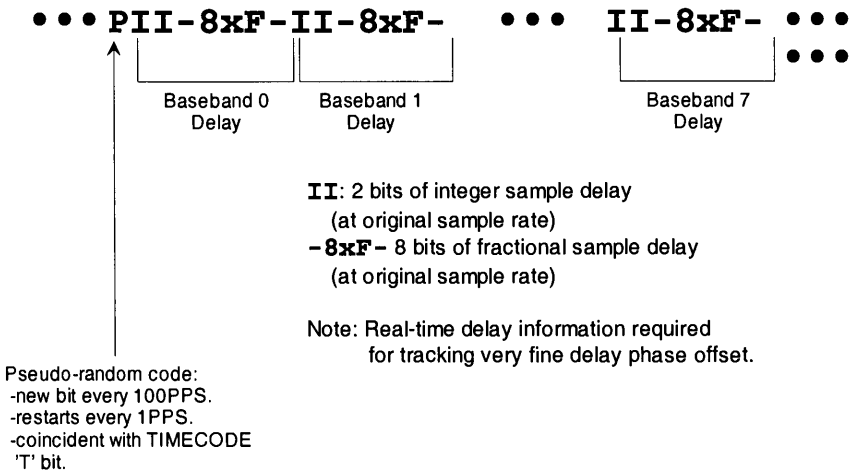


Figure 4-23 Proposed DELAYMOD format. The signal contains real-time delay models for all 8 basebands. There is a synchronization bit every 100PPS that is synchronized to TIMECODE so that proper framing can be done. The 8 bits of fractional sample delay is for very fine delay in the EVLA. The additional 2 bits of integer sample delay is for VLBI.

Support for VLBI

DELAYMOD also contains 2 bits of integer sample delay that are needed if the Baseline Board is to be used for VLBI[22]. Integer and fractional sample delay information is required to determine the baseline fine (a.k.a. 'Vernier') delay and phase modifier. This operation is described in detail in [1]. If the Baseline Board is to be used for VLBI, then DELAYMOD must be made available to the correlator chip—which is the only place that the baseline-based operations can be performed.

4.4.7 PHASEMOD Format

The PHASEMOD signal contains the point-slope phase models for every sub-band of every baseband. The station receiver FPGA on the Baseline Board captures the appropriate models from this stream and, when it receives a trigger, loads it into the appropriate phase generators (frequency synthesizers). Note that PHASEMOD (similar to DUMPTRIG) is not synchronized to TIMECODE. This allows phase models for each sub-band of each baseband of each antenna to be updated at an appropriate rate. For example, antennas near the phase center of the (sub-) array do not need to be updated as often as antennas far from the phase center. This mechanism has the necessary, but inconsequential, restriction that phase models for the same antenna/station must be updated at harmonically related rates. Since PHASEMOD is not synchronized to TIMECODE, it will be necessary to run tests with it synchronized to TIMECODE so that timing skews in the system can be removed or compensated for[23]. A proposed format for

[22] VLBI on data already recorded as sub-bands. If wideband demultiplexed data is recorded, then the operation is the same as the EVLA.

[23] Or, one could just define the "Start bit" to be synchronous with the "T bit" in TIMECODE (with some loss of generality) and be done with it.

PHASEMOD is shown in Figure 4-24. PHASEMOD also contains an 8-bit antenna/station identifier that should be written to the correlator chips in the station receiver FPGA's row or column. This ID serves to unambiguously identify the antenna (or antenna site) that the data came from.

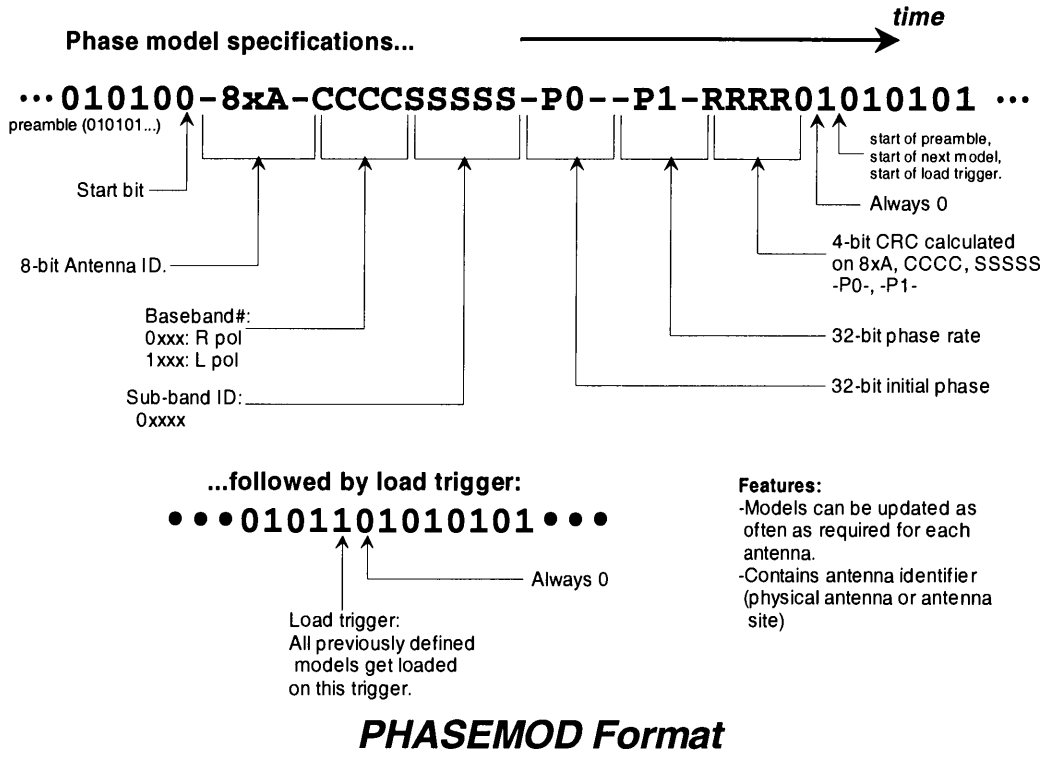


Figure 4-24 Proposed PHASEMOD format. This signal contains the phase generator coefficients for all basebands and sub-bands for the station. It also contains an Antenna ID that uniquely identifies the source of the data (i.e. in the cable the PHASEMOD signal is part of).

4.5 Power System, Hot Swapping, and Remote Monitoring

As currently understood, there are reasonably stringent reliability and maintenance requirements for the correlator. There is no *complete* 'hard' specification on what this means so in an attempt to quantify this, the following requirements for the correlator are defined.

- Be able to determine to the board or cable level where a fault is. Most of this can be performed on-line while observing, but some of it cannot. For example, it will be impossible to determine if there is a lag failure in a correlator chip while observing, but it can be determined in a short off-line test run, say, once per day.

- Hot-swappable circuit boards. The correlator system should be able to tolerate hot replacement of malfunctioning boards with only a small loss of data. The implication here is that the MTTR (Mean Time To Repair) should be less than ~5 minutes (provided someone is on site).

- Reliable operation. The VLA site is at a remote location and will be operated remotely a large fraction of the time. This implies a reasonably high MTBF (Mean Time Between Failures). An MTBF that is reasonable to aim for is 6 months for a fully debugged and burned-in system. It may be impossible to calculate and/or determine the MTBF but this serves as a guide for the specification of connectors[24], power supplies, and power supply derating.

- The health of circuit boards must be monitored remotely. Typically this means being able to monitor temperature and voltage(s) at several locations on the board via the standard computer interface. Monitoring of functionality will be done both on and off-line as appropriate. Temperature and voltage monitoring allows early determination of impending power supply or cooling system (i.e. fans) failure. For this monitoring to be useful, it must be possible to remotely power off and power on each circuit board (Station Board, Baseline Board, Phasing Board).

- Health of PCs should also be remotely monitored. There are some commercial systems available that can provide this capability. Most likely, industrial-strength rack-mount PCs will be used for high reliability operation. These are about twice the cost of commercial grade units.

- There is no plan to provide fully redundant fault-tolerant operation. Doing so would double the cost of the system.

In addition, it is proposed to include the following fail-safe features into the correlator:

[24] E.g. If a certain device is used, then the calculated MTBF for all of those particular devices in the system must be >> than 6 months if the entire system is to have an MTBF of 6 months.

- Per-board (Station Board, Baseline Board, Phasing Board) dead-man temperature protection. Simple electro-mechanical thermostats[25] will disconnect power to the board (or inhibit the power supply) when over-temperature conditions occur. This provides a fail-safe mechanism to protect the board in case of monitoring/control computer failure. This should be done on a per-board basis to minimize the loss of data.

- Main power system shutdown. Power to the entire correlator will be shutdown in the event of an over-temperature condition, or manual intervention via a panic switch.

- Perhaps use smoke detectors to trigger main power system shutdown. This will help to prevent a catastrophic failure. Smoke detectors must be very reliable and not prone to false alarms.

4.5.1 Power System Design

The size and operating requirements of the correlator share many attributes that are similar to central-office telephony communications systems. Reliability, hot swapping, high MTBF, and low MTTR are all important for communications systems so that they can provide 100% availability to their customers[26]. Thus, it is proposed that the power distribution system for the correlator be modeled after those used in central office systems. This system has the following characteristics:

- 48V DC power distribution to all racks and boards. Each board has a high efficiency DC-DC converter to provide the board with its required voltages.

- Central 48V DC power source with battery back-up. Commercial systems, such as the Lucent Technologies "GALAXY Power System" 4812 provide this capability with up to 60 kW capacity (1200 A at 48V) in a single solution. This is more efficient than a conventional UPS which includes a DC-AC inverter as well as rectifier and filter. The GALAXY system is designed for high reliability and on-line maintenance.

- 48V DC bus-bars can be distributed throughout the system without high voltage safety concerns.

A correlator powered from a 48V DC system would have the following attributes:

- Each circuit board has its own compact DC-DC converter[27]. Failure of a power supply means failure of only one circuit board rather than many if a sub-rack

[25] Available with various fixed temperature limits in TO-220 packages.
[26] Communications systems provide redundant fault-tolerant operation which, as mentioned, would be expensive to do with the EVLA correlator.
[27] DC-DC converters are pretty much inexpensive commodity items with many manufacturers offering many models.

power supply were used. The failed circuit board could be quickly hot-swapped with a new circuit board thus minimizing down time of the telescope.

- The number of power connections is minimized. Only one blind-mate power connection is made when the board is plugged into the sub-rack. There are no power cords, fuse holders, plugs and sockets, high voltage power switches or other failure points that could fail and bring down the system.

- Telephony central-office power reliability. Power products designed for 48V systems are designed for reliability.

- Industrial PCs come with optional 48V DC power inputs. These can plug directly into the 48V DC power grid and thus provide reliable PC operation. If cost is prohibitive[28], then commercial-grade PCs could plug into an AC power grid generated with a conventional UPS.

Perhaps the most important motivation for using a 48V DC system is the immunity of the system to single power supply failures and the low MTTR provided by hot-swap capability. A single power supply failure only causes one board to go down and hot-swap capability makes it easy to replace the board and return the system to full on-line status.

An additional factor that must be considered is the capability of providing total power supply shutdown in the event of a catastrophic failure. This requires that 48V DC be interrupted to all racks. Since the output of the GALAXY Power System goes to a bank of batteries for UPS capability, a large switch is required on the output of the battery bank. The capability of providing this (or similar functionality) must be seriously considered.

4.5.2 Power Requirements Estimate

The following is an estimate of the power requirements of the correlator system. This assumes that the Station Boards are located at a central site with the rest of the correlator and that a 48V DC power distribution system is used.

- Each Station Board and Phasing Board uses ~120W. 4x40=160 Station Boards and 16 Phasing Boards for a total power requirement of 21.1 kW. DC-DC efficiency is minimum 80%, for a total power requirement at 48V DC of **26 kW**.

- Each Baseline Board uses ~150W (1.5Wx64=96W for correlator chips). There are 15x16=240 boards for a total of 36 kW. With DC-DC efficiency of 80%, the total power requirement at 48V DC is **45 kW**.

- Each PC uses a maximum 300W (including power supply efficiency losses). If there are 10 PCs for Station Boards (one PC for every 4 antennas), and 16 PCs for

[28] Or, if otherwise desired for a back-end supercomputer where on-line reliability is not of utmost importance.

Baseline Boards (one PC for every sub-band for ~40 Mbytes/sec/sub-band dump rate), and 1 PC for the phasing boards, then the total PC power requirement is **8 kW**.

Thus, if the PCs use 48V DC power as well, the total 48V DC power requirement is 80 kW. If the GALAXY Power System DC power supply is 84% efficient, then the total mains-AC power requirement for the correlator is **95 kW**. The GALAXY data sheet does not say that it is power-factor corrected which means (since it uses switching power supplies), that it has a power factor of about 0.7. This means that the mains-AC must supply 135 kVA to the correlator. Also, two GALAXY systems would be required to provide the 80 kW required by the correlator (each GALAXY system can deliver 1200A at 48V).

4.5.3 Hot Swapping

Assuming that a 48V DC mains power distribution system is used, hot swapping of individual circuit boards (Station, Baseline, and Phasing Boards) is envisioned to occur as follows:

- A front panel lighted power switch is used to enable or disable the on-board DC-DC converter. This switch does not switch any current but rather inhibits the power supply via a 'logic disable' line. This same line will be used for over-temperature board shutdown and probably remote power control.

- To remove a board, turn off the power switch, disconnect the SCSI cable, and remove it.

- To replace a board, plug in the new board, plug in the SCSI cable, and turn on the power switch. Once configured, the board will quickly and automatically synchronize to the embedded synchronization pulses (i.e. in TIMECODE, DUMPTRIG etc.) and be on-line and generating data.

An alternate method for hot-swap power delivery is to use power pins that ensure that make/break contact with power is done before signals are made/broken. This method may suffer from rapid power cycling when pins first make contact and there may be some difficult interactions with the SCSI cable (i.e. is the cable would have to be removed when power is applied and this could cause some interruption on the SCSI bus).

The SCSI controller in the PC and driver software must be able to tolerate interruption of service to devices on the SCSI bus. The software should be designed so that when service is interrupted, a reconfiguration occurs when the PC senses that that SCSI device is back on-line.

4.5.4 Remote Monitoring

The correlator is going to be operated remotely (presumably) from the AOC in Socorro. This requires that automatic monitoring of the health of the system be performed with the

additional capability of allowing someone at the AOC to check it as well. PCs connected to Station, Baseline, and Phasing Boards will automatically monitor temperature and voltage and shutdown boards[29] if the measured quantities are out of range. These same PCs will act as gateways for remote checking. Additionally, the health of the main (e.g. GALAXY) power supply must be remotely monitored.

4.5.5 Remote Operation

Presumably, a skeleton operations staff will be at the VLA site to change archive tapes, change VLBI tapes, and perform simple operations on the correlator system on request from the AOC. If no staff is at the VLA, then obviously tape devices must be at the AOC and data must be shipped from the PCs via high-speed fiber to the AOC. For VLA archive data, this can reasonably easily be done with the addition of a fiber-optic transceiver board plugged into each PC that could then be connected to the fiber system. For VLBI data, this would require the addition of a fiber-optic transceiver on the Phasing Board.

[29] The mechanism for shutting down boards needs to be investigated. Probably, it will use the inhibit of the on-board power supply controlled, perhaps by a central PC so that each board can be individually remotely power controlled.

5 Preliminary Cost Estimate

This section provides a preliminary estimate of the cost of a 40-station WIDAR correlator system for the EVLA described in the previous sections. The cost of the high-speed 4-bit samplers, the FOTS, the mains (e.g. GALAXY) power system, air-conditioners, and office space is not included in this estimate. All figures are in U.S. dollars. Items outlined in bold are based on actual quotations. Other items are based on past experience, current industry trends, and information from the U.S. NRAO ALMA correlator group. NRE is "Non-Recurring Engineering" charge.

1. FIR filters. NRE~$100k; $100/chip; 40x4x16x2=5120 chips. TOTAL: **$600k**

2. **Station-Baseline cable**: 6 m GORE cable (16x40x$454)=$290k; 1.5 m GORE cable (40x5x16x$263)=$841k. TOTAL: **$1150k**

3. Correlator chips: NRE: $500k; 15x64x16 chips x $100/chip= $1536k. TOTAL: **$2036k**

4. Misc Station Brd logic + cct boards: $10k/boardx40x4 boards=$1600k. TOTAL: **$1600k**

5. Misc Baseline Brd logic + cct boards: $10k/boardx15x16 boards=$2400k. TOTAL: **$2400k**

6. Phasing subsystem. Cable: 2x40x$454=$32k; 8 boardsx$20k/board=$160k; Cable: 8x40x$200=64k. TOTAL: **$256k**

7. **Computers (Dual-PIII 550MHz CPU Industrial-grade PCs)**. Station: 10; Baseline: 2x16=32; @$5k each: TOTAL: **$210k**. (does not include archive tape drive costs)

8. Labour: 5 eng x 4 yrs x $70k/year[30]. TOTAL: **$1400k**

TOTAL: Parts: $8.3 million (U.S.) + 15% contingency = $9.5 million
Labour: $1.4 million (U.S.) + 20% contingency = $1.7 million

It should be noted that there is some contingency already built into the parts costs because, for example, there is some indication that the correlator chips will cost <$100 each.

[30] Based on engineers working on-site in Canada.

Based on these estimates, a total cost curve was developed for this design showing the potential costs for up to a 100 station correlator for N=16 (256 MHz clock rate) and N=32 (128 MHz clock rate). This cost curve is shown in Figure 5-1.

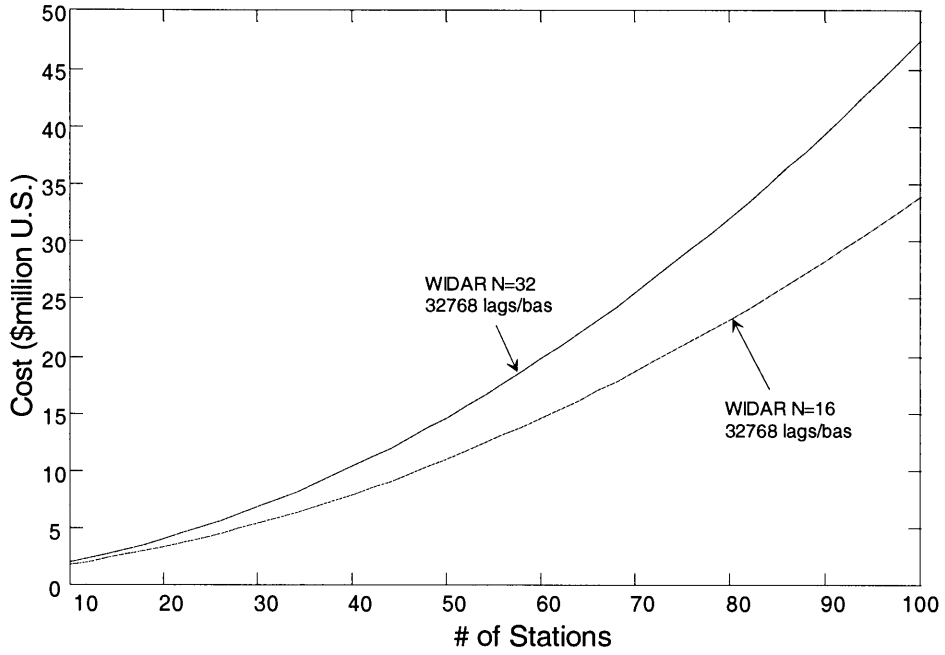


Figure 5-1 Total cost curve for an N=16 and N=32 WIDAR correlator for up to 100 stations/antennas. This cost includes the phasing subsystem and NRE costs but does not include labour or explicit contingencies. An N=32 correlator (for the same spectral resolution) is more expensive because of the additional Baseline Boards and overhead logic.

6 Conclusions

This document has discussed in some detail a proposed implementation of the EVLA correlator based on WIDAR signal processing. Observing Modes and tables were defined that demonstrate that the correlator offers a wide range of flexibility with the capability of efficiently trading off bandwidth for spectral resolution. In wideband modes, it can provide 1024 spectral channels across every 2 GHz baseband[31] for all baseband pairs and all polarization products. As the number of baseband pairs is decreased, the number of spectral channels per 2 GHz baseband increases. With a single 2 GHz baseband, it is possible to get 16384 spectral channels on every baseline. Each antenna can have an independent delay/phase center on the sky for each baseband. The WIDAR technique is able to accomplish fully digital, precision $1/16^{th}$ sample delay tracking. This eliminates the need for modifying the phase of the sampler clocks. The correlator uses reasonably large digital filters to decrease the processed bandwidth for targeted spectral-line observations. Many filters allow many narrow bands to be correlated—the placement of the bands being very flexible across many GHz of bandwidth with multiple independent LOs. With narrow bands, and data recirculation methods, it is possible to get up to a ¼ *million* spectral channels per baseline for extremely high spectral resolution experiments. A full 4-bit correlator is proposed that, along with WIDAR anti-aliasing methods, aims to provide $>10^5$ spectral dynamic range even in the presence of extremely powerful narrowband interference. This, coupled with high spectral resolution, will permit precision post-correlation, narrowband interference mitigation. The WIDAR architecture is also well suited to adaptive interference cancellation techniques—should such techniques ultimately prove to be feasible. A flexible fully digital phased-VLA system is proposed that can operate concurrently with VLA interferometer mode. This system does not require a D/A and A/D step thus maintaining maximum efficiency and predictability.

A reasonably detailed hardware architecture of the correlator and phasing system was presented with discussion describing data paths, signal processing, and potential hardware implementations. This architecture is not a final design but provides a reasonably solid and detailed framework from which the final architecture can be built. A proposed correlator installation scheme was presented and discussed. This scheme aims to provide reliable and remote operation with the goal of minimizing correlator, and thus telescope downtime at reasonable cost. The total power requirements of the correlator were estimated to be about 100 kW. Finally, an initial cost estimate indicates that the correlator (not including the 4-bit samplers) could be constructed for ~$11 million U.S. including about 20% contingency.

[31] A baseband is defined as the output of a single 4 Gs/s sampler—normally with 2 GHz of bandwidth. A "baseband pair" is 2 sampler outputs—one for each polarization.

7 References

[1] Carlson, B.R., Dewdney, P.E., Burgess, T.A., Casorso, R.V., Petrachenko, W.T., Cannon, W.H. The S2 VLBI Correlator: A Correlator for Space VLBI and Geodetic Signal Processing, Publications of the Astronomical Society of the Pacific, 1999, 111, 1025-1047.

[2] Carlson, B.R., Dewdney, P.E. Efficient wideband digital correlation, Electronics Letters, IEE, 2000, *In prep.*

[3] Thompson, A.R., Moran, J.M., & Swenson, G.W. 1986, Interferometry and Synthesis in Radio Astronomy, Wiley, New York.

[4] SETI Institute: http://www.seti-inst.edu/science/1ht.html

[5] S.W. Ellingson, J.D. Bunton, J.F. Bell, "Cancellation of GLONASS signals from radio astronomy data", Astronomical Telescopes and Instrumentation 2000—Radio Telescopes, SPIE conference 4015, Munich, March 2000.

8 Appendix I – WIDAR Signal Processing

This appendix will describe the fundamentals of WIDAR (Wideband Interferometric Digital ARchitecture) signal processing. A good understanding of this processing is required to fully understand the flexibility, power, and limitations of the proposed correlator system.

8.1 Basic WIDAR Technique

In [2], the fundamentals of WIDAR signal processing are described. In that paper, the following block diagram is presented and includes all of the critical signal processing elements:

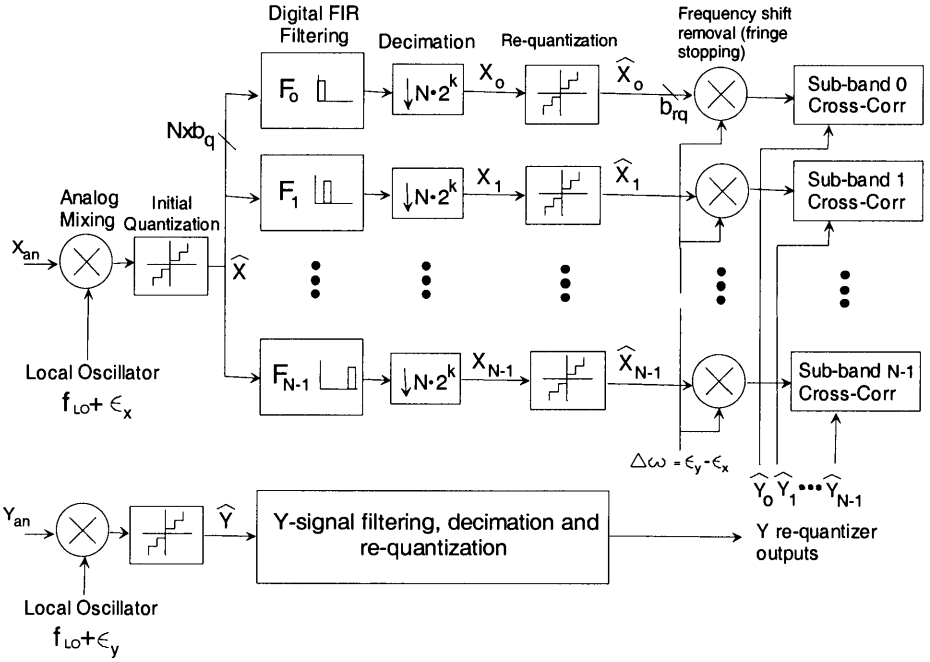


Figure 8-1 Basic WIDAR signal processing block diagram. There is a different Local Oscillator offset in each antenna. This offset is removed in the correlator, causing aliasing to decorrelate and allowing digital filter generated sub-bands to be seamlessly stitched together to produce the wideband result. It is important to note that the mixers in the correlator are simple digital multipliers. In an XF correlator, the mixers use either 3-level, or as proposed for the EVLA correlator, 5-level approximations to sine and cosine functions—a standard technique in VLBI.

The fundamental problem in correlating high-speed sampled data is that the data is running at a much higher rate than downstream correlator hardware is capable of handling. Few-bit samplers inherently run at very high rates because they consist of relatively simple electronics that, with care, can be made to run very fast. Conversely, downstream digital correlation hardware is inherently relatively slow due to its increased complexity.

Correlation of high-speed sampled data can be managed by using a sub(sample)-rate demultiplexing scheme. One scheme is to demultiplex the data in time "bursts" (chunks of contiguous-time sampled data) and correlate each time burst in a separate correlator. Final results are obtained by adding all of the individual correlator results together to produce the final result. This scheme works well, but is inefficient when it is desired to produce many spectral points across the wideband since many correlations[32] must be done to produce each spectral point. Another scheme is to demultiplex the data by filtering into sub-bands and correlating each sub-band in a separate correlator. This scheme is efficient in that only one correlation must be performed for each spectral output point. A complication with this scheme is that sub-band overlaps and aliasing must be properly handled, and if analog filters are used, the filters suffer from time and temperature variability.

WIDAR uses the frequency/filter demultiplexing scheme and solves the sub-band overlap/aliasing and analog filter instability problems in novel ways. Looking at Figure 8-1, it is desired to cross-correlate the data from two antennas 'X' and 'Y'. With WIDAR, a slightly different Local Oscillator is used in both antennas. This offset introduces a relative frequency shift in the baseband signal that is different for each antenna. The wideband signal is sampled at the high-speed rate and then time demultiplexed (not a burst demultiplex) and fed into N digital FIR (Finite Impulse Response) filters. Each FIR filter produces a sub-band output that, when decimated by N, is 1/N as wide as the original sampled data. The FIR filter bandpass must also reside in an integer 'slot' on n/N boundaries. The data out of each FIR filter is requantized and N correlators are used to correlate each sub-band and remove the different frequency shifts. Because the frequency shift is removed after filtering, any aliasing of the filter transition band that crosses the 'slot'/sub-band boundary decorrelates and does not contribute to the sub-band cross-power spectral amplitude. Also, digital FIR filters are stable, perform consistently, have calculable amplitude responses, and have linear-phase (which, for this application means zero phase) characteristics within the bandpass of interest.

The process of frequency shifting, filtering, and correlation is outlined in steps in the following Figures. In Figure 8-2, the wideband signal with a frequency shift in it is sampled at the high-speed rate. This signal is then bandpass filtered with the digital FIR filter to yield the filtered signal which, when decimated becomes a baseband signal at the lower sample rate (Figure 8-3). The arrows in Figure 8-3 show the frequency sense of the main part of the band and the aliased (dotted lines) and non-aliased (solid lines) transition bands. The different frequency shifts are removed with the fringe stoppers in the correlator—resulting in the decorrelation of the aliased transistion bands. (Indeed, any signal in the band with an incorrect frequency shift—such as a residual unwanted sideband—will decorrelate.) The final integrated cross-power spectrum of the signal in the example is shown in Figure 8-4. The aliased transition bands decorrelate and all that is left is the desired sub-band spectrum. The overlap of the transition bands outside of the strict sub-band boundaries correlates and allows adjacent sub-bands to be seamlessly

[32] If the demux factor is N, then N correlations must be performed.

stitched together (i.e. a signal right on the boundary will not suffer funny edge effects so data right up to and including the sub-band boundary can be used).

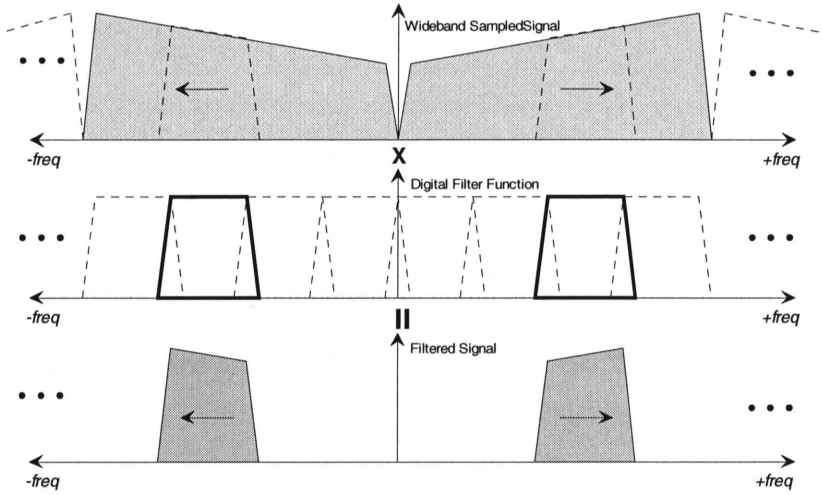


Figure 8-2 Basic sub-band filtering operation. The wideband sampled signal is shown in the top of the figure. The slope in the bandpass is just used to indicate frequency sense. The arrows indicate the direction of the frequency shift from the Local Oscillator which is different in each antenna. The middle of the figure shows the digital filter function in the frequency domain. The bold is the bandpass that's going to be used for this sub-band and there are N=4 possible sub-bands, the outline of the others shown in dotted lines. The bottom of the figure shows the filtered sub-band before decimation—a signal that is not actually available anywhere since filtering and decimation occurs in one step for FIR filter implementation simplicity.

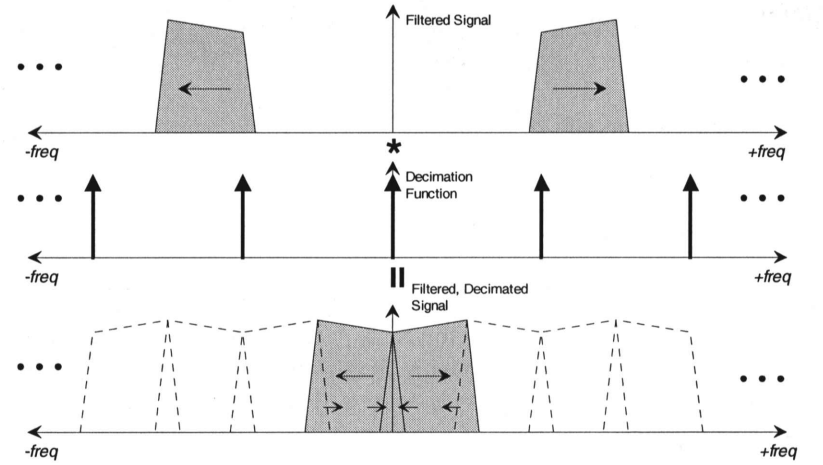


Figure 8-3 Sub-band decimation. The bandpass filtered signal is mixed to the new baseband with the decimation/sub-sampling function. In this case, since the bandpass is on an even frequency slot (slot 2 of 4—counting from 0), the signal shows up with the same frequency sense. If the bandpass was on an odd slot, then the frequency sense would be flipped and can be corrected by changing the sign of every other sample. The arrows show the frequency sense of the main part of the passband and the different aliased and non-aliased transistion bands.

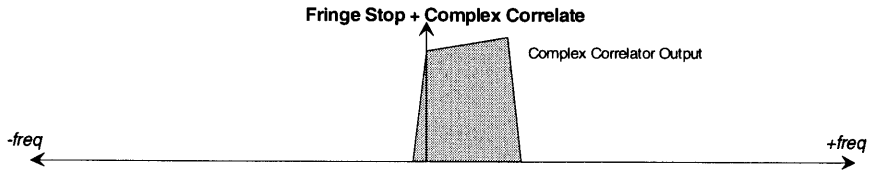


Figure 8-4 The final integrated cross-power spectrum of the sub-band from the previous figures. The aliased transition bands decorrelate and all that is left is the desired sub-band. Note that in this example, the transistion bands overlap into frequencies outside the strict sub-band boundaries. This overlap correlates, suffers $\sqrt{2}$ SNR degradation, but provides the required information to allow adjacent sub-bands to be seamlessly stitched together.

A key measurement that allows any arbitrary wideband spectrum to be re-assembled from the individual sub-band cross-correlations is the measurement of the total power of the sub-band filtered signal after FIR filtering, but *before* re-quantization (X_n, Y_n in Figure 8-1). Since the normalized correlation coefficients produced by the sub-band cross-correlator will cause total power[33] of the sub-band to be lost, measuring it before re-quantization allows the level of a particular sub-band cross-power spectrum to be established relative to all other sub-bands. The measurement is a simple lag 0 autocorrelation but must be performed with all (or nearly all—a number that needs to be determined) of the bits per sample generated by the FIR filter. Additionally, in order for these measurements to be meaningful, sub-band filter coefficients must be normalized to a common coefficient before they are scaled to real integer quantities to be used in the real FIR filter. To maintain maximum dynamic range in each FIR (because internally each FIR uses a finite number of bits), and simplify the system[34] this should be done in two steps: 1. normalize each FIR tap coefficient to its maximum tap coefficient before scaling to real hardware FIR integers; 2. use the ratio of the maximum tap coefficient of each FIR to some system-wide coefficient—such as the maximum coefficient for 1/16 bandpass—to properly scale a particular FIR's output total power measurement.

8.2 Sub-band Cross-power Amplitude Corrections

From [2], and assuming proper FIR coefficient normalization as described above, the general equation that is used to correct sub-band correlator spectral amplitudes so that the seamless wideband spectrum can be obtained is:

$$\rho_{nf} = \rho_{onf} \cdot \frac{P_{XY_n}}{P_{XY_T}} \cdot P_{XYFIR_T} \cdot \frac{S_{XYFIRideal_{nf}}}{S_{XYFIR_{nf}}} \cdot P_{XYideal_T}$$

Where the terms in the equation are defined as follows:

- ρ_{nf} A final corrected cross-power spectral point at frequency bin f of the n^{th} sub-band from antennas X and Y.

[33] The 'real' total power is lost after initial quantization. However, if we're interested in producing the same results as a perfect wideband correlator, then individual sub-band total power is of interest.
[34] Since there are many FIRs of different bandwidths on different antennas, that could be changing on-the-fly at different times etc. etc.

- ρ_{onf} A normalized sub-band correlator spectral point. Normalized sub-band correlator data is produced using standard normalization techniques [3] that use the raw sub-band cross-power data and statistics of the sub-band X and Y re-quantizer outputs.

- $\frac{P_{XY_n}}{P_{XY_T}}$ This is the **sub-band scaling term** and scales each sub-band correlator output to its correct amplitude relative to the gain of the sub-band filter[35], the other sub-bands and also relative to the initial quantizer output[36].
  $P_{XY_n} = \sqrt{\langle x_n^2 \rangle \langle y_n^2 \rangle}$ is the geometric mean of the n^{th} sub-band total power measurement after FIR filtering from the X and Y antennas. P_{XY_T} is the 'average' of the total power measured out of 'all' X and Y antenna sub-band FIR filters. Just exactly what 'average' and 'all' means here is defined in a following subsection.

- P_{XYFIR_T} This is the **gain differential compensation term** and removes any total gain differences in sub-band FIR filters inherently present in the denominator of the sub-band scaling term. This term is a bit of a slippery character to understand. Basically, the measurement of $\langle x_n^2 \rangle$, $\langle y_n^2 \rangle$, and subsequent calculation of P_{XY_T} inherently factors in calculable, but non-ideal behaviour/gain of each FIR filter (i.e. FIRs have finite transition bands rather than infinitely steep transistion bands) that generally will be different for each sub-band. To remove the effects of these non-ideal gains, they are calculated and applied here to cancel out the effect in P_{XY_T}. Note that P_{XYFIR_T} must be calculated in the same manner as P_{XY_T} so that the compensation is correct.

- $\frac{S_{XYFIRideal_{nf}}}{S_{XYFIR_{nf}}}$ This is the **bandshape correction term** and corrects for any sub-band FIR filter shape that is not the same as the ideal filter it is designed to implement. $S_{XYFIRideal_{nf}}$ is the ideal cross-power filter shape that we want for the sub-band. Nominally, this is simply a "box-car" function, but it can be virtually any shape if it is desired to perform some gain equalization across the sub-band before correlation. $S_{XYFIR_{nf}}$ is the actual cross-power response of the implemented FIRs. This can normally be calculated by taking the square of the Fourier

[35] For example, increased filter gain due to a wider transition band causes ρ_{onf} to drop—by multiplying by the gain of the filter, the effect of this drop is cancelled.
[36] That is, so that the reassembled wideband spectrum is the same as that produced by a perfect wideband correlator.

transform of the tap coefficients (zero-padding if more frequency points[37] are required to match the number of points in the cross-correlator output) before any integer scaling is done for hardware.

- $P_{XYideal_T}$ This is the **wideband power gain term** and simply scales the cross-power output to the total gain of the wideband gain equalization function. If the wideband gain equalization function is a box-car with a gain of 1, then this is simply 1. However, if the function is some arbitrary shape to compensate for gain losses in the analog system, then this restores the output amplitude to what it would have been if the analog system was perfect.

Figure 8-5 shows the cross-power spectra of 16 individual sub-band correlations before the sub-band correction equation is applied. Figure 8-6 plots the sub-band scaling term for each sub-band. Figure 8-7 shows the same data (along with a reference wideband correlator result) after the equation has been applied with arithmetic mean total power normalization (see later in this section).

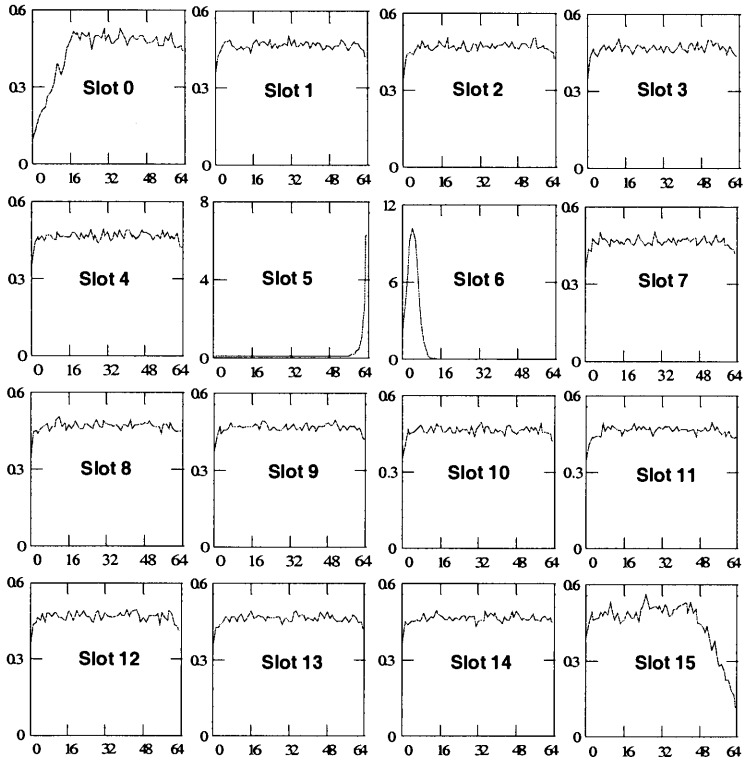


Figure 8-5 Example plots of individual sub-band spectra. In each plot, the vertical axis is correlation coefficient amplitude, and the horizontal axis is frequency bin number. These are individual sub-band correlator results before applying the sub-band correction terms. Note the amplitude scales of the individual sub-bands. Also, the transition band roll-off is visible in each case.

[37] Curiously, if the FFT of the tap coefficients yields more frequency points than the sub-band cross-power spectrum provides, simple truncation of tap coefficients and FFT did not yield the correct amplitude response. This point will require further investigation.

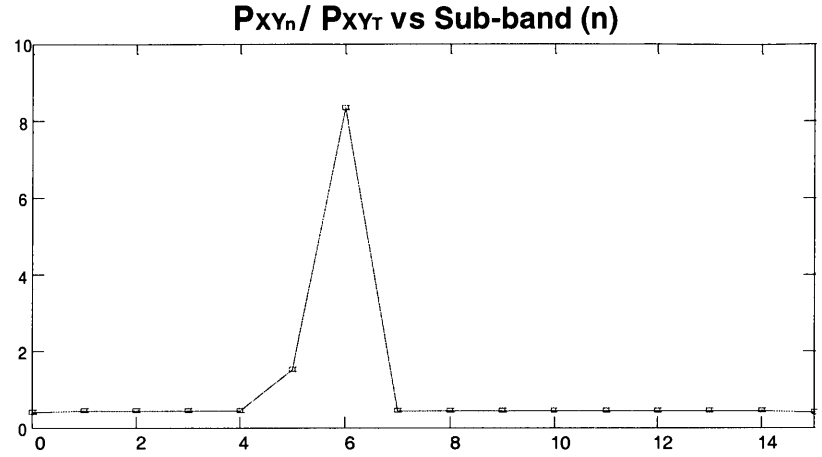


Figure 8-6 Plot of the sub-band scaling term for every sub-band for the correlations shown in Figure 8-5. P_{XYT} arithmetic mean total power normalization is used and, because of the powerful narrowband line, continuum sub-bands are depressed by this term and sub-bands containing the narrowband signal are amplified (i.e. depressed and amplified compared to the amplitudes shown in the previous figure).

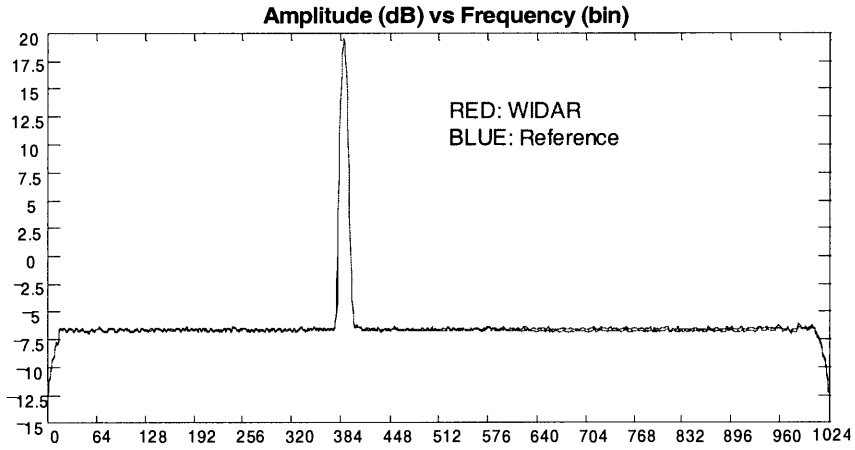


Figure 8-7 Example corrected and stitched-together WIDAR correlator data (from Figure 8-5) and reference (ideal wideband) correlator data. Sub-bands can be seamlessly stitched together once the WIDAR correction equation has been applied to each sub-band. In this case, the arithmetic mean total power normalization is used.

8.3 Power Normalization Options (P_{XYT})

In the previous section it was mentioned that there may be different ways in which the denominator of the sub-band scaling term (P_{XY_T}) could be calculated. Each method yields a different result if there are narrowband signals within the wide band.

8.3.1 Total Power (Arithmetic Mean)

This method yields the same result as normalized correlation coefficients from a wideband correlator. In this case, P_{XY_T} is the total power out of all of the sub-band filters before requantization and is calculated according to the following equation:

$$P_{XY_T} = \frac{1}{N} \cdot \left[\sum_{n=0}^{N-1} P_{XY_n} \right] = \frac{1}{N} \cdot \left[\sum_{n=0}^{N-1} \sqrt{\langle x_n^2 \rangle \langle y_n^2 \rangle} \right]$$

Where N is the number of sub-bands. As mentioned, this does yield the same result as a "normal" wideband correlator, but the normalized correlator output will be modulated by varying narrowband signals that show up in the measurement of $\langle x_n^2 \rangle$ and/or $\langle y_n^2 \rangle$. If the main continuum interference mitigation technique is post-correlation excision of narrowband interference, then even though the interference has been excised, the modulating effect of the interference is still present. This effect will be present even if the narrowband interference does not correlate.

8.3.2 Total Power (Geometric Mean)

This method is similar to the above method except that the geometric mean calculation results in a P_{XY_T} that is not as affected by "outlyers"—sub-bands with strong narrowband interference in them—as in the arithmetic mean case. In this case, the calculation of P_{XY_T} is:

$$P_{XY_T} = \left[\prod_{n=0}^{N-1} P_{XY_n} \right]^{\frac{1}{N}} = \left[\prod_{n=0}^{N-1} \langle x_n^2 \rangle \langle y_n^2 \rangle \right]^{\frac{1}{2N}}$$

8.3.3 "Clean" Sub-band Power

In this method, all sub-bands are normalized to those one or more sub-bands that do not have any narrowband interference in them. This removes the modulating effect from varying narrowband interference in some parts of the band and can have a significant impact on the continuum dynamic range of the data. In this case the calculation of P_{XY_T} is:

$$P_{XY_T} = \frac{1}{K} \cdot \left[\sum_{\substack{n=0 \\ n \neq nbIF}}^{N-1} P_{XY_k} \right] = \frac{1}{K} \cdot \left[\sum_{\substack{n=0 \\ n \neq nbIF}}^{N-1} \sqrt{\langle x_n^2 \rangle \langle y_n^2 \rangle} \right]$$

Where K is the number of sub-bands that do not contain narrowband signals. An arithmetic mean calculation is shown but a geometric mean calculation could be used and would yield essentially the same results. Note that all that is required is for one sub-band

to be interference free. In fact, in the face of interference in all sub-bands, a special purpose filter could be used to find just a sliver of the wideband that is interference free and use its output in the calculation. Of course, as the sub-band gets narrower, P_{XY_T} will get noisier.

What follows is a short discussion of how this normalization option works.

"Clean" Sub-band Power Normalization Discussion

Fullband/Wideband Correlator

From [3], the normalized correlation coefficient of a sampled but unquantized signal is:

$$\rho = \frac{\langle x \cdot y \rangle}{\sqrt{\langle x^2 \rangle \cdot \langle y^2 \rangle}}$$

If x and y are quantized in amplitude ($\hat{x}$, $\hat{y}$), then a suitable normalization algorithm will yield the above results, with some added noise. It is thus sufficient to look at the problem without considering quantization and the effect of an AGC (Automatic Gain Control) circuit that optimizes quantizer performance[38].

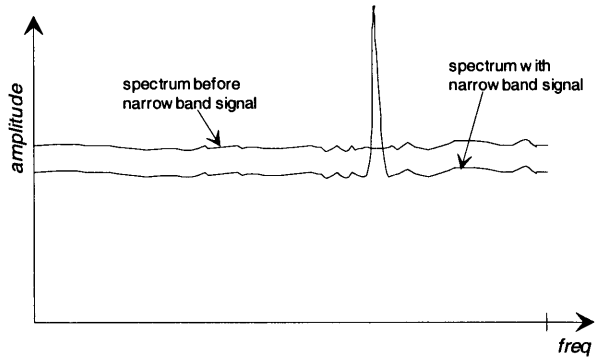
At any particular frequency bin f, the normalized correlation coefficient in *power/Hz* (or *power/total bandwidth*) is:

$$\rho_f = \frac{\langle x_f \cdot y_f \rangle}{\sqrt{\langle x^2 \rangle \cdot \langle y^2 \rangle}} \qquad (1)$$

In a wideband correlator for a pure continuum (flat-band) signal, ρ_f at all frequencies is the same and, if in terms of *power/total bandwidth*, then $\rho_f = \rho$ for all f, albeit with more noise. Note that x_f and y_f are not signals present anywhere in the system but that $\langle x_f \cdot y_f \rangle$ is a correlator output product transformed to the frequency domain.

Now, suppose a narrowband signal is added to some part of the band. The denominator in equation (1) increases since the total power has increased. At all frequencies except where the narrowband signal is, there has been *no change*. However, since the ratio of the correlated power to the total power has changed, ρ_f at every frequency bin in the band drops:

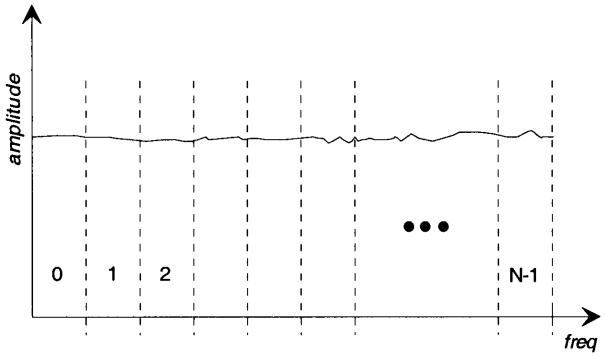
[38] It is possible to go through a simple "thought experiment" *with* amplitude quantization and an AGC and the results are the same. (Indeed, the effect was first noticed in simulation results with an AGC and with amplitude quantization.) In this case, adding a narrowband signal to some part of the band will cause the denominator of equation (1) to stay the same (because of the AGC), but the numerator in the equation (where there is no narrowband signal) will drop (because the AGC gain has dropped). This results in a ρ_f that is the same as without amplitude quantization and an AGC.



To remove this effect and restore the true level of the cross-power spectrum, it is necessary to determine the power of the signal that was added—possible but difficult to do in practice. Note that it is not necessary for the narrowband line to correlate since its power shows up in the lag 0 autocorrelation in the denominator of equation (1).

WIDAR Correlator

With WIDAR, the wideband is split into N sub-bands, each of which is separately correlated:



The normalized correlation coefficient for each sub-band *n* is:

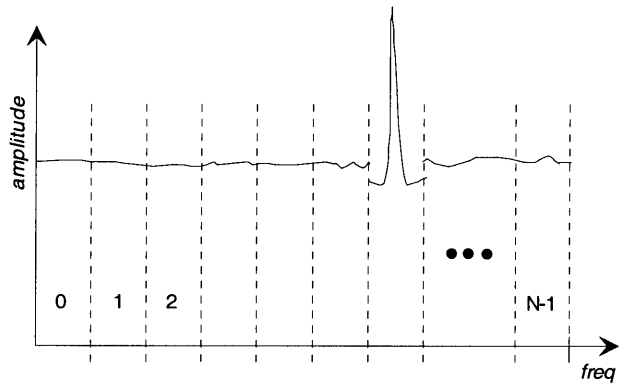
$$\rho_{onf} = \frac{\langle x_{nf} \cdot y_{nf} \rangle}{\sqrt{\langle x_n^2 \rangle \cdot \langle y_n^2 \rangle}} \qquad (2)$$

Similar to previous arguments this equation is for unquantized signals after filtering—if the signals are quantized, then a suitable normalization algorithm will produce equivalent results. With WIDAR, the total power in each sub-band after filtering (and if requantized, before requantization) is measured. These are the terms in the denominator of equation (2).

Once ρ_{onf} for each sub-band has been determined, it is normalized using the following equation (including just the **sub-band scaling term** and ignoring the calculable non-ideal filter effects that introduce the other terms):

$$\rho_{nf} = \rho_{onf} \cdot \frac{\sqrt{\langle x_n^2 \rangle \cdot \langle y_n^2 \rangle}}{\frac{1}{N} \cdot \sum_{n=0}^{N-1} \sqrt{\langle x_n^2 \rangle \cdot \langle y_n^2 \rangle}} \qquad (3)$$

Now, let's introduce the narrowband signal into sub-band 6:



According to equation (2), ρ_{onf} for n≠6 has not changed since $\langle x_n^2 \rangle$ and $\langle y_n^2 \rangle$ have not changed. However, the level in sub-band 6 *has* changed according to previous arguments for the wideband correlator (equation (1)).

If equation (3) is applied to the data, then the continuum level in all sub-bands drop, and the results are the same as the wideband correlator. However, if we modify equation (3) so that it does not include the sub-band total power from sub-band 6 in the denominator, then:

$$\rho_{nf} = \rho_{onf} \cdot \frac{\sqrt{\langle x_n^2 \rangle \cdot \langle y_n^2 \rangle}}{\frac{1}{N-1} \cdot \sum_{\substack{n=0 \\ n \neq 6}}^{N-1} \sqrt{\langle x_n^2 \rangle \cdot \langle y_n^2 \rangle}}, \qquad (4)$$

then sub-bands without the narrowband signal do not change in amplitude.

In sub-band 6, the decreased level due to equation (2) is precisely matched by the numerator of equation (4), and the correct level for the entire sub-band 6 is re-established and the same level as the other sub-bands.

This same effect occurs even if the signals are re-quantized after filtering—an assertion that has been demonstrated with simulation.

Thus, if the denominator in equation (3) does not include sub-bands with narrowband signals, then the correct levels for the entire spectrum can be obtained—making the WIDAR correlated output immune to the effects of time variable interference. It is possible to use only one sub-band (or a small part of one sub-band produced with a

separate filter) in the denominator of equation (4). However, as the number of sub-bands used decreases, the SNR drops—although for any given bandwidth used, SNR is maximized since the lag 0 autocorrelation is used in the equation.

As a final note, there is a 3rd order effect from time-variable interference that cannot be corrected with WIDAR. This is the changing quantization noise present from time variable interference which will affect correlated amplitudes. This affect is >~30dB (for 4-bit quantization) lower than the primary effect described in this discussion and is effectively negligible.

8.4 Aliased Transition Band Decorrelation

As previously mentioned with WIDAR, aliased transition bands decorrelate with integration time. This is the key feature that allows individually correlated sub-bands to be seamlessly stitched together to yield the wideband spectrum. Decorrelation requires that the Local Oscillators be offset differently in each antenna in the array so that frequency shift removal in the correlator has the desired effect. . This is illustrated with simple sine waves in the frequency domain in Figure 8-8.

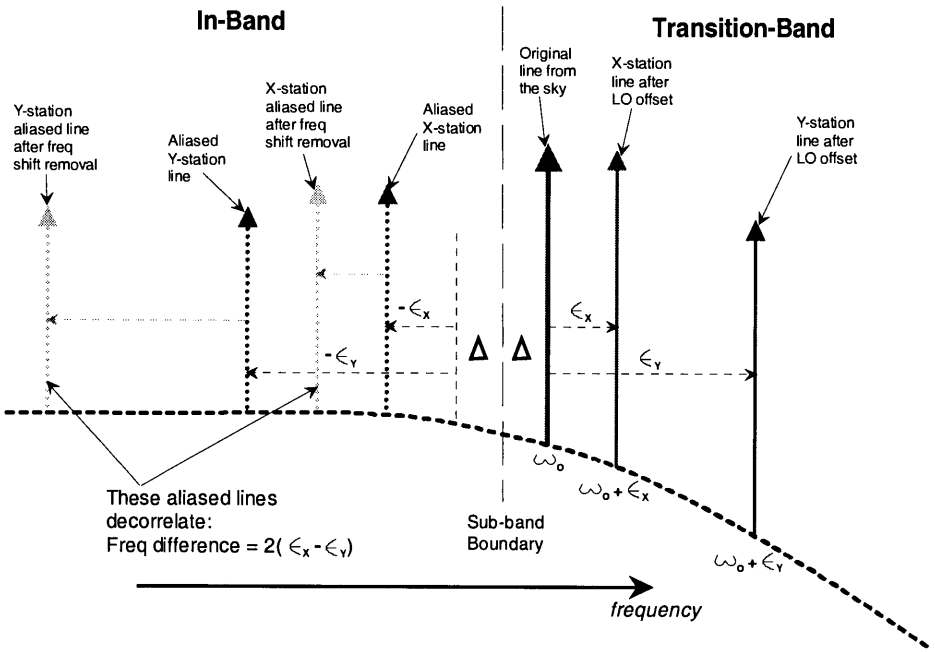


Figure 8-8 Diagram illustrating how, with sine waves (δ functions in the frequency domain), the aliased transition band decorrelates. The solid signals are the originals, and the dotted signals are the aliased components. The "Original line from the sky" does not actually exist in the system except in the integrated cross-power spectrum. A fine point to note is that the Local Oscillator frequency shifts should be in the same direction (same sign), otherwise the aliased components move closer together after fringe stopping in the correlator—reducing the desired decorrelation effect.

Obvious questions are: how much does the Local Oscillator offset have to be? and how much does the aliased signal decorrelate?

The decorrelation amplitude is given by the fringe washing function $\frac{\sin(2\omega T)}{2\omega T}$; where ω is the (radian) frequency difference between the two antennas' Local Oscillators and T is the *incoherent* integration time in seconds. The factor of '2' results from the fact that frequency sense of the aliased transition band is opposite to that of the main part of the band. (i.e. when fringe stopping is done in the correlator, the difference between the fringe stopping frequency and the aliased transition band frequency shift is double the original frequency shift difference.) This amplitude drops rapidly with integration time. Figure 8-9 shows the decorrelation amplitude versus time with a 10 kHz frequency offset (i.e. antenna X and Y with Local Oscillators that are different by 10 kHz). Note that the decorrelation occurs proportional to the integration time and not the square-root of the integration time for noise as indicated in the Figure (noise envelope is for ~250 Ms/s).

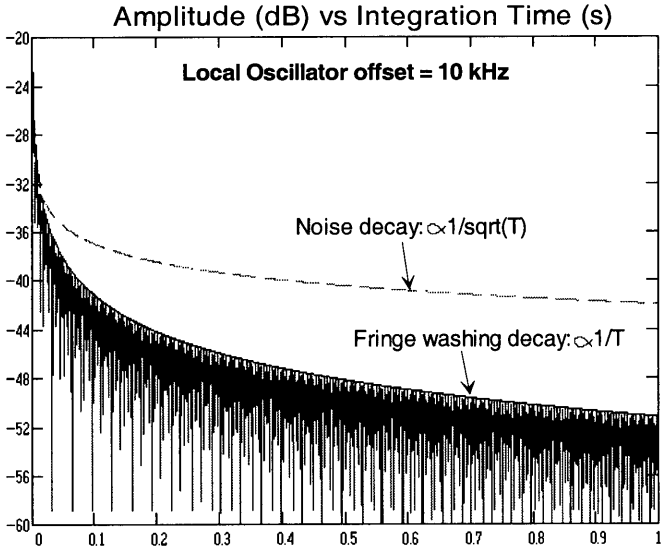


Figure 8-9 Fringe washing decay (decorrelated amplitude) versus integration time for a 10 kHz X and Y antenna Local Oscillator frequency difference. For comparison, the noise decay is shown for a ~250 Ms/s sample rate.

If many antennas are used in the array, then it is necessary to ensure that a minimum offset of, say, 10 kHz occurs between any two antennas. Thus, the maximum frequency offset is *#of antennas x minimum offset*. For the EVLA, with 40 antennas and a 10 kHz proposed minimum offset, the maximum offset is 400 kHz. Since this is << sub-band bandwidth and because of the correlated overlap of transition bands, no loss of spectral information occurs. It should be noted that since frequency shift removal occurs after filtering, the **bandshape correction term** in the sub-band normalization equation should take this into account. That is, X and Y bandshapes should be shifted in frequency before being multiplied together to obtain the correct $S_{XYFIR_{nf}}$. This will only be necessary if the frequency shift is a reasonably significant fraction of the transition band.

A final subtle point here is that decorrelation of the aliased transition band occurs with incoherent integration time. However, an increase in decorrelation will not occur when

many correlator dumps are integrated *coherently*. This is because phase for the decorrelated component is wrapping with time. That is, with every incoherent dump, the decorrelated component has a phase and this phase will be changing with time—if the change is linear, then it coherently integrates. This effect is inherently mitigated by Doppler phase acceleration on the sky since this will cause the decorrelated phase to change non-linearly with time. Thus, the antennas closest to the phase center of the array should have the largest relative frequency offsets since they will generally have the lowest Doppler phase acceleration. If necessary, explicit mitigation of this effect can be achieved by varying, in some random way, the frequency offsets of the antennas. This will cause the residual decorrelated phase to be random with time—effectively preventing coherent integration.

8.5 Quantization Noise Decorrelation

Because the frequency shift is introduced in the analog system before quantization and removed after quantization (sub-band filtering or not), quantizer noise and harmonics from strong narrowband signals[39] generated by the quantizer decorrelate. This is because quantizer noise and quantizer harmonics 'splattered' across the band are at frequencies that are some multiple of the *offset* fundamental frequencies in the X and Y antennas. Thus, when the frequency shift is removed in the correlator, the harmonics don't get shifted properly to allow them to correlate. This is illustrated with a narrowband signal in Figure 8-10.

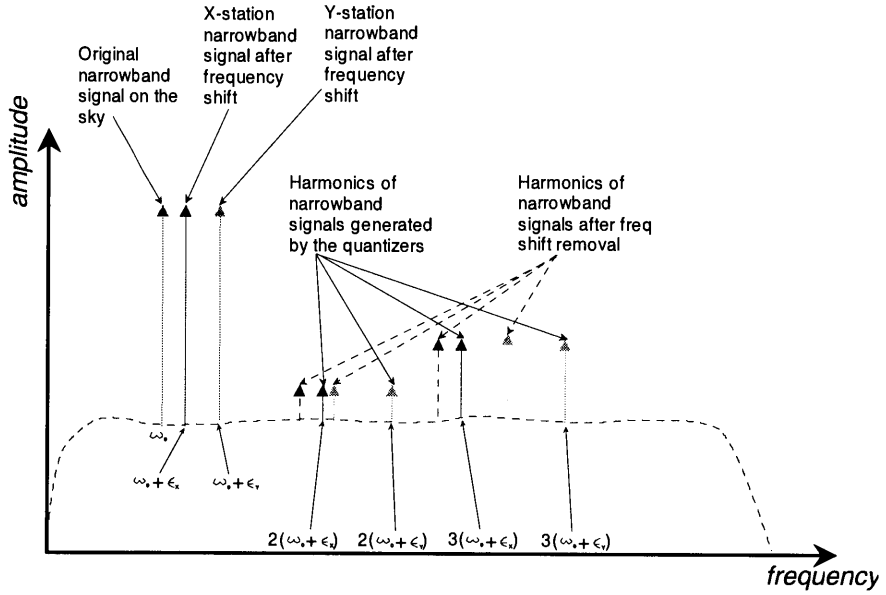


Figure 8-10 Decorrelation of quantizer generated harmonics. Harmonics of narrowband signals decorrelate since the frequency shift that they have is the Local Oscillator shift times the harmonic number. When the frequency shift is removed in the correlator, it prevents the harmonics from overlapping in frequency and therefore prevents them from correlating.

[39] Correlated quantizer noise from strong continuum signals is not a concern since these levels would be much lower than the continuum correlation level anyway.

The quantizer also generates broadband quantization noise from narrowband signals. This noise decorrelates somewhat but appears to be limited to about 15 dB better than what would happen if there were no frequency shift. This does not appear to be a limitation of the digital fringe stopper (described shortly), but rather is a result of many intermodulation products of harmonics aliased many times overlapping in frequency and eventually correlating. It is difficult to quantify this effect but simulation (and some study with quantized pure sine waves) has shown that this appears to be the case.

The harmonic decorrelation shown in Figure 8-10 has a fundamental limitation due to the use of digital mixers in the correlator. As it turns out, the harmonics[40] of the digital mixer function always fringe-stop the quantizer-generated harmonics of the narrowband signal. Since the digital mixer only contains odd harmonics this means that all odd harmonics of narrowband signals get fringe-stopped by the digital mixer with a resulting amplitude that is a function of the signal harmonic amplitude and the digital mixer harmonic amplitude. The time and frequency domain representation of a 3-level [3] fringe rotation/digital mixer function is shown in Figure 8-11.

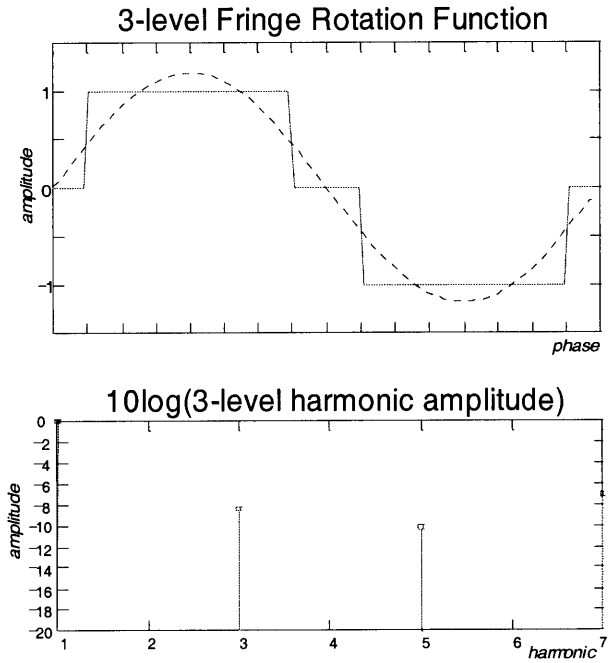


Figure 8-11 3-level fringe rotation/digital mixer function time domain (top) and frequency domain (bottom) representations. The dotted line in the top plot is a perfect sine wave that the digital function is trying to model. Harmonics of the digital mixer are about 6 dB lower than the fundamental. Note that even though the digital mixer function is a voltage, only 10log() of the harmonic amplitude is used to obtain its performance since the mixer is a "one sided" voltage multiplier in the correlator chip.

It is possible to suppress this effect somewhat by choosing a digital mixer function that reduces its 3rd harmonic amplitude since typically, the 5th and higher-order quantizer-

[40] In the April 2000 meeting, it was stated that only the 3rd harmonics of the digital mixer stop the 3rd harmonics from the quantizer. More careful investigation has revealed that all harmonics behave the same.

generated harmonics are >10 dB below the 3rd[41]. The choice of an alternate digital mixer function and its benefits must carefully consider the impact on the implementation of the correlator chip.

As luck would have it, there is a simple 5-level digital mixer function that suppresses the 3rd harmonic component by an additional 6 or 7 dB. This function is shown in Figure 8-12 and requires logic in the correlator chip that is similar to the 3-level function except that an additional bit shift (to multiply by 2) is performed. It is not anticipated that this is an unreasonable hardware requirement for the correlator chip.

Use of the 5-level fringe rotation function generally enables the WIDAR correlator to achieve about 14 dB better spectral dynamic range than a correlator without frequency shifting and digital mixing. This is a significant improvement in the performance of the correlator especially in the presence of strong narrowband signals where quantizer-generated harmonic amplitudes can be significant.

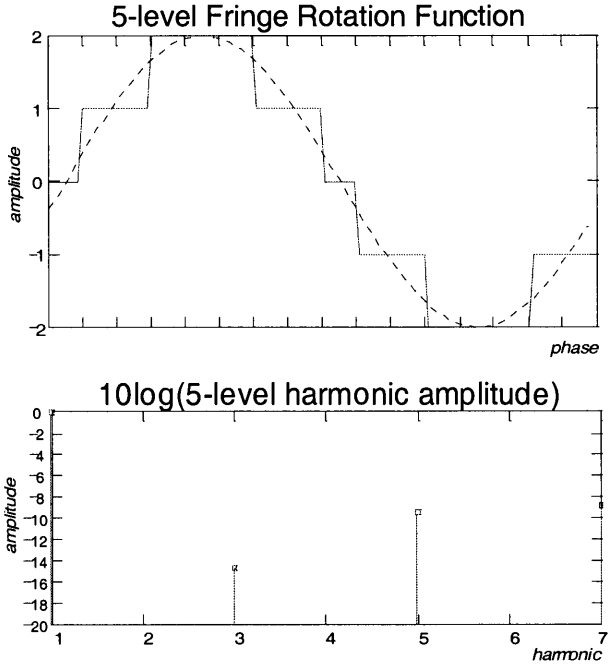


Figure 8-12 5-level fringe rotation/digital mixer function. The 3rd harmonic is about 8 dB lower in this function than the 3-level function. This will enable the WIDAR correlator to generally provide about 14 dB better spectral dynamic range than a correlator without frequency shifting. This is especially important in the presence of powerful narrowband signals. The 5-level fringe rotation function also reduces the sensitivity loss in the correlator to about ½ the 5% loss from a 3-level function.

8.6 Digital Mixer Phase Dithering

In the proposed correlator chip discussed in Section 4.2, the *baseline* phase is calculated at each lag by differencing the X and Y-station *quantized* phase before generating the digital sine and cosine functions. Generally, this operation will cause some phase

[41] As demonstrated with simulations of 3-level and 15-level quantization of strong narrowband signals.

dithering to occur when changing baseline phase states. This phase dithering will subsequently yield sine and cosine amplitude dithering. Dithering introduces a slight SNR degradation (0.62% for a 3-level function as shown in [1]) over the case where the baseline phase is first calculated with many phase bits and then quantized. For VLBI, dithering is inherently guaranteed since the X and Y station phase rates are a function of antenna position, earth rotation, and sky frequency—generating phase rates that are not frequency locked.

For the EVLA, where the fringe rate is mostly artificially imposed by offsetting the Local Oscillators, and the correlator (presumably) removes the antenna Doppler shift, it is possible to get a situation where the fringe rotators alternate between dithering and no dithering if the Local Oscillator offsets are harmonically related. This will modulate the output cross-correlation amplitude—a decidedly unwanted effect. This situation can be avoided if the Local Oscillator offsets are chosen so that no two antennas' offsets are harmonically related. This is a trivial but important requirement. Figure 8-13 shows the output of the 5-level fringe rotation function when dithering is present. Also note that different baseline-based Local Oscillator offset differences will generate different dithering patterns—it is not completely known if these different dithering patterns have a *subtle* effect on correlator amplitudes. If this effect is significant, then it may be necessary to: 1. choose Local Oscillator offsets so that random dithering patterns occur on all baselines; 2. stop Doppler fringes at the antennas so that frequency shifts and thus correlator amplitudes stay constant on all baselines; or 3. in the worst case do not use the quantized lag-based fringe stopping method.

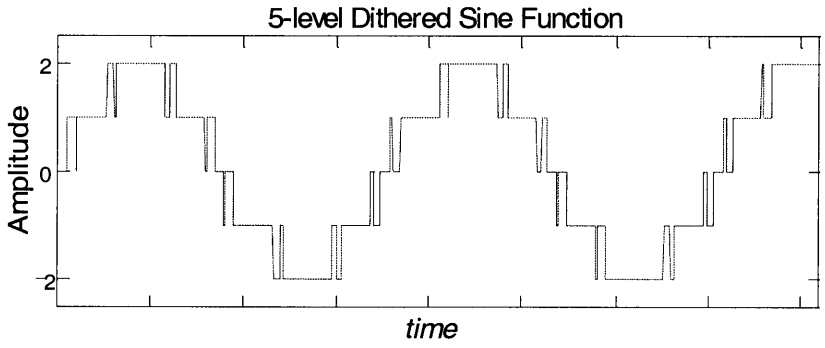


Figure 8-13 5-level dithered sine function produced when baseline phase is formed from differencing quantized station-based phase. The Local Oscillator offsets should be chosen so they are not harmonically related to ensure that phase dithering always occurs in order to avoid possible modulation of correlator amplitudes if phase differencing wanders in and out of dithering.

8.7 Digital Sub-sample Delay Interpolation/Tracking

A requirement of the EVLA correlator (or associated subsystems) is that at least 1/16th sample delay tracking must be performed. Traditionally, this is done in one of two ways[42]: modify the phase of the sampler clock in 1/16th sample steps; or dump the correlator fast and remove the frequency-domain phase slope (equivalent to a delay)

[42] ...in an 'XF' correlator. Provided the delay rate is not too large, an 'FX' correlator can apply a phase-slope correction in freqeuncy in real-time before correlation.

before integrating with other dumps. The first method allows the sub-sample delay tracking to be performed properly but it complicates the sampler/quantizer design and, since it is done in real analog electronics, there is some indication that there are subtle effects due to an imperfect hardware realization. The second method (post-correlation correction) works well, but at the high sample rates that are going to be used, prohibitive dump times are required to perform the operation properly—especially on longer baselines.

A side-benefit of WIDAR is that fully digital $1/16^{th}$ (or $1/N^{th}$ – N being the demultiplexing factor) sample delay tracking can be fairly easily realized. The method is shown in Figure 8-14. Integer (+/-0.5 sample) delay tracking introduces a phase slope in the wide band that is 0 at DC and ranges between +/- 90°, depending on what the delay error is, at the top edge of the band (dotted sloped lines in the figure). With WIDAR, when a sub-band is selected, the phase excursion[43] within that sub-band due to the wideband integer delay tracking is $1/16^{th}$ (for N=16) of that in the wide band plus a phase offset that changes with delay. The phase offset is tracked by offsetting the phase of the fringe stoppers in the correlator chip and all that is left is a phase slope, centered in the middle of the sub-band, that alternates between +/-90°/16. Actually, if phase-offset tracking is set to the middle of the sub-band, the phase alternates between +/-90°/32 at both edges of the sub-band—yielding better performance than required.

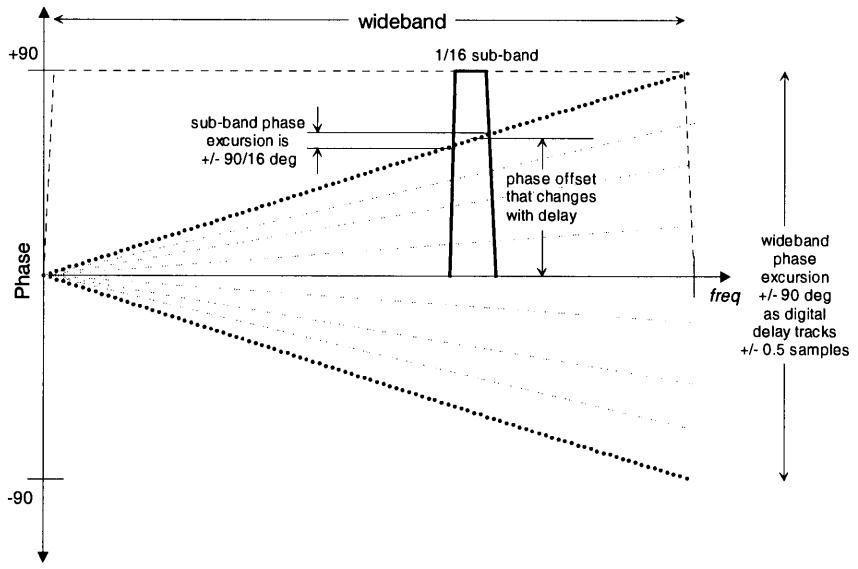


Figure 8-14 Fully digital sub-sample delay tracking is realized with the WIDAR correlator. Integer delay tracking at the wideband sample rate introduces a phase slope that is 0 at DC and alternates between +/-90° at the upper edge of the band. When sub-band filtering is performed, the phase excursion within the sub-band due to the wideband tracking is $1/16^{th}$ (N=16) of the wideband plus a phase offset that changes with delay. The phase offset is tracked by offsetting the phase of the fringe rotators in the correlator.

[43] It is the phase excursion across a correlated band within an integration time that causes decorrelation and loss of signal amplitude.

8.8 Phased-Array Operation with WIDAR

It is a requirement of the EVLA correlator to be able to simultaneously provide phased-antenna data for VLBI. This is easily done with WIDAR by using desired sub-bands of desired basebands, removing antenna frequency shifts (including phase offsets for 1/N delay tracking) with a complex digital mixer[44], adding the complex data from the antennas together, and then performing a broadband phase shift to finally yield a phased output. A block diagram of this process is shown in Figure 8-15. For the VLBI application, the sub-band data is requantized to 2 bits before being recorded on VLBI tape. Provided the antenna samplers are using standard VLBI rates and are frequency locked to the VLBI frequency standard, it is not necessary to perform a D/A and A/D operation.

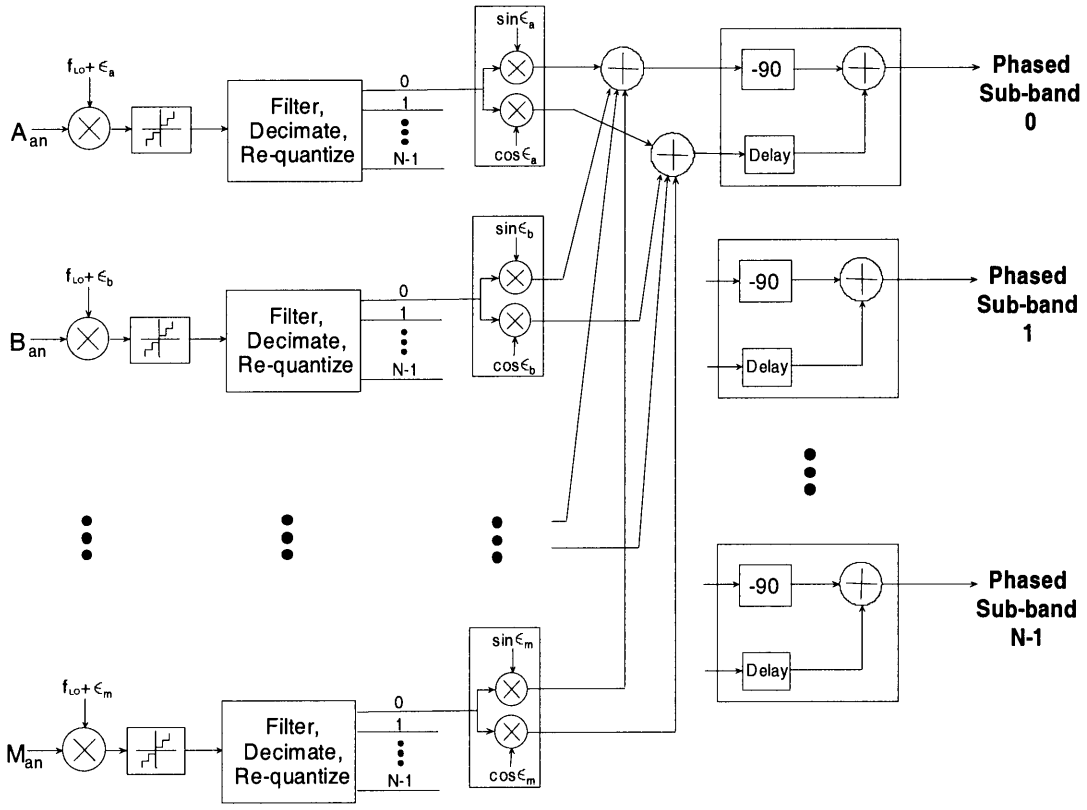


Figure 8-15 Phased-array operation with WIDAR. FIR filter sub-band outputs (which could be the same ones used for correlation) go to complex mixers where the frequency shift (and antenna Doppler shift) is removed with many-level sine and cosine functions. The complex data from all antennas in the phased-array are then added together—the results of which get converted to simple data with a -90° phase shift (implemented in a Hilbert FIR filter) and add as shown. Finally, the simple data is requantized (if desired and not shown in the figure) for further processing.

[44] This mixer must use many levels to minimize cross-correlation of mixer harmonics from both antennas. This is a well-known problem and is the reason why XF VLBI correlators perform only one mixing function in the correlator chip.

A slight variation from the phased-array application for VLBI is if bunches of antennas are to be phased together before cross-correlating the bunches as interferometer elements. For example, this could be required if many small antennas are to be used for sensitivity and it is infeasible to cross-correlate them all in a correlator (such as in the 1hT for SETI [4]). In this case it is not necessary for every antenna in the bunch to have its own frequency shift. It is merely sufficient for the bunches of antennas to have different frequency shifts so when they are correlated, the aliased transition bands of the sub-bands decorrelate. Thus, the complex mixers shown in Figure 8-15 would simply remove antenna Doppler shift—not a problem since many-level sine/cosine functions are used—and track the delay-dependent phase offset. A subtle penalty with this configuration is that 1/N delay-tracking would not be able to be performed[45] in the cross-correlation of the bunches of antennas since, at that point, integer delay tracking would be at the sub-band sample rate rather than the original wideband sample rate.

8.9 Adaptive Interference Cancellation

The interference mitigation that is planned for the WIDAR-EVLA correlator is straightforward and, given the capabilities of the correlator, feasible. This interference mitigation is most effective for continuum observations where it is acceptable to throw away some bandwidth where there is interference with the penalty being a loss of sensitivity. However, in spectral-line observations it may be the case that the emission or absorption line of interest is at the same frequency location as a much more powerful interfering source. The only way to be able to recover the desired line in this case is to remove the interference. This can potentially be done with adaptive cancellation.

There is some adaptive cancellation research being performed at various institutes around the world and the basic method is the same. The method is to obtain a copy of the interference signal that does not have the radio astronomy signal in it, shift its phase by 180°, and add it to the interference+radio astronomy signal to remove the interference. The "copy of the interference signal" can be obtained with the use of a separate antenna not sensitive to the radio astronomy signal, or synthesized [5] if its waveform is predictable in time. A WIDAR "add-on" that could perform the adaptive cancellation is shown in Figure 8-16. WIDAR is ideally suited to adaptive cancellation because the operation can be performed on the sub-band data rather than the more difficult to handle wideband data. The FIR filters can be programmed to change the gain of the interference-only signal as well as filter out those parts of the band where there is no interference (i.e. to minimize the additional noise added to the radio astronomy signal).

[45] However, the bunches of antennas have been phased together with 1/N delay tracking.

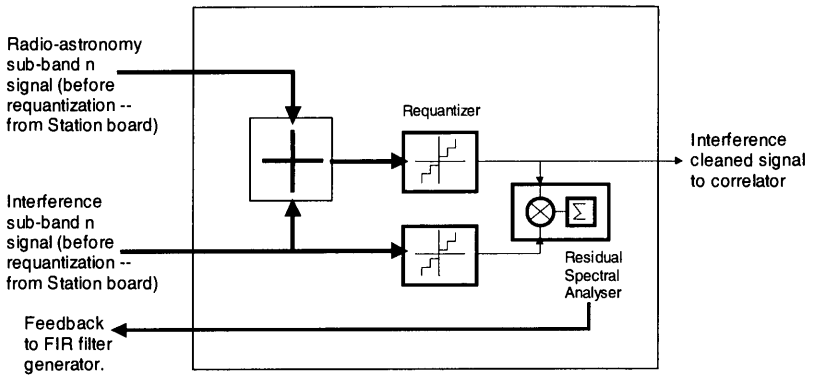


Figure 8-16 Block diagram of an adaptive cancellation add-on to the correlator. The radio astronomy signal is added to a cancellation signal containing only the inverted interference signal. The result is requantized and compared with the cancellation signal to allow fine-tuning of the cancellation signal amplitude. Ideally, the interference-cleaned signal that goes to the correlator contains no interference.

The method seems straightforward, but in practice it is difficult to obtain a high SNR interference-only signal from a secondary antenna. In this case the attenuation of the interference signal is limited by the SNR of the copy. 20 dB attenuation is typically obtained. High SNR interference signals can be synthesized, but only for those signals that are predictable in time—not generally the case with a communications signal. Nevertheless, WIDAR is well suited to adaptive cancellation should the techniques ultimately prove to be worthwhile.

8.10 Windowing

Windowing of lag data before Fourier transform to the frequency domain is a necessary operation if the spectral dynamic range of the signal is greater than ~10 dB or if there are discontinuities in the spectral signature of the signal (e.g. unresolved lines). The latter effect causes ringing near the spectral discontinuity known as Gibbs phenonmenon and the former causes duplicates of high dynamic range spectral features to appear across the band due to the sidelobes of the window function (e.g. without windowing—a.k.a. rectangular windowing—these are sidelobes of the *sinc()* function). There are many window functions in use but perhaps two of the most suitable are the Hanning window and the Kaiser window functions. Hanning windowing provides good performance for a spectral dynamic range of about 30 dB. Kaiser windowing is flexible and involves a tradeoff between sidelobe level and main lobe width. With Kaiser windowing and $\theta = 4\pi$ (θ is the adjustable parameter) a spectral dynamic range of ~80 dB can be achieved at the expense of ~2.5 times the main lobe width.

Windowing is a non-trivial exercise when adjacent WIDAR sub-bands are to be seamlessly stitched together. This is because there are two conflicting requirements at work:

1. To minimize the size of the SNR degradation region near the sub-band boundary, the filter cutoff (i.e. filter gain at the boundary) should be minimized but still stay within the linear-phase region of the filter. This is

typically greater than –15 dB. This ensures that the transition band is very steep resulting in a steep and rapidly declining SNR degradation away from sub-band boundary.

2. Windowing is a function of the bandpass response (shape of the spectrum) *and* the spectral features present within the bandpass. Thus, in order to provide a final windowed result that is generally 'correct'[46], it is desired to have (ideally) a flat amplitude response across the sub-band. If the bandpass response at the edge of the sub-band is not flat (i.e. steep as desired by 1. above) then the windowed spectrum, if there are spectral features present, will not be correct. That is, the windowed spectrum will not be the same as if the bandpass response was flat.

In an effort to resolve these two conflicting requirements and provide general windowing capability, the following three windowing methods have been developed. These methods assume a reasonably flat bandpass response and were tested with a –1.2 dB sub-band filter cutoff. For very narrow bandpass response filters (e.g. 1/256) it is important to minimize the SNR degradation in the region near the edge of the sub-band requiring a cutoff of about –10 dB—with an accompanying steep bandpass response. In this case it may not be possible to perfectly stitch adjacent sub-bands together if anything other than Method 1 is used. However, normally if very narrow sub-bands are used, it is not necessary to stitch adjacent sub-bands together anyway.

Windowing Method 1: Rectangular windowing

Use rectangular windowing when the spectral-dynamic range of the data that is being Fourier transformed is <= ~10 dB. The procedure is to FFT each sub-band's lags, apply the calculated bandshape correction (FFT FIR tap coefficients and square the result), and concatenate sub-band frequencies (use DC to Nyquist/2 – 1 frequency bins in each sub-band). In this case, the sub-band boundary SNR degradation is $\sqrt{2}$[47].

Windowing Method 2: Frequency convolution

Use a desired window function when the spectral-dynamic range of the data that is being Fourier transformed is <=~20 dB. The procedure is to FFT each sub-band's lags, apply the bandshape correction, and then convolve with the FFT of the window function. With Hanning windowing, a good approximation of the convolution function has three points: 0.5, 1.0, 0.5. This method requires a spectral overlap point just before DC and at Nyquist/2 so the bandshape correction must apply over an appropriate range of data. Here, the sub-band boundary SNR degradation is (for a –1.2 dB cutoff filter) about $\sqrt{3}$ (70%). This is because the frequency bin just below DC and at Nyquist/2 have a worse degradation than $\sqrt{2}$. If it is desired to ensure that the SNR degradation does not exceed $\sqrt{2}$, then the cutoff of the filter should be reduced to say –0.2 dB. In this case, the sub-band overlap region is larger—which can be mitigated somewhat with a longer FIR filter.

[46] 'Correct' is defined here to be the same as the result obtained with a flat bandpass filter.

[47] If there are no spectral features. For example, if there is a strong line right on the sub-band boundary more than $\sqrt{2}$ SNR degradation can occur.

Windowing Method 3: Sub-band overlap merge

This is probably the best general purpose windowing method that should be used when strong narrowband signals are present. It will ensure that even if a strong narrowband signal sits right on (or, worst yet, slightly offset from) a sub-band boundary, the result will be very close to the ideal, flat-passband result. The procedure is to window the sub-band lags, apply a windowed sub-band bandshape correction[48] to the spectral points, and use a weighted average of overlapping frequency points between adjacent sub-bands. Thus, the frequency bins near the sub-band boundary consist of averages of data from both adjacent sub-bands. With a –1.2 dB cutoff filter and using three overlapping points the weights were found to be (1.0,~0.2), (1.0, 1.0), (~0.2,1.0)—the first number in the pair is the weight for the lower sub-band point and the second number is the weight for the upper sub-band point of the adjacent sub-bands. This windowing method was found to work well for all arbitrary data sets tested and with Kaiser windowing. More overlapping points can be used for better performance over more frequency bins.

More research will be required to further quantify and refine the windowing methods presented here. The bottom line is that there will be a tradeoff between precision in stitching adjacent sub-bands together, length of sub-band FIR filters, and the size of the SNR degradation region.

[48] This is obtained by Fourier transforming the tap coefficients, squaring the resulting spectra, inverse transforming, applying the window function, and finally Fourier transforming back to frequency.

9 Appendix II – Representative WIDAR Correlator Simulation Results

A large number of simulations were performed in an effort to determine the feasibility of a WIDAR correlator. Much of the behaviour of the correlator system presented in previous sections of this document were first determined with the simulator and then analyzed and theoretically explained. Many of the signals that were simulated were very complex in nature and so the WIDAR correlator simulation results were compared with ideal correlators that, in some cases, can only be constructed inside a computer. This comparitive simulation has allowed subtle effects to be determined and has served to quantify the behaviour of the signal processing architecture.

This appendix explains the methods that were used in the simulator and presents a subset of the simulation results that were obtained.

9.1 Simulation Methods

The simulator was built in 'C' and run on a "Beowulf" computer cluster consisting of 16, Dual PIII (Pentium III) PCs. The PCs are connected together with 100 Mbps Ethernet through a network switch. All of the PCs see one master disk with the same paths and directory structure. The simulation jobs were split between sub-band correlators for the WIDAR correlator, and between different parts of the lag chain for the wideband correlator(s). Each job generated its own simulated quantized signal so that there was no need for process-to-process or CPU-to-CPU communication. Three distinct correlators were used in the simulations. The WIDAR and fullband correlators are not simulated—they are software representations of correlators that could be built with real hardware[49]. These three correlators are:

1. **WIDAR correlator**. This correlator is functionally the same as it would be implemented in hardware. The sub-band FIR filters have finite length coefficient*data products (implemented in a lookup table as would be done in hardware). The re-quantizer is as precise as it could be built in hardware. The lag 0 autocorrelation of data out of the FIR filters used all of the available bits on all of the samples. Fringe stopping and complex correlation is as it would be in the correlator chip—except that the multiplier output is unbiased. (For efficiency, the correlator chip multiplier output is biased—the bias is removed in software after correlation.)

2. **Reference correlator**. This is a perfect wideband correlator with floating-point fringe stopping. Fringe stopping was used in this correlator so that the harmonic decorrelation benefits of fringe stopping in this correlator could be compared to what is obtained with digital 3-level and 5-level mixers in the WIDAR correlator.

[49] Because it uses floating-point, the reference correlator would be very expensive to build in actual hardware although it is theoretically possible.

3. **Fullband correlator**. This is a perfect wideband correlator but without fringe stopping. This allows a comparison between what WIDAR can achieve and what a "normal" correlator could achieve if fringe-stopping is done by changing the phase of the signal at the antennas (as is normally done in connected-element interferometers).

The quantizer thresholds were set to 0.374σ for 4-bit quantization, 0.65σ for 3-bit quantization and 1σ for 2-bit/4-level quantization (n=3 correlation). Where 4-bit (15-level) and 3-bit (8-level) data was being correlated, normalized correlation coefficients[50] were generated simply by taking the correlator output and dividing by the lag 0 autocorrelation of quantized data. This is very close to a linear correction for 4-bit data but it is unknown how close a correction this is for 3-bit data. 4-level (2-bit) Van Vleck corrections were performed by using the slope of <r4> vs ρ, σ=1, n=3 at ρ=0.3 and applying a linear correction to all data.

The digital sub-sample delay tracking capability of the WIDAR correlator was not simulated. It was not done because this process appears to be straightforward and the extra complication adds compute time and code complexity. If necessary, the simulator could be upgraded to provide this capability. The X and Y station noise generator was constructed as shown in Figure 9-1.

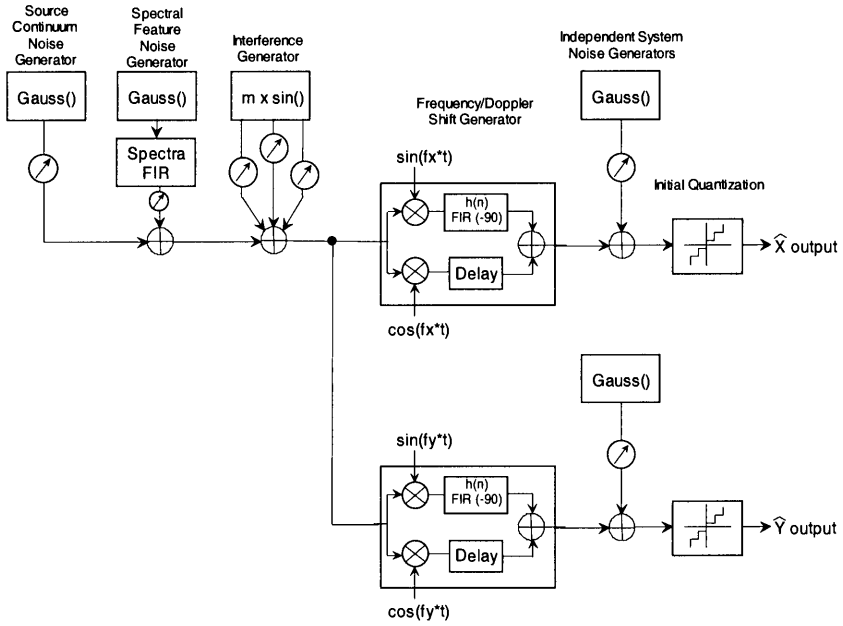


Figure 9-1 Simulator's noise generator block diagram. Common noise, spectral features, and interference that will correlate goes into separate frequency shift generators implemented as digital single-sideband mixers. Independent/uncorrelated noise is added to the frequency shift data before initial quantization. All operations before initial quantization are double precision floating-point.

[50] For WIDAR, this is the first step in the normalization process—produce normalized sub-band correlation coefficients which are then scaled according to the WIDAR sub-band normalization equation. For the reference and fullband correlators this is the only normalization that is required or performed.

All operations in the noise generator before initial quantization are double precision floating-point. Broadband noise was generated[51] with a Gaussian noise generator taken from Numerical Recipes in 'C'. Complex spectral features were generated by FIR filtering a broadband noise source—the FIR tap coefficients were generated using an FFT of a desired shape. Narrowband interference was generated with simple sine waves whose amplitude and placement in the passband could be changed as desired. The frequency shift was implemented with a digital single-sideband mixer as shown in the figure. It used a 165-tap Hilbert FIR filter that provided >50 dB of sideband rejection. The initial quantizers have the capability of having their decision levels randomly offset with Gaussian distribution (i.e. the "random" function here was Gaussian). Many tests used a 1-sigma offset of 0.05ρ of the input signal—this ensured that 3-sigma offsets in adjacent decision level points did not delete one output level. An error in a real quantizer that was not simulated was differing decision levels depending on whether the input signal was slewing positively or negatively. Nevertheless, it is asserted that the effect of any initial quantizer errors will be the same in the WIDAR correlator and the reference correlator since, with WIDAR, each sub-band filter sees all the initial quantizer data.

Normalized complex-lag correlator output data is written to individual ASCII files. Lag 0 autocorrelator data after sub-band filtering but before requantization from the WIDAR correlator is also written to ASCII files. This data is required to scale and normalize sub-band spectra before stitching them together. All simulator ASCII files are read into MathCad where final corrections, windowing, and stitching are performed. The MathCad programs also perform any desired comparisons between the WIDAR correlator and the reference or fullband correlators. Unless otherwise stated, arithmetic mean total power normalization was performed on WIDAR data so the results could be directly compared with wideband correlator data (whose normalized correlation coefficients are always normalized to total power).

A number of simulations were performed taking anywhere from a few hours to several weeks to complete. The simulations ran the following gamut of tests:

- Continuum—high/"low" SNR.
- Mixed continuum + emission lines.
- Mixed continuum + emission lines + narrowband interference.
- High dynamic range (crowded spectrum) emission lines + narrowband interference.
- VERY high dynamic range (crowded spectrum) emission lines + narrowband interference.

[51] The Gaussian noise generators were reset and used different random number generator seeds for each correlator 'dump'. This was done to try to avoid non-random behaviour after generation of many (~200e6) samples. Also, except for one test, the WIDAR correlator used different seeds in the noise generator than the wideband correlator so that statistical comparison of correlated data was performed.

- Absorption spectrum - high and medium dynamic range.
- Broad structure emission spectrum.
- 2, 4-bit initial quantization. 2, 3, 4-bit re-quantization.
- Initial quantizer decision level errors.
- Real (Vela) pulsar spectrum.

9.2 Plots of Some Simulation Results

The following figures are plots of some representative simulation results. The figure captions briefly explain the simulation and any noteworthy results. Text on the plots themselves explains most simulation parameters.

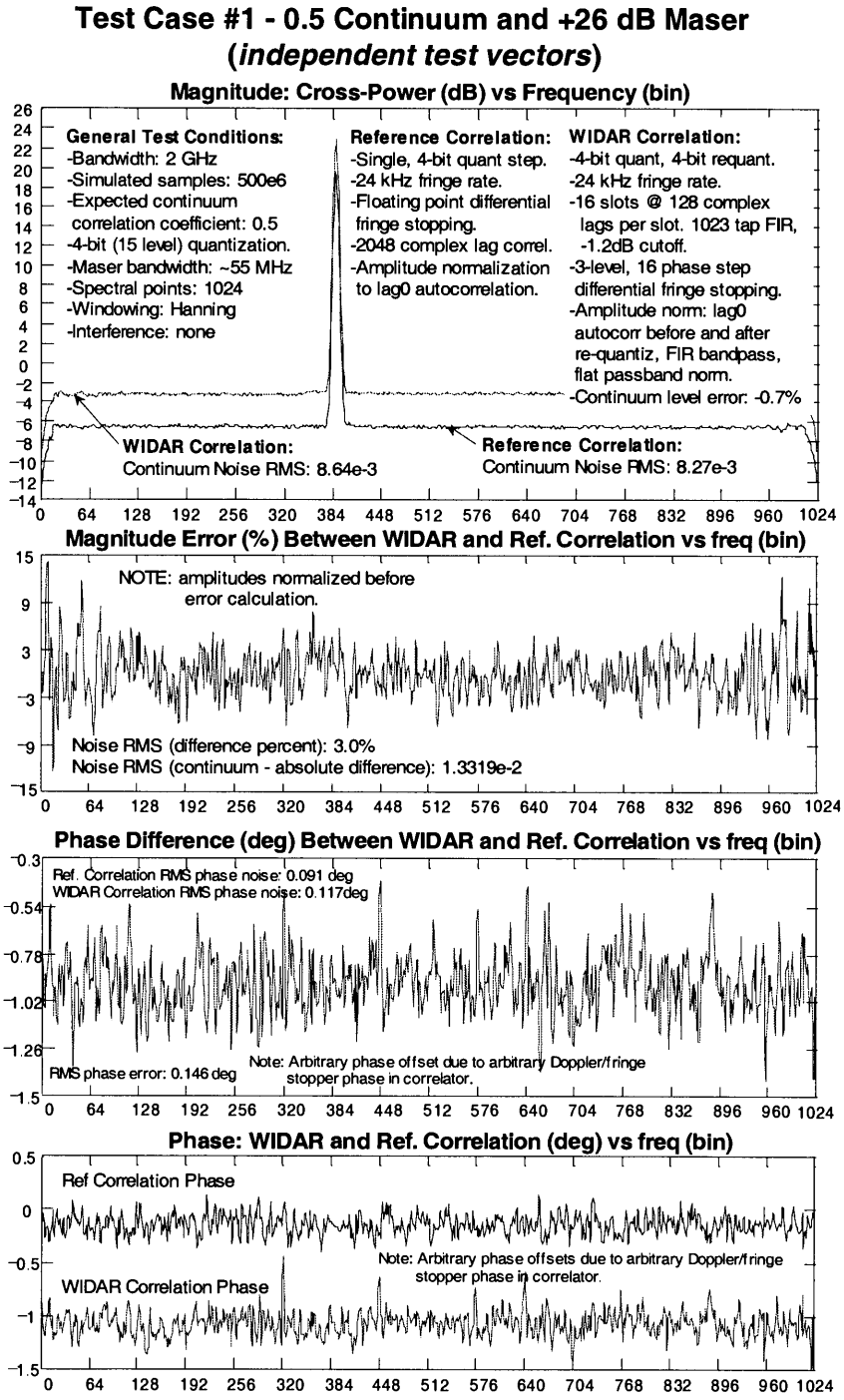


Figure 9-2 Continuum and strong maser simulation with independent test vectors. The WIDAR amplitude was normalized using the "Sub-band Power" method for P_{XYT} so that the correct continuum level is obtained in the presence of the strong narrowband signal. Only one integration of 500 million samples was simulated—the degradation of the noise generator is seen in the ragged spectra near the edges of the band. Window method #2 is used.

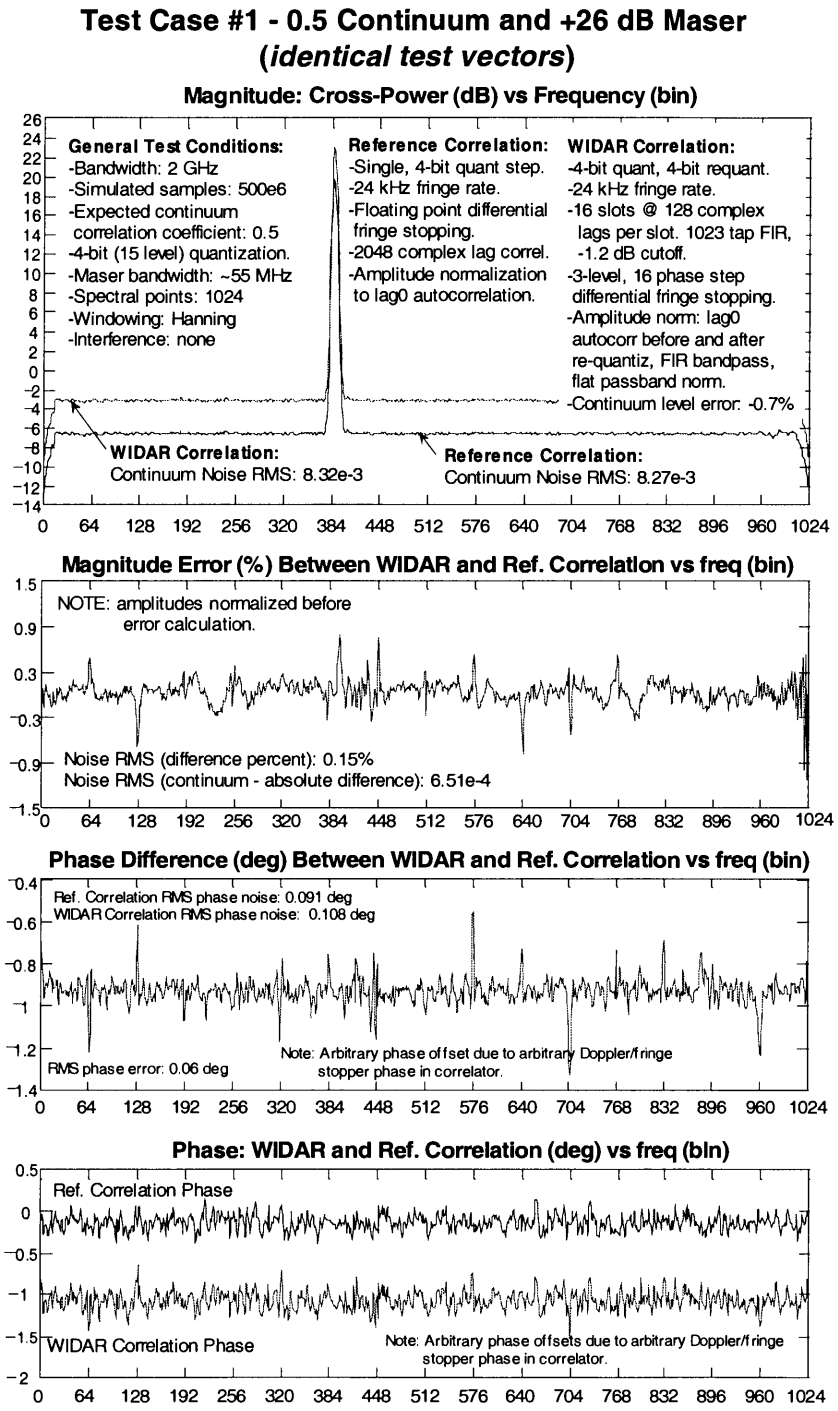


Figure 9-3 Continuum and strong maser simulation with *identical* test vectors. This is the same as the previous figure only the test vectors into the reference correlator and the WIDAR correlator are the same—resulting in a very sensitive test. In this case, noise waveforms are being compared—resulting in seemingly large amplitude differences at sub-band boundaries since this is where aliased and decorrelated independent noise will be present (i.e. the aliased transistion band represents independent noise compared to the reference correlator). Window method #2 is used.

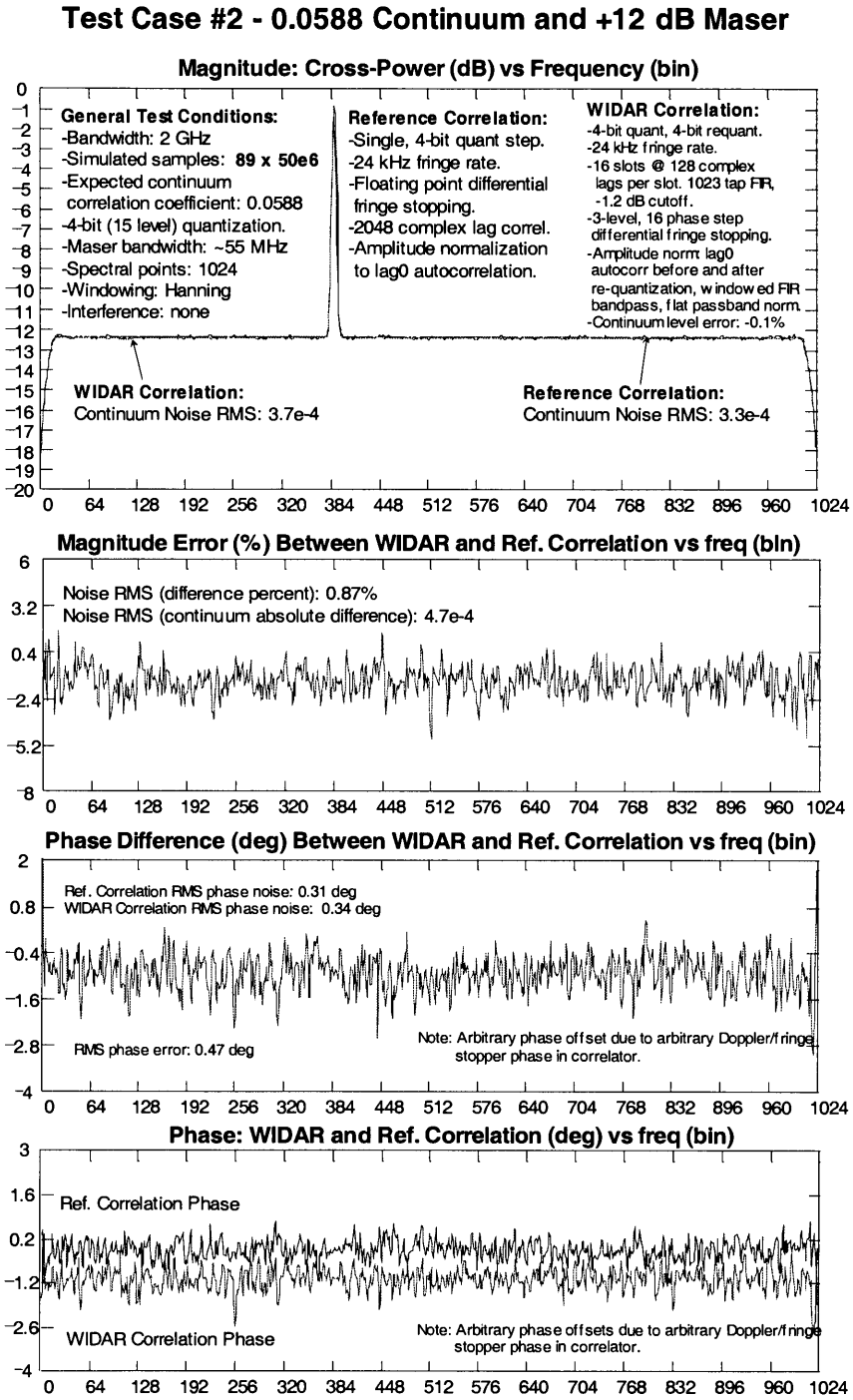


Figure 9-4 Continuum and weaker narrowband signal. 89 dumps of 50 million samples each were incoherently integrated to produce this result. The noise behaviour is better than the previous two plots. Independent test vectors were used and different noise generator seeds were used for each dump. Arithmetic mean total power normalization was used for both correlation results.

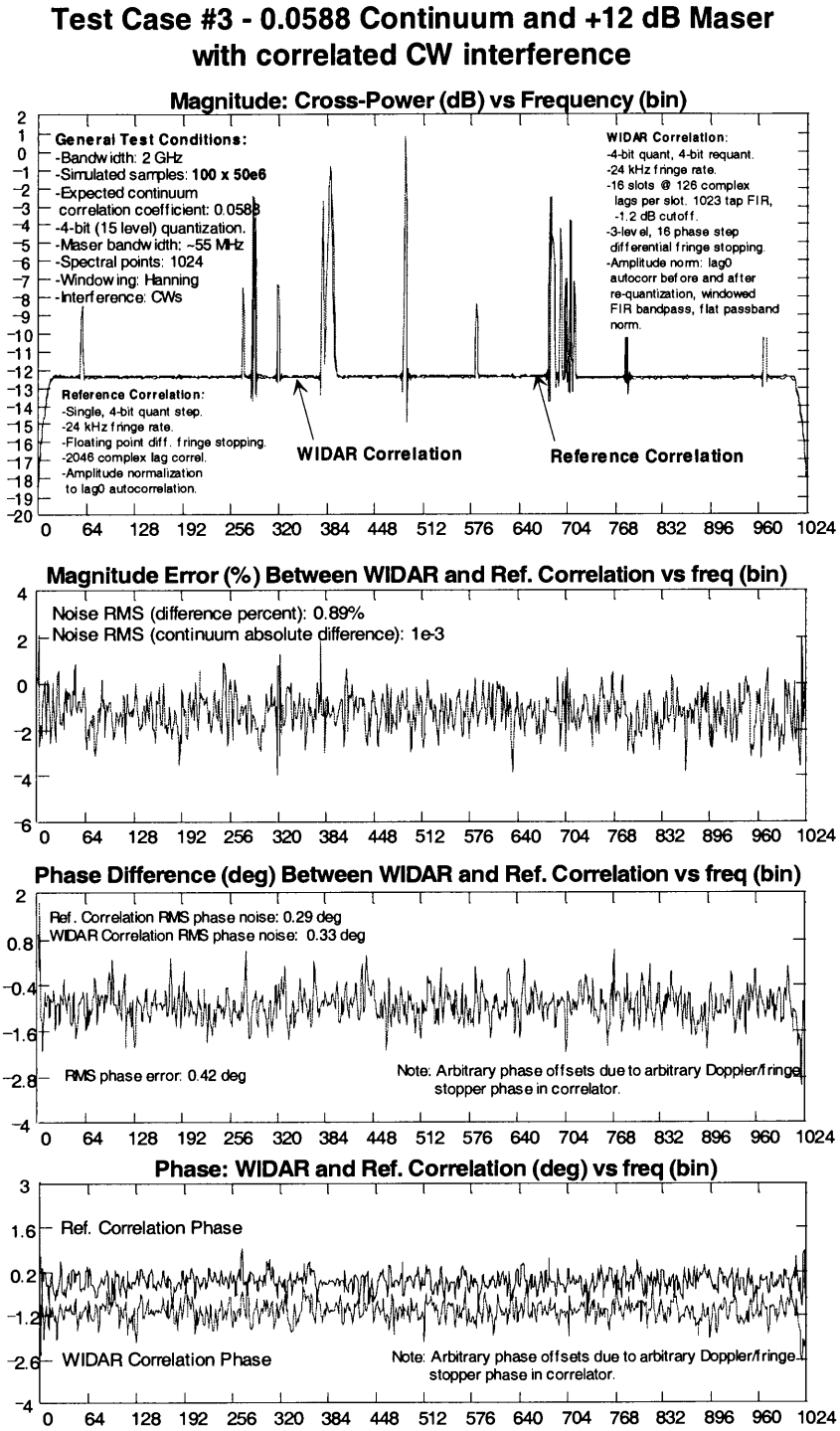


Figure 9-5 Continuum with narrowband signal and narrowband interference. WIDAR does a good job with unresolved lines that are on sub-band boundaries like the interference line near frequency bin 320. This simulation integrated 100 dumps of 50 million samples each—125 milliseconds of real-time correlator time at 4 Gs/s.

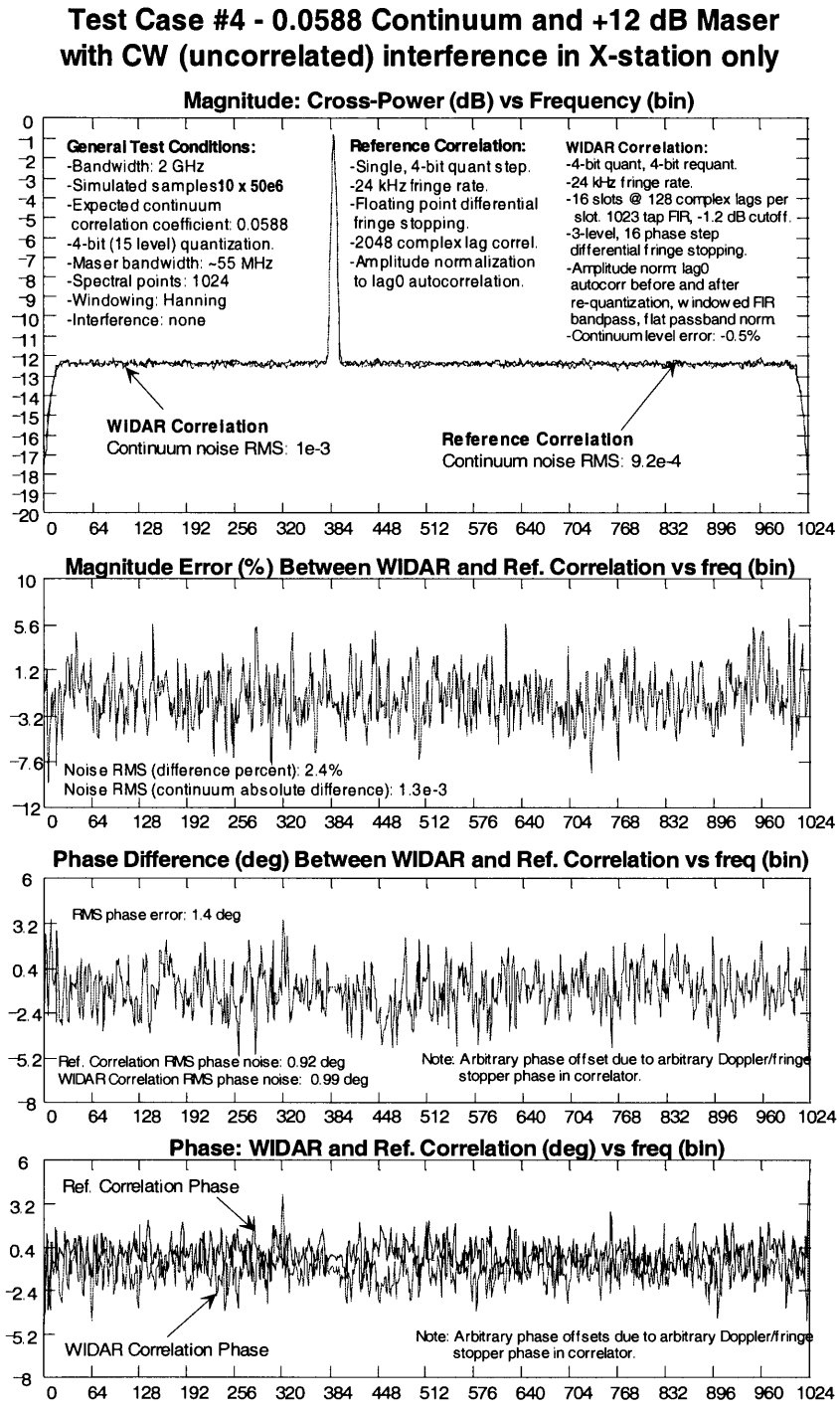


Figure 9-6 Continuum with narrowband line and uncorrelated interference (in the same locations as the previous figure) in the X-station only. This demonstrates that the WIDAR normalization equation can stitch together sub-bands even though there may be some narrowband power that does not correlate. Ten dumps of 50 million samples each were integrated for this test.

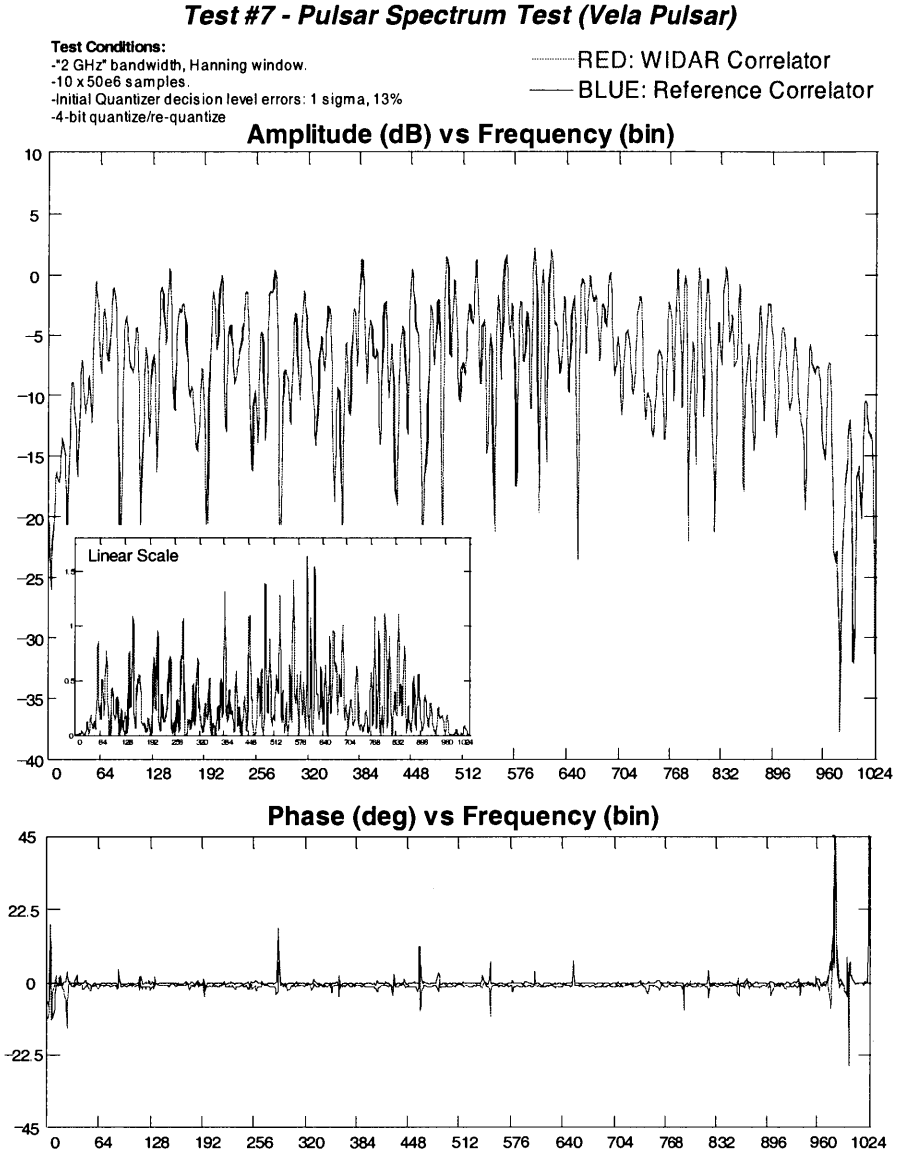


Figure 9-7 Simulation of the on-pulse spectrum from a real pulsar. The actual bandwidth of this spectrum is 16 MHz, but in the simulator it is used as 2 GHz (its all discrete-time so it doesn't matter). Significant phase differences occur where the power is weak and therefore noisy—as expected. This spectrum was chosen for simulation because of its complex amplitude shape.

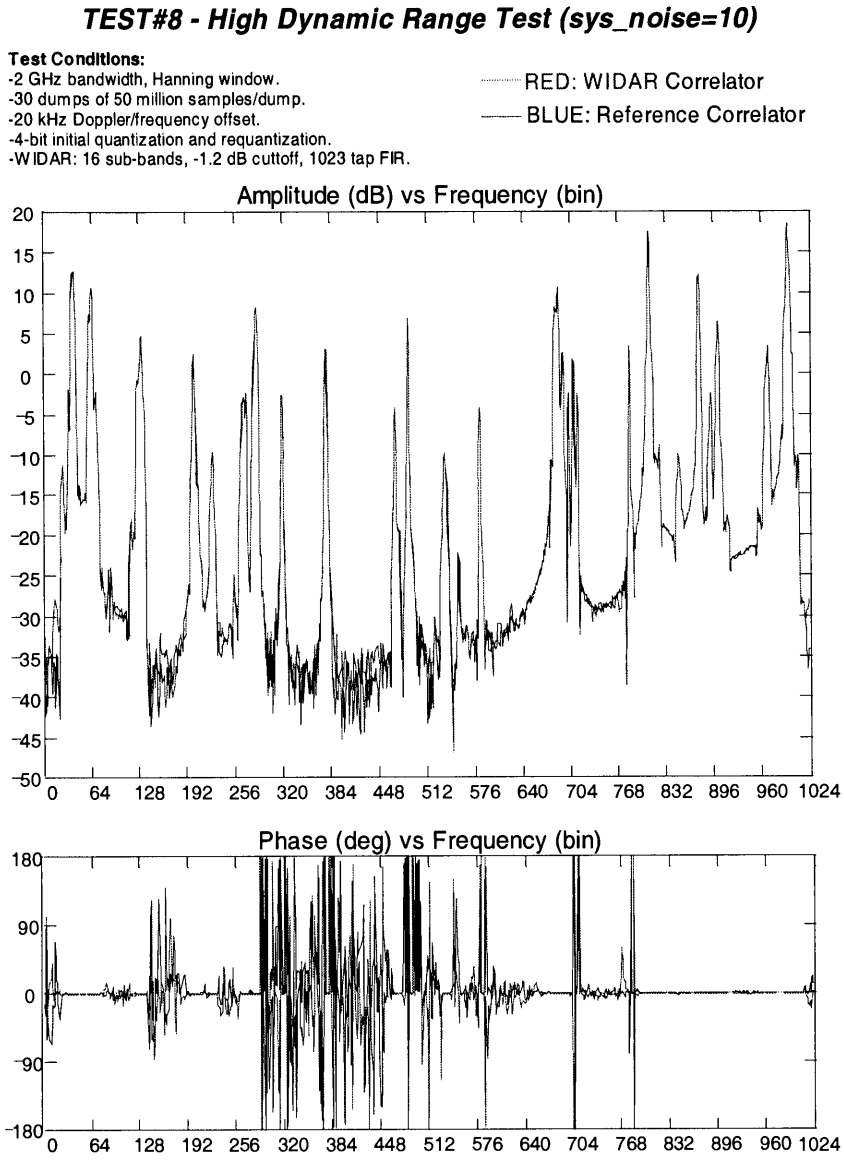


Figure 9-8 High dynamic range test with a number of (FIR filter generated) narrowband lines and interference lines (sine waves). The WIDAR correlator does as well as the reference correlator (4-bits, floating-point fringe stopping) except at about frequency bin 20. This artifact was later found to be due to digital mixer (fringe stopper) harmonic correlation with re-quantizer generated narrowband harmonics. In this test a 3-level fringe stopper was used—5-level fringe stopping would yield better results.

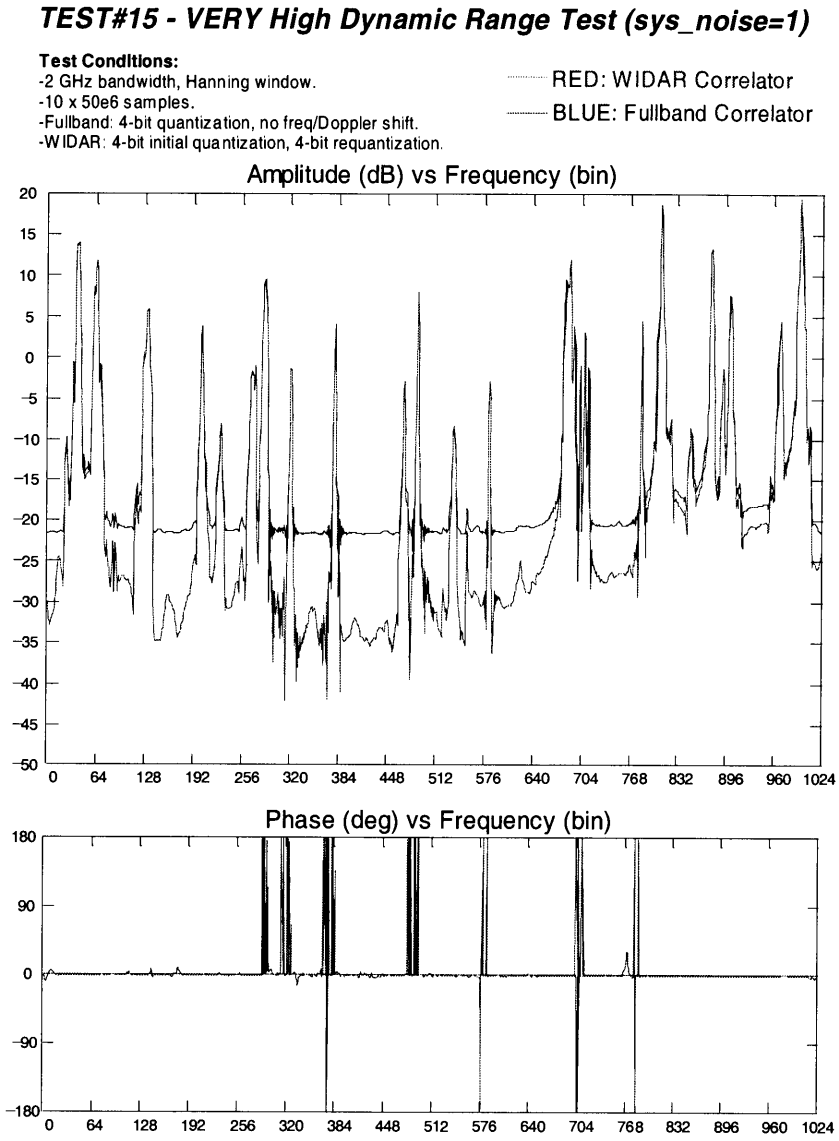


Figure 9-9 Very high dynamic range test comparing WIDAR correlator results and fullband/wideband correlator results without frequency shifting. The decorrelation effect is clear. The fullband correlation result is the best that could be achieved in this "pathological" case with a full 4-bit correlator without frequency shifting.

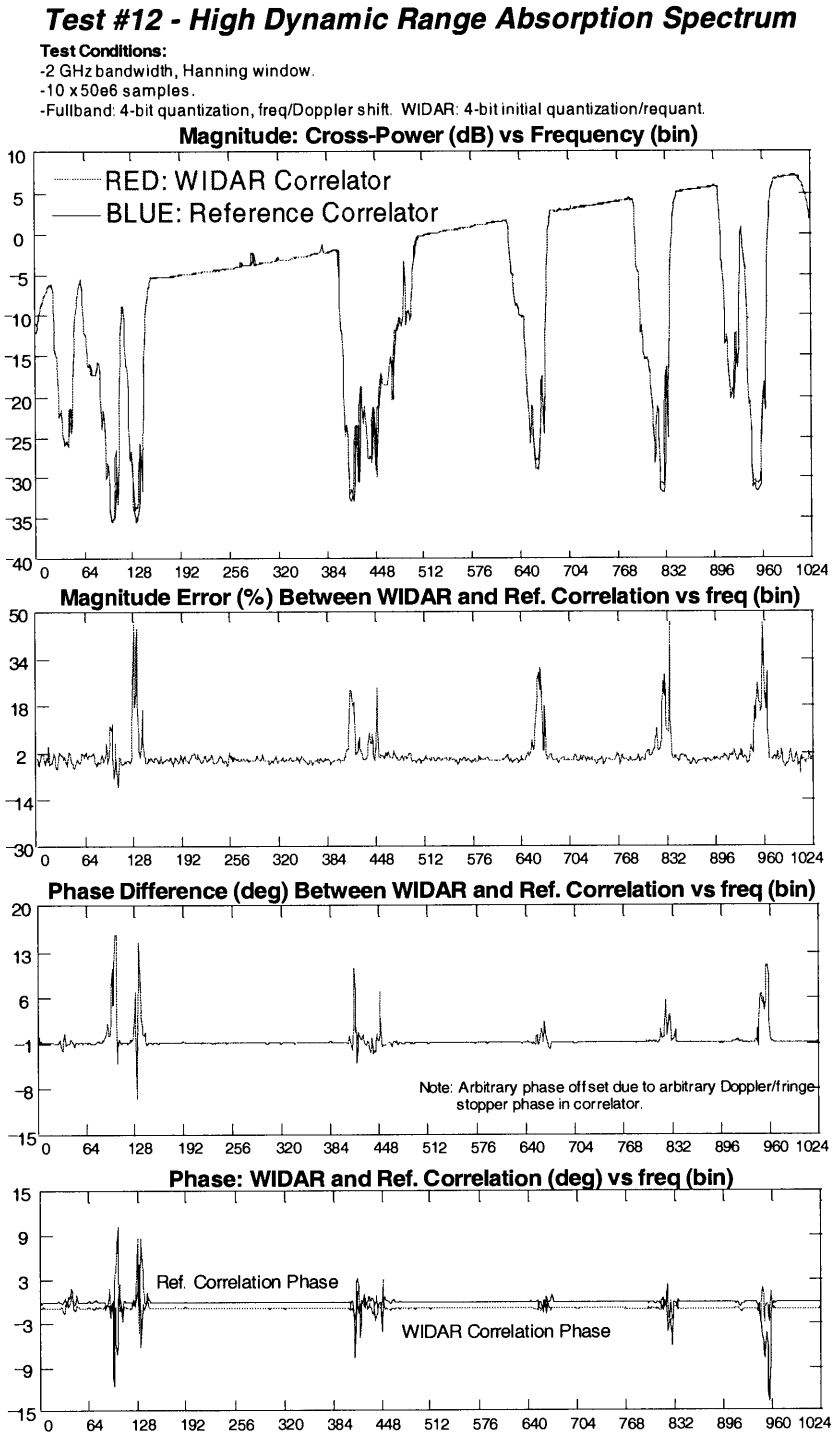


Figure 9-10 High dynamic range absorption spectrum. The spectral shape was generated with the noise generator's FIR filter. Differences between WIDAR and the reference correlator occur where the signal is weak—as expected.

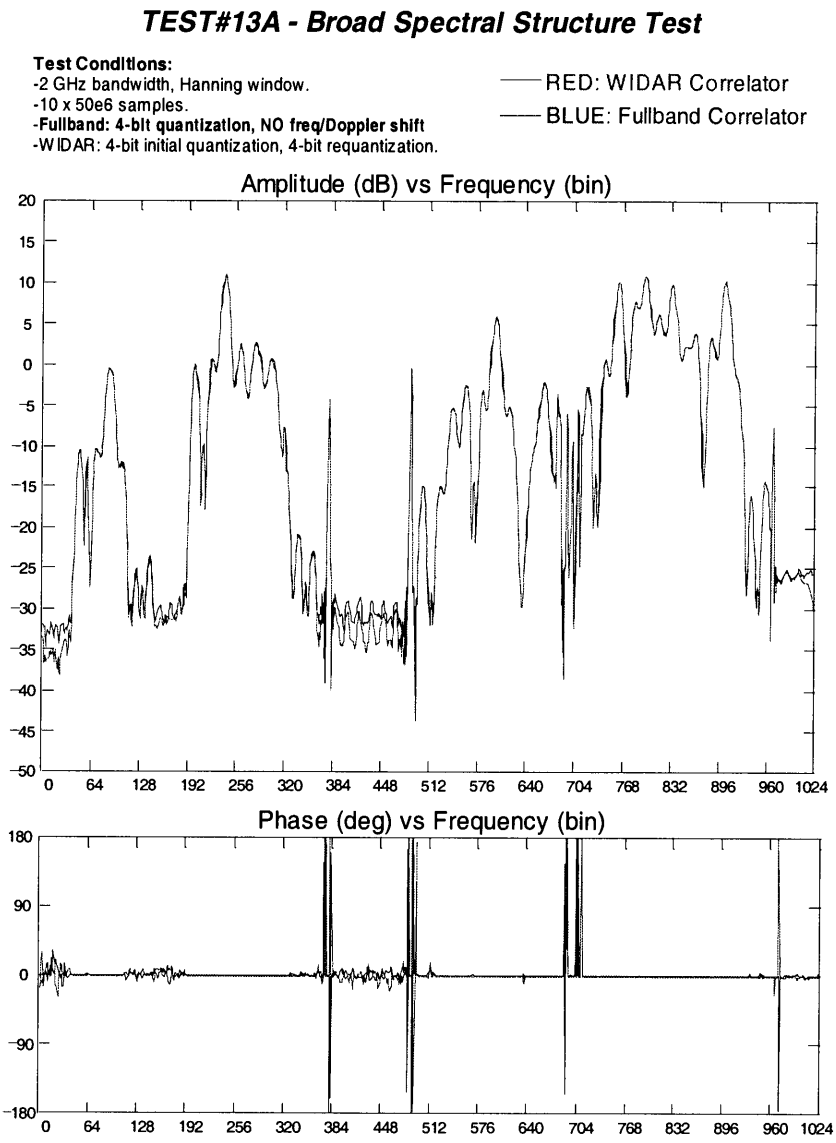


Figure 9-11 This high dynamic range and high SNR broad spectral structure test compares WIDAR correlation results with a fullband/wideband correlator with no frequency shifting. The decorrelation effect yields better results (see bins 384 to 448) but, as it turns out, not quite as good in this case as the reference correlator. Three-level fringe stopping was used in the WIDAR correlator. The rather artificial-looking structure was generated with the noise generator's FIR filter.

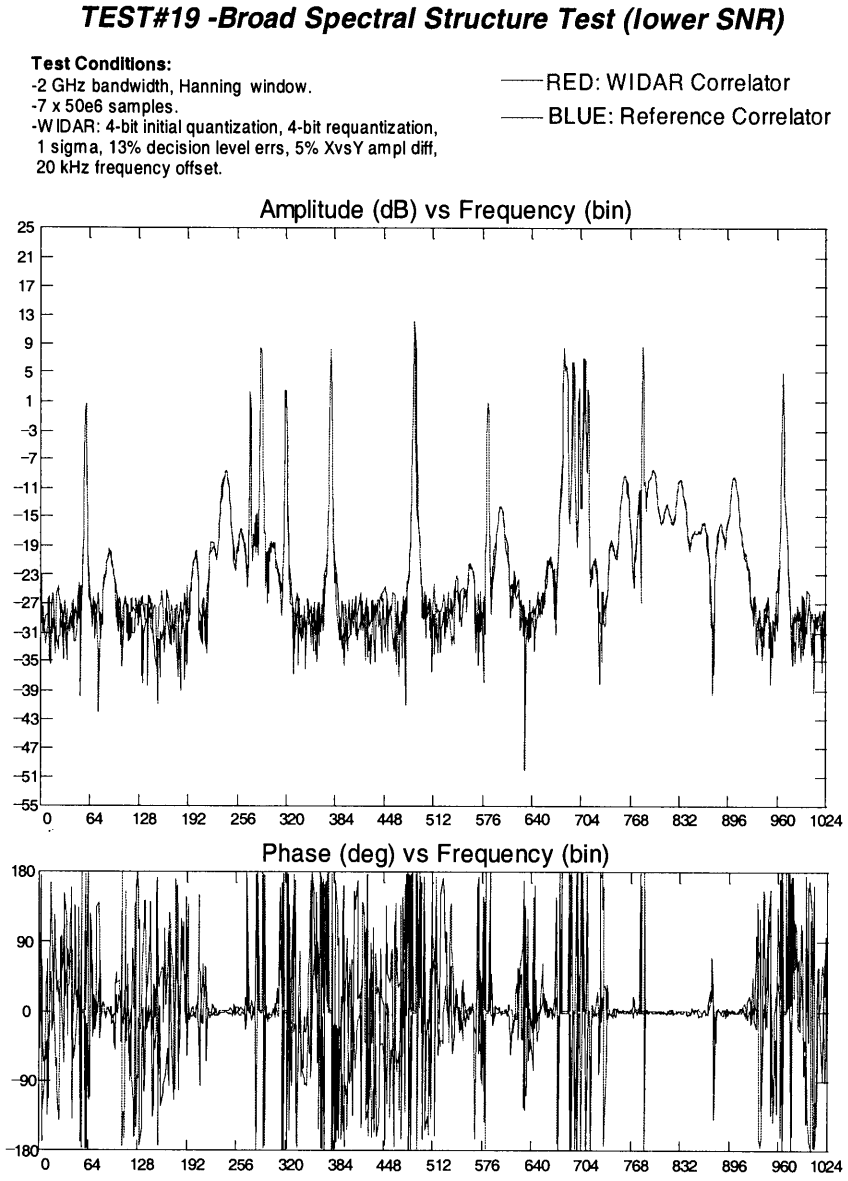


Figure 9-12 Lower SNR broad spectral structure test comparing the WIDAR correlator results with the reference correlator with floating-point fringe stopping. This plot illustrates that there are no strange effects occurring at the sub-band boundaries where there is no correlation.

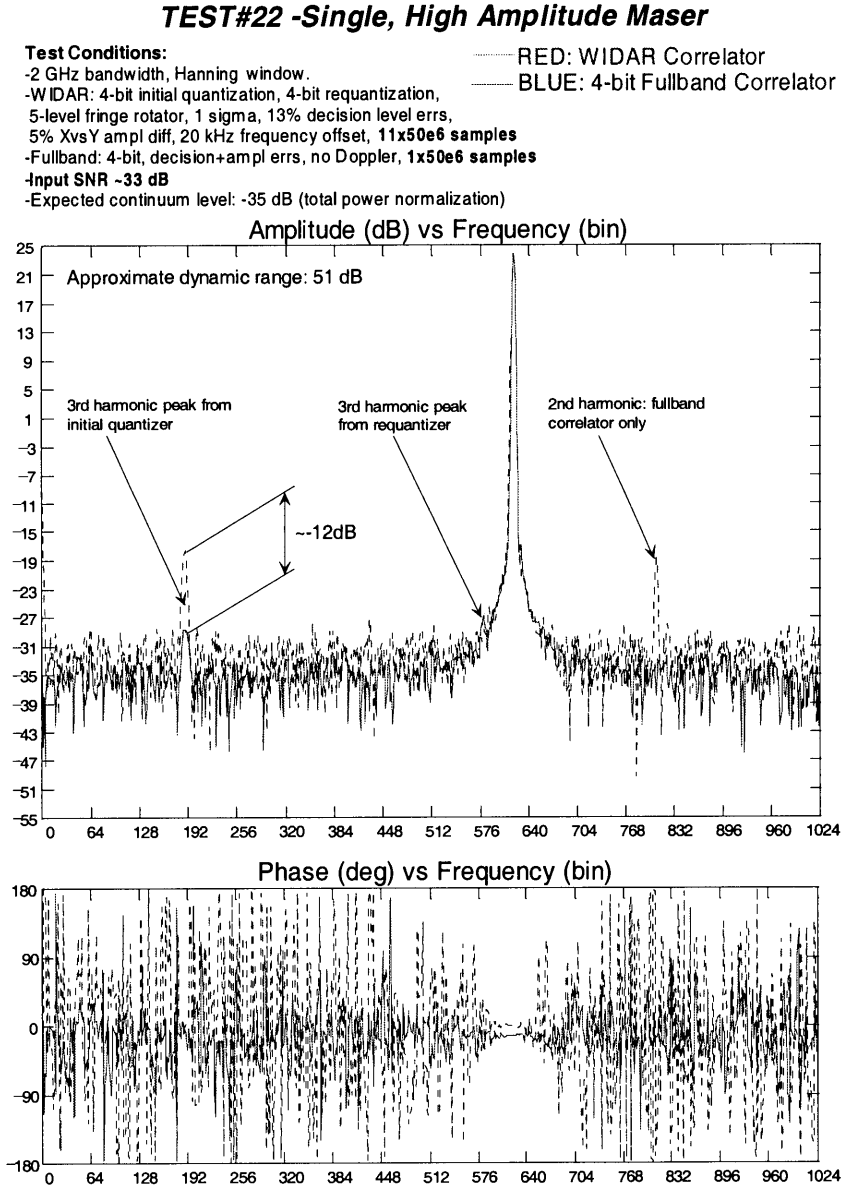


Figure 9-13 WIDAR correlator simulation results with a single "+33 dB maser". This is a test that clearly illustrates the spectral dynamic range limitation of the WIDAR correlator in the presence of powerful narrowband signals. About 90% of the total power in the band (going into the initial quantizer) is contained within this "+33 dB maser". The dotted line is simulation results for a fullband 4-bit correlator without frequency shifting. The spectral dynamic range improvement with WIDAR is clear. The "skirting" of the narrowband line starting at about –20 dB is due to an effect of the noise generator FIR filter. This effect is not entirely known but does occur with very narrow FIR bandpasses and is not predicted by the Fourier transform of the FIR's tap coefficients.

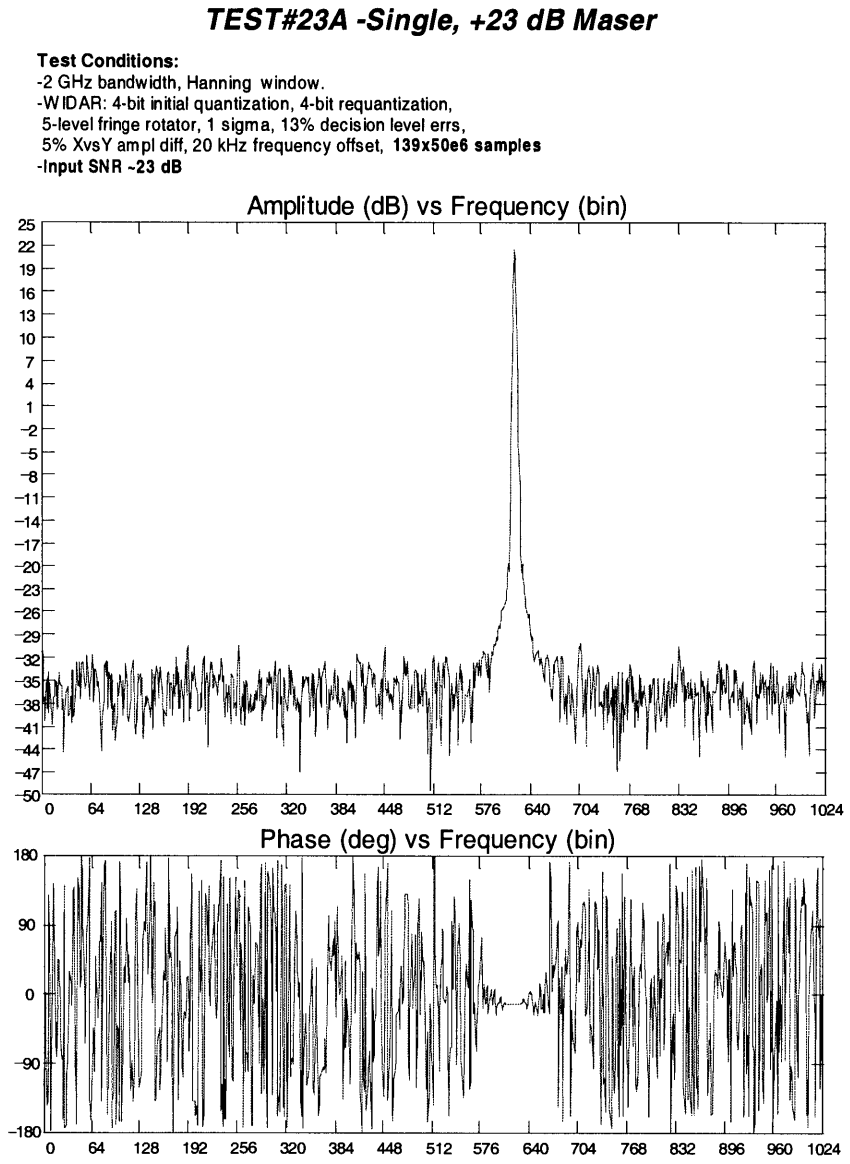


Figure 9-14 WIDAR correlator simulation results with a single "+23 dB maser". In this case, about 30% of the total power going into the quantizer is contained within the narrowband signal. This simulation ran for several weeks and, because of the relatively low SNR, the noise was only "beaten-down" to demonstrate a spectral dynamic range of about 50 dB. Harmonics of the narrowband signal are not detectable in this simulation. In this case, it would take a few *years* of simulation to drop the noise level to be able to demonstrate 60 dB of spectral dynamic range.

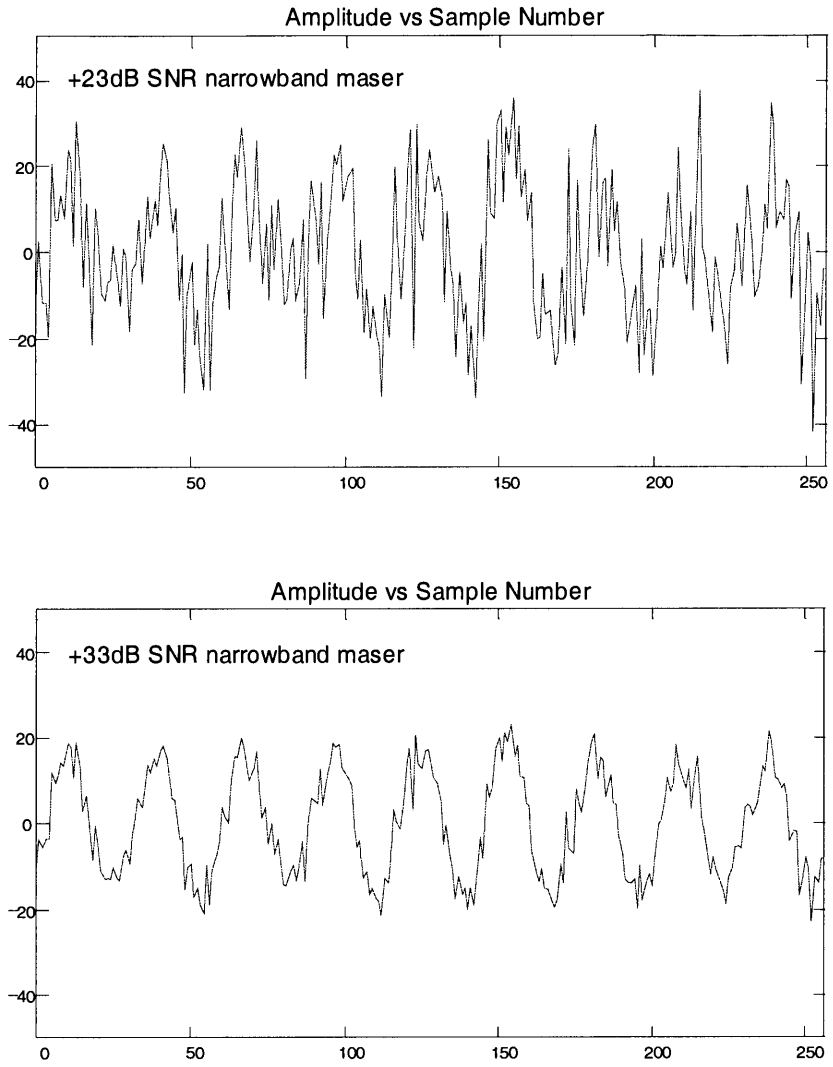


Figure 9-15 Plots of voltage waveforms (for +23 dB and +33 dB narrowband line tests) into the initial quantizer. For these plots, the line was moved down in frequency so the sine-wave pattern would be clear. The "+33 dB maser" plot clearly shows a sine pattern which will result in reasonably strong quantizer-generated harmonics.

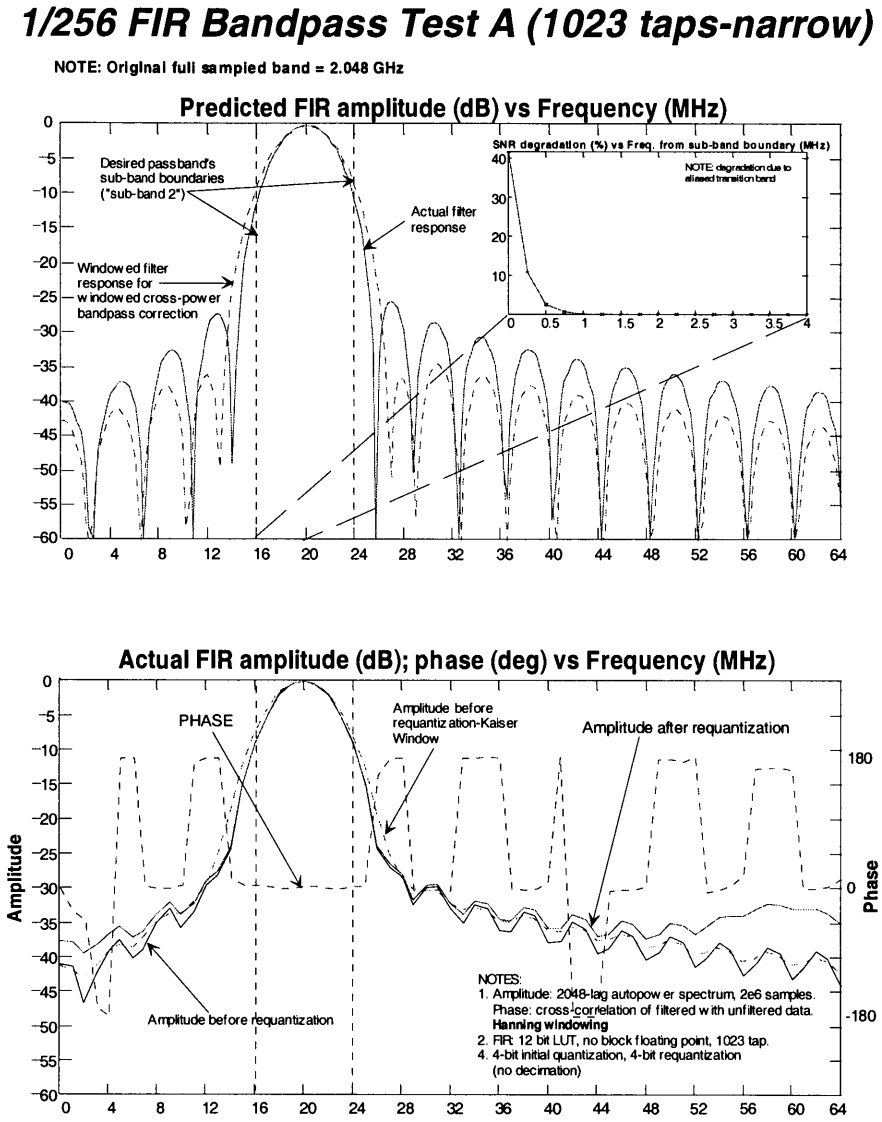


Figure 9-16 Test of 1023-tap FIR filter with a 1/256 (of 2 GHz) bandpass. The top is the predicted FIR response from the FFT of the tap coefficients and the bottom is the actual response—including phase. Note that phase is well behaved across the entire sub-band and so all data within the sub-band can be used and properly corrected in amplitude. From this test, it seems reasonable to conclude that 1/256 bandpass filtering with 1023-tap FIR filters is feasible although the sidelobe levels in this plot may be too high.

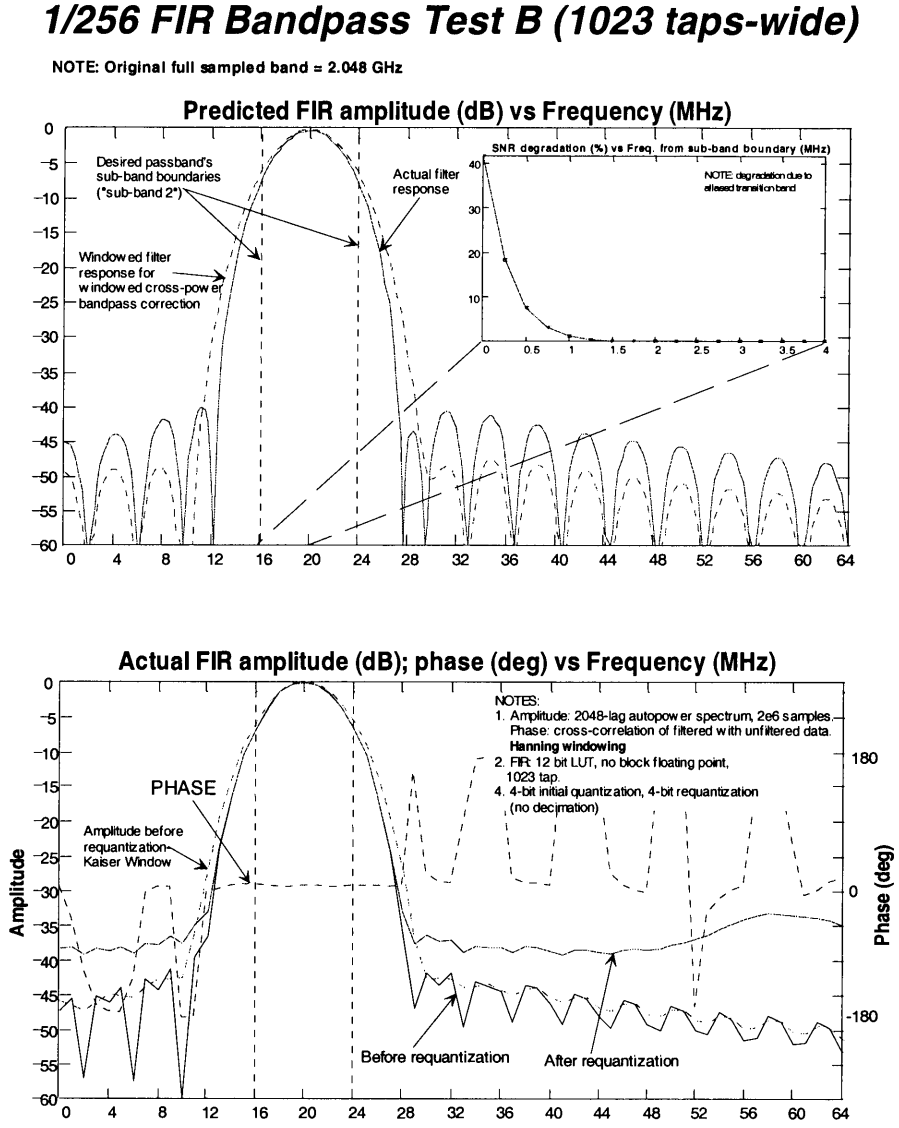


Figure 9-17 Another test of 1/256 bandpass with a 1023-tap FIR filter. In this case, the sidelobes are lower at the expense of a wider main lobe with the penalty being a larger SNR degradation region near the sub-band boundary (insert plot in top plot).

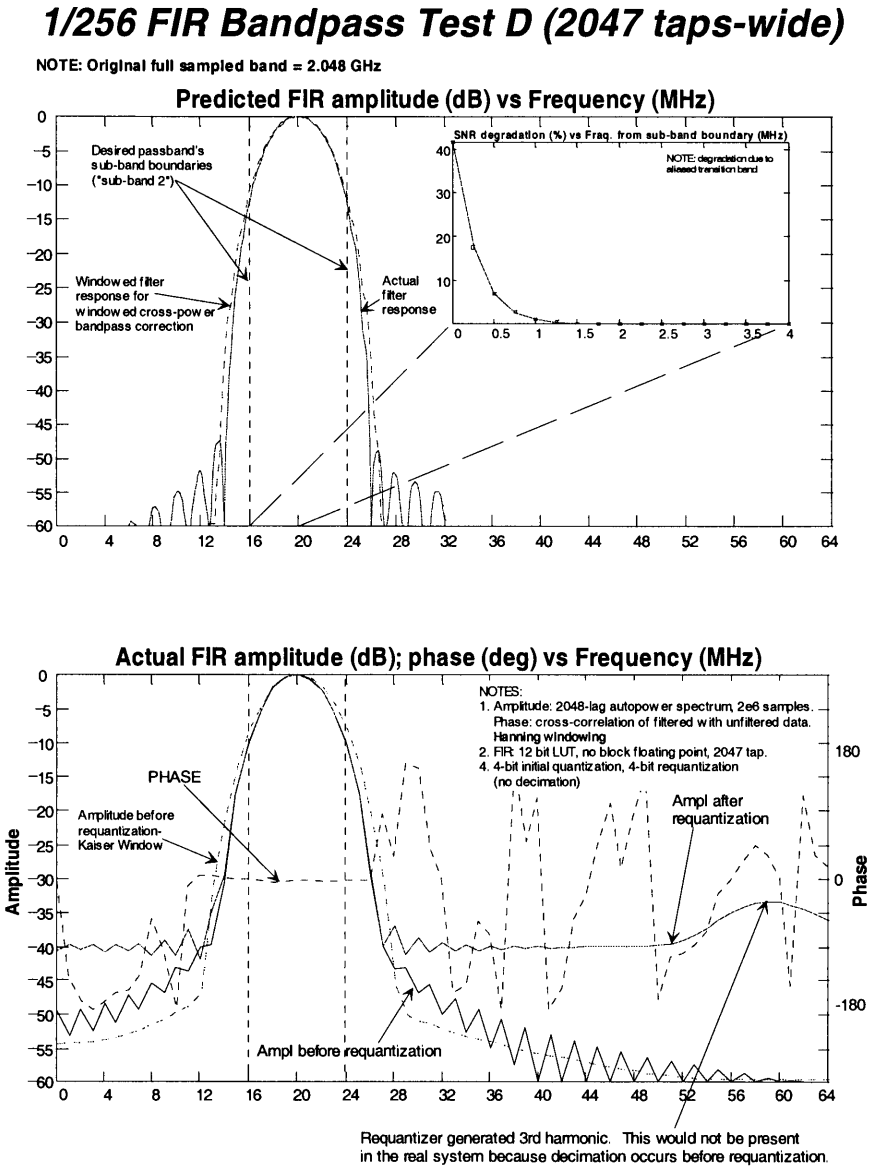


Figure 9-18 Test of 1/256 FIR filter response with a 2047-tap FIR filter. There is some improvement over the previous plot but the sidelobe levels are more than predicted by the FFT of the tap coefficients. This is the same "skirting" action noticed in the noise generator FIR filter when generating very narrow passbands and is not completely understood. Note that the sidelobes in the sub-bands adjacent to the passband will decorrelate.

10 Glossary of Terms

A/D (converter) Analog-to-digital converter. Same as a quantizer or sampler.

adaptive cancellation Refers to the process of removing unwanted signals by adding an inverted and amplitude matched version of the signal.

AGC Automatic Gain Control circuit. Generally, this refers to an electronic circuit with feedback to keep a voltage or a current at a constant value no matter what the input signal is doing. This is used in front of samplers (in radio astronomy) to maintain optimum quantization efficiency.

aliasing Analog signals that are discrete-time sampled must be sampled at a rate at least twice the signal's highest frequency otherwise part of the signal will alias, or appear as, a signal at a lower rate.

analog Referring to continuous-time and continuous amplitude signals.

AOC Array Operations Center for the VLA and VLBA in Socorro New Mexico.

baseband The output of one high-speed sampler/quantizer. Each baseband is 2.048 GHz wide and is sampled at 4.096 Gs/s. In the EVLA, there are 8 basebands per antenna.

baseband pair Consists of 2 basebands, one of which is Right circular polarization and one of which is Left circulator polarization. In the correlator it must be possible to obtain the cross-power spectrum from all 4 polarization products from a baseband pair.

baseline Referring to two antennas of an array. Also refers to the projected distance, normal to the radio source vector, between two antennas.

BGA Ball Grid Array. A chip package where the pins of the device are solder balls on its bottom surface. This requires an oven/reflow process to solder the device to the board. BGA devices have lower inductance per pin, better thermal characteristics, small sizes, and higher reliability than conventional gull-wing or J-lead surface mount devices.

complex-lag Referring to a lag with complex multiplication and accumulation. Each lag yields an in-phase and quadrature result. Only ever used for cross-correlation.

correlator A machine that correlates data. If data is from two antenna elements of an interferometer then it is referred to as cross-correlation. If from one antenna, it is referred to as autocorrelation.

correlator chip The heart of any correlator system. The abilities of the correlator chip largely determine the abilities of the correlator system. This is where correlation finally occurs.

D/A (converter) Digital-to-analog converter. Convert discrete-time digital signals to analog form.

decorrelation Referring to the drop in cross-correlation amplitude with integration time. This is a desired effect if the amplitude drop occurs for undesired part of the signal.

delay tracking Compensation by the correlator or related systems for radio-signal wavefront arrival time mismatch at the antennas. That is, generally the radio signal from space arrives at one antenna before another—an effect that must be compensated for.

DFIR Refers to a FIR filter architecture that operates on time-demultiplexed data. Bandpass filtering and decimation is done in one step yielding an efficient FIR architecture. This is not an industry-standard term—it was invented for the WIDAR application.

digital mixer A time-domain mixing function that is performed entirely with digital representation of the signals involved.

dynamic range The ratio of the strongest to the weakest detectable signal or feature in a system.

EVLA Referring to the Expansion Very Large Array (VLA) project.

fanout In the electronics industry, this refers to the ability of a device to drive several receivers.

FGA Fine Grid Array. Same as BGA only the lead spacing is smaller.

FIR Finite Impulse Response filter. This is a digital filter that continuously operates on a finite amount of time-shifting data. This filter has calculable and stable amplitude and phase characteristics.

FOTS Fiber Optic Transmission System. Nominally the system that will transfer sampled data from antennas to a central correlator site in real time.

FPGA Field Programmable Gate Array. An electronic chip whose logic can be custom programmed. Most FPGAs download a configuration program that can easily be changed—thus modifying its logic functions—without changing the chip.

fringe stopping Also known as "lobe rotation" or Doppler compensation. This removes phase wrapping resulting from delay tracking that is not performed at the radio source's original RF frequency.

FX correlator The cross-power spectrum is obtained by Fourier transforming the data (F) followed by multiplication in frequency (X).

interference Any unwanted signal that degrades the detection or processing of desired signals.

lag Refers to the element of an XF correlator where accumulation of cross or auto correlation data at a particular relative delay occurs.

Local Oscillator Single frequency that down-converts the received signal from RF to baseband. In practice there may be many Local Oscillators (LOs) in the receiving system, but to the correlator only the net result is important.

LTA Long Term Accumulator. Refers to hardware where longer time-accumulation of correlator data occurs off the correlator chip. This is done for efficiency so the correlator chip accumulators don't have to be too large.

LVDS Low Voltage Differential Signaling. A standard for high-speed, reliable, digital data transfer. Differential transmission requires two wires for the data but it is more noise-immune than single-ended transmission—especially in high-speed digital systems.

narrowband Referring to signals that are narrow in frequency width compared to the frequency space that is is being considered.

NRAO (U.S.) National Radio Astronomy Observatory with headquarters in Charlottesville Virginia.

NRC National Research Council of Canada.

phased-array When signals from an array of antennas are coherently added together to provide combined (additive) collecting area and sensitivity.

power factor Cosine of the effective phase difference between voltage and current. The smaller this number, the higher the power rating the power delivery system must be for a given power requirement. Unless power factor correction is included, off-line switching power supplies normally have a power factor of about 0.7.

quadrant architecture An alternative architecture for the correlator where all baselines and sub-bands for one baseband pair are correlated in one conceptually separate subsystem.

quantization noise Output noise generated by a sampler from quantizing the continuous input amplitude to a discrete number of levels.

quantizer Same as a sampler.

radar mode This is a mode supported in the correlator that provides at least 1 Hz spectral resolution over tens to hundreds of kHz of bandwidth. In the proposed correlator, this will be provided with a combination of two-stage FIR filtering and recirculation. This mode was designed to meet the requirement for inter-planetary radar.

recirculation Refers to the ability of an XF correlator to make maximum use of available correlation resources when processing signals that are slower than the correlator is capable of. Chunks of data are burst through the correlator at the highest rate, with

each burst at a different relative delay so that the final result is over a much larger delay (lag length) yielding improved spectral resolution.

RMEM Acronym for recirculation memory.

sampler Analog to digital converter. This device quantizes the analog baseband signal in time and amplitude.

SCSI Acronym for Small Computer Systems Interface. There are several versions of SCSI, among them the Ultra-Wide SCSI2 interface currently in use mainly for tape and disk storage devices.

SDRAM Synchronous Dynamic RAM. Same as SRAM only dynamic RAM is used. Dynamic RAM is more efficient—resulting in larger memories—than static RAM except that there are memory cell refresh requirements and CAS latency that must be met. SDRAMs are used for large core-memory applications such as in PCs.

SIMM Single In-line Memory Module. A small circuit board with many memory chips on it that is convenient to plug into a socket on a board. This package achieves high memory density.

SNR loss Signal-to-Noise Ratio loss. Refers to a receiver or correlator system's loss of sensitivity.

spectral dynamic range The ratio of the strongest spectral feature to the weakest (unambiguously) detectable feature in a frequency spectrum.

spectral resolution The width of one spectral frequency channel. Calculated as the total spectrum bandwidth divided by the number of spectral channels.

SRAM Synchronous (static) RAM. For speed, reads and writes are synchronized to a clock.

station Used interchangeably to refer to an antenna but normally refers to associated downstream hardware from the antenna itself.

sub-array A collection of antennas that is smaller than the entire array of antennas. These antennas would generally share attributes such as pointing, observing frequency, and bandwidth so that cross-correlation of antennas within the sub-array yields useful results.

sub-band Output of one FIR filter which has bandpass-filtered and then decimated the data from one baseband.

sub-band correlator This is a correlator (rack) that processes a sub-band from all basebands from all antennas on all baselines. This is effectively a 128 MHz bandwidth correlator.

time-burst demultiplexing A high-speed time series of samples is split into streams of bursts of contiguous-time samples.

time-demultiplexing A high-speed time series of samples is split into multiple streams by simply re-directing samples. This is a method of slowing down the digital clock rate while still being able to transmit all of the high-speed data.

transition band That part of a band where a transition is occurring from one level to another. In this document this is normally the transition from the pass-band to the reject-band.

VLBI Very Long Baseline Interferometry. Because of the separation between antennas, data is recorded on magnetic tape, along with atomic clock (Hydrogen Maser) time signals, and shipped to a central site for cross-correlation.

WIDAR Acronym for Wideband Interferometric Digital ARchitecture. This is the proposed architecture for the EVLA correlator.

wideband Colloquially referring to a 2 GHz baseband.

XF correlator The cross-power spectrum is obtained by multiplying in frequency (X), followed by Fourier transform (F).

11 Index

C

D

E

F

G

H

I

Q

R

S

T

U

V

W

X

Z