ALMA Memo N° 476

Enhancing the Baseline ALMA Correlator Performances with the Second Generation Correlator Digital Filter System

B.Quertier[1], G.Comoretto[2], A.Baudry[1], A.Gunst[3], A.Bos[3]

[1]Observatoire Aquitain des Sciences de l'Univers, UMR 5804, BP89, 33270 Floirac, France
[2]Osservatorio Astrofisico di Arcetri, Largo E. Fermi 5, 50125 Firenze, Italy
[3]ASTRON, P.O. Box 2, 7990AA Dwingeloo, The Netherlands

2003-August-27

Abstract : *Characteristics of the digital filter bank initially proposed for the Second Generation Correlator (2GC) are summarized and we show that this filterbank can be adapted to the Baseline ALMA Correlator design to enhance the number of spectral channels and its overall performances. Details on required Field Programmable Gate Array (FPGA) resources, power consumption and cost are given and we discuss implementation of the digital filters in HardCopy devices to further improve performances and cost.*

Keywords : *correlator, digital filtering, filter architecture and design*

1. Introduction

The European component of the ALMA Correlator Integrated Product Team has worked on a new correlator concept named Digital Hybrid XF whose architecture is intermediate between the XF and FX architectures (see [1] and ASAC reports). One basic idea in this design is to diminish the required cross-correlation resources by splitting the input bandwidth into several sub-bands. The digital filter bank thus plays a central role in this design. No time multiplexing is required when the sub-band bandwidth matches the correlator clock frequency. In the case of ALMA with 125MHz clock rate each 2GHz baseband would thus be sub-divided into 32 sub-bands of 62.5MHz each. The feasibility study of the hybrid design now called Second Generation Correlator was concluded in 2002 [2] and the European team proposed a digital 2-stage filter bank architecture with digital oscillator and mixer for sub-band overlapping and tunability across 2GHz (see e.g. [3]).

The parallel array architecture of the Baseline ALMA Correlator could, as well as in the 2GC case, be used to process all 32 sub-bands from each ALMA baseband. This would enhance the number of available channels in spectral line projects [4]. In this paper we give some advantages and describe the principles of the 2-stage filter architecture initially proposed for the 2GC, and we show that this filter bank can be adapted to the design constraints of the Baseline ALMA Correlator to greatly enhance its performances.

2. Simple Examples of Science Enhancement

For each antenna in the array and both polarizations the 2GC filter bank splits each 2GHz baseband into 32 independent sub-bands tunable across the baseband. Furthermore, by using a filter resource recirculation scheme, each sub-band bandwidth can be chosen to be 62.5MHz, 31.25MHz, 15.63MHz or 7.81MHz (filtering mode 1, 2, 3, or 4). This possibility together with independent sub-band filters opens a wide range of scientific observing modes.

In Table 2.1 we give the spectral resolutions which can be achieved when the 2GC scheme is implemented in the Baseline ALMA Correlator ; the bandwidth of each sub-band filter is assumed to be equal across the total bandwidth. With the Baseline filter design there are 64 spectral points available per baseline across 2GHz for all 4 polarization products whereas we have now 32 times more spectral channels and the possibility to use sub-bands smaller than 62.5MHz. With 62.5MHz per sub-band and 64 spectral points we thus obtain 976.56kHz resolution. Note that all resolutions in Table 2.1 may still be improved by a factor of 2 if no cross-polarization products are useful. Resolutions achieved with the filtering modes 2, 3 and 4 are possible only if the correlator would be run at a real or virtual clock lower than the nominal 125MHz in order to process bandwidths smaller than 62.5MHz.

Resolution (kHz)				
Total correlated bandwidth (MHz)	Sub-band bandwidth (MHz)			
	Mode 1 : 62.5	Mode 2 : 31.25	Mode 3 : 15.63	Mode 4 : 7.81
2000	976.56			
1000	488.28	488.28		
500	244.14	244.14	244.14	
250	122.07	122.07	122.07	122.07
125	61.04	61.04	61.04	61.04
62.5	30.52	30.52	30.52	30.52
31.25		15.26	15.26	15.26
15.63			7.63	7.63
7.81				3.81

Table 2.1. Spectral resolution for 2-stage filter in combination with the Baseline Correlator (see text for limitations on modes 2, 3 and 4)

An interesting possibility with independent tunable sub-bands consists in observing specific spectral windows with different resolutions. An observing 'zoom' mode can be obtained when several contiguous sub-bands in the original spectrum are synthesized with different sub-band bandwidths and then analyzed with different resolutions. Examples of zooming capability are illustrated in Fig. 2.1 and 2.2 where the original spectrum (Fig. 2.1) is further filtered with 977kHz and 244kHz resolutions.

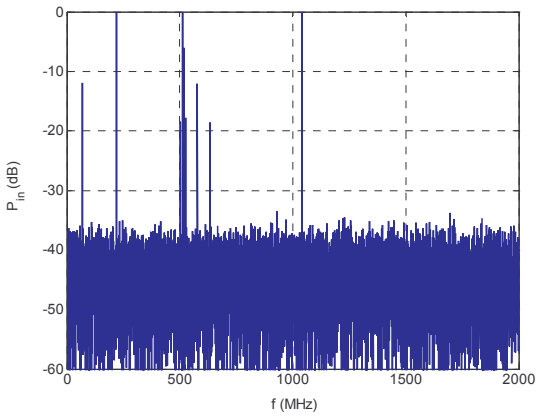

Fig. 2.1 : Input baseband with lines at 70.3, 218.8, 503.9, 511.7, 519.5, 527.3, 578.1, 632.8 and 1039.1MHz.

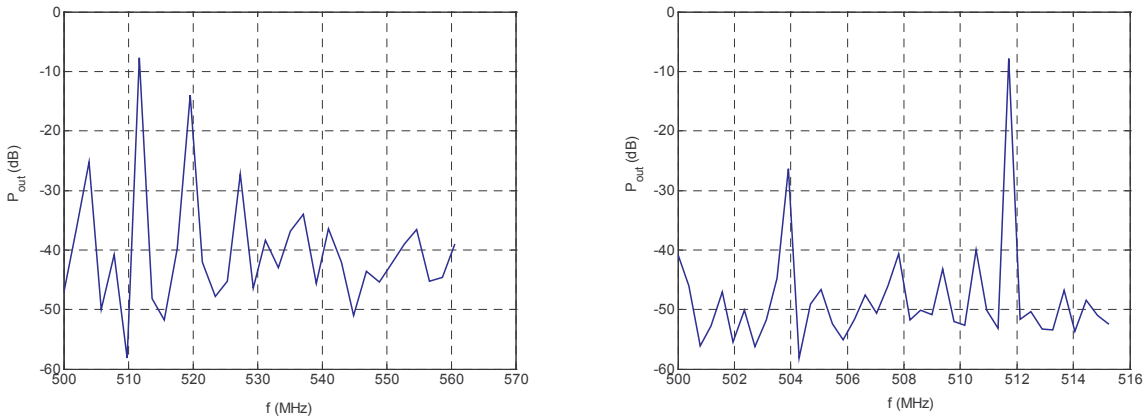

Fig. 2.2 : Filtered sub-band with 62.5MHz bandwidth and 977kHz resolution (left panel) and filtered sub-band with 15.63MHz bandwidth and 244kHz resolution (right panel).

Finally we note that the availability of independent sub-bands, tunability and zooming allow for simultaneous continuum and high spectral resolution observations.

3. Digital Filter Schematics and Modelizations

3.1. System Description

Each 2-stage filter board is designed to be 'plug compatible' with the Baseline Correlator filter board. The basic functionality of the latter is retained, using the same system architecture. Each board implements 32 filters for one baseband and one polarization. Each filter is assigned to one correlator plane. Since all filters are identical, independent and fully programmable, each correlator plane can analyze any portion of the input baseband or different planes can analyze the same data for increased resolution.

The board receives 3-bit digital samples from the antenna electronics backplane, parallelized in groups of 32 consecutive samples. Two FPGAs, whose design is taken from the original board, derives the fine delays (i.e. adjusts delays in steps of one 4GHz clock, up to 32 steps), and sends them to an array of identical filter chips. Each filter selects a slot of adjustable width and position in the input band and converts it to baseband, requantized to 2 bits. Each filter output is presented to one of the 32 outputs of the filter board. A bypass mode should be available in order to transfer directly the input samples to the output without any processing except power measurement. Chip programming and personality downloading is performed using a CPLD derived with minor modifications from the standard CPLD2 chip used in the Baseline Correlator.

The conceptual design for one individual filter is described in detail in [5] and a simplified diagram is shown in Fig. 3.1. The input signal is processed using a complex time-multiplexed digital mixer driven by a DDS Local Oscillator, in order to have the central frequency of the desired sub-band translated to frequency zero. The resulting signal is complex and is processed by a 2-stage filter. This filter is implemented as two identical 2-stage filters operating on real and imaginary portions of the signal. A first rough low-pass filter provides enough selectivity to allow decimation by 32 of the sample frequency without significant aliasing, then a second filter synthesizes the desired final bandpass. Thereafter, the signal is decimated and converted to real by shifting it up by ¼ of the final frequency.

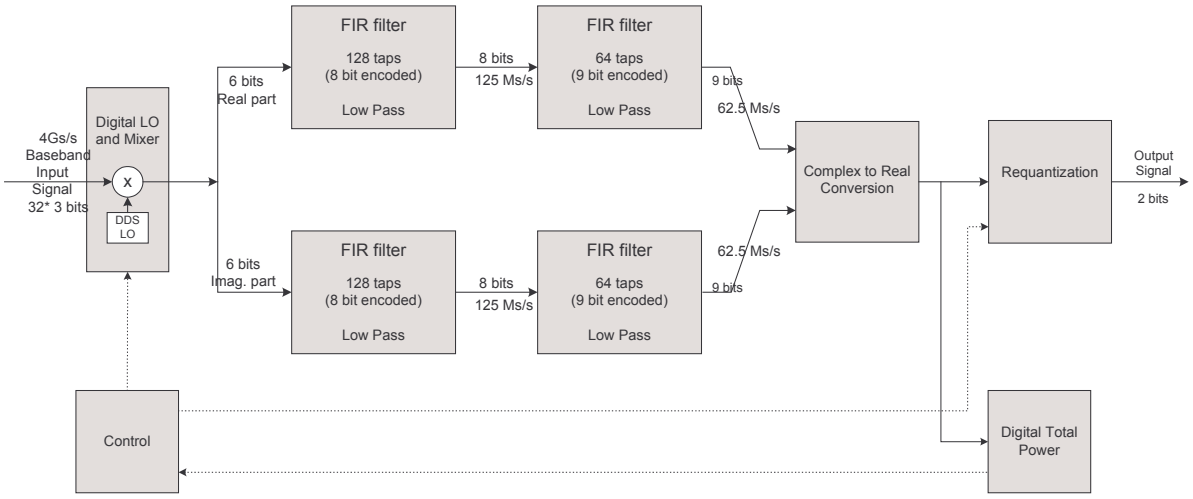

Fig. 3.1 : Simplified diagram of the 2-stage digital filter. The input signal is frequency translated using a complex mixer, and then filtered by a 2-stage lowpass filter. The output signal is converted to real, and requantized after appropriate rescaling. A digital total power meter allows for subsequent calibration. The number of bits shown in the diagram are indicative of what has been implemented in this work (see Section 4).

Fig. 3.2 shows the effects of the signal processing in the frequency domain. The first filter passband extends from -1/64 to +1/64 of the input band in the frequency range [-31.25MHz, 31.25MHz], while the stopband begins at +/-93.75MHz. Stopband attenuation depends on the actual implementation parameters, with 128 taps and 8 bit coefficient resolution, the stopband attenuation is 47dB. After 1/32 decimation, the complex signal conveys useful information in a band of 62.5MHz wide, centered on frequency zero, while the transition band is folded in the two external regions of the spectrum.

The second filter determines the actual final bandshape, and can be programmed for different final bandwidths and decimation factors from 62.5MHz (no decimation) to 7.81MHz (1/8 decimation). The second filter has 64 taps when operating at the nominal 1/32 decimation. This corresponds to 2048 taps at the original input frequency of 4 GS/s. The number of taps increases at higher decimation. For a maximum decimation factor

of 8 (global decimation factor of 256) corresponding to a sub-band bandwidth of 7.81MHz (filtering mode 4) the number of taps increases up to 512. The second filter passband extends from -15/32 to 15/32 of its output band. The final sub-band bandwidth is then, after complex to real conversion, 15/16 of the nominal sub-band bandwidth (independent from decimation factor) and the in-band ripple is about 0.15 dB. In this way, a continuous uniform coverage of the input bandwidth is possible with an overlap of 1/16 (corresponding to 4 spectral points for 64 points per sub-band) between adjacent sub-bands. This is discussed in Section 3.2.3.

The complex output is multiplied by $\exp(2\pi i\ \nu_o\ n/4)$, where ν_o is the final sampling frequency ; thereafter the real part of the result is taken. The exponential has values +/- 1 and +/-i, thus the multiplication corresponds to selecting the real or imaginary part of the filter output. Real samples are used only on even cycles and imaginary samples on odd cycles ; therefore, only alternate samples have to be computed in the second filter stage. After conversion to real, the digital total power is computed and the signal is requantized to 2 bit values with the correct statistics required by the correlator.

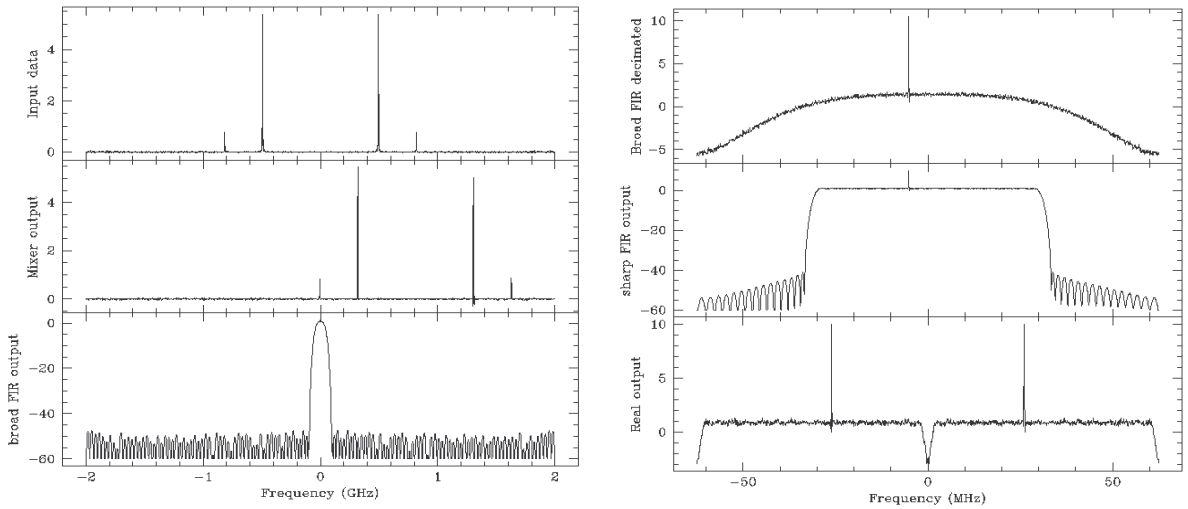

Fig. 3.2 : Signal processing in a 2-stage tunable filter. Left panel, from top : original signal ; signal after conversion by the complex mixer ; signal after first filter (undecimated). Right panel, from top : signal after first filter (decimated) ; signal after 2nd filter ; signal converted to real. The vertical scales are logarithmic (dB)

3.2. Signal Processing and Sub-band Stitching

The hybrid correlator produces 32 separate cross-product spectra per baseline which have to be separately calibrated and corrected for instrumental effects and then combined together. Steps of signal processing are :

- *a)* Correction for quantization. The signal is quantized several times during sample processing. The most severe quantization steps occur at the 2GHz baseband sampling stage and at the digital filter output (respectively 3 bit and 2 bit encoding). Less severe requantization effects occur after the DDS mixer and between the first and second stage filters.
- *b)* Tapering of the cross correlation function for apodization of the cross spectrum.
- *c)* Correction for individual channel response. This correction takes into account effects due to ripple in the digital filter, aliasing, and distortion due to the tapering introduced in the previous step. The correction includes also a data dependent component because the channel response depends on the slope in the cross spectrum.
- *d)* Overall spectrum denormalization. Each channel is rescaled prior to final requantization. Total power information collected at the filter output is used to re-align the cross-product spectra to the same overall gain. This problem is not considered here.

3.2.1. Correction for Quantization

The relationship between the cross-correlation of two digital signals, quantized to a finite number of independent levels, and the cross-correlation of the two original unquantized signals can be derived analytically. It is much more problematic to find out such a relationship when the signal undergoes several quantization stages. Therefore, the corresponding relationship has been found experimentally from simulation of artificial data.

3.2.2. Tapering and Fourier Transform

Tapering is usually adopted to prevent spectral leakage from adjacent channels, at the expense of a reduced spectral resolution. In a hybrid correlator, a robust tapering is essential in order to prevent (or reduce) edge effects at the extreme channels of each sub-band. This is necessary as these channels typically convey useful information, and may fall in any portion of the observed spectrum, even on the spectral feature of interest. Hanning tapering is assumed in this report.

The Fourier transform is usually computed on frequency points that are integer multiples of the channel separation. This means that the first channel corresponds to frequency 0 (DC component of the signal) with no phase information. Also, the last channel (at the Nyquist frequency) has no phase information. It is more convenient to compute the Fourier transform on semi-integer multiples of the channel separation thus obtaining complex only and equidistant channels with same nominal bandwidth.

3.2.3. Correction for Passband Shape

The cross-product spectrum R_j evaluated for channel j is given by the integral $\int R(\nu)P_j(\nu - \nu_j)d\nu$, where $R(\nu)$ is the actual cross-product spectrum and $P_j(\nu)$ is a function, different for each channel j and for the real and imaginary parts of the cross spectrum, that depends on both the filter response and on the adopted tapering function. For a more detailed discussion, see [6].

The global effect of the integral is, at first order, to multiply the "true" cross-spectrum $R(\nu_j)$ by a quantity a_j and to evaluate R at a frequency that is offset by a small quantity m_j. Both a_j and m_j are easily computed from $P_j(\nu)$. Both quantities exhibit a ripple corresponding to the ripple in the passband filter, and are ill-behaved near the band edges. Since they are different for the real and imaginary parts of the signal, phase errors occur near the sub-band edges and in the presence of strong gradients in the spectrum.

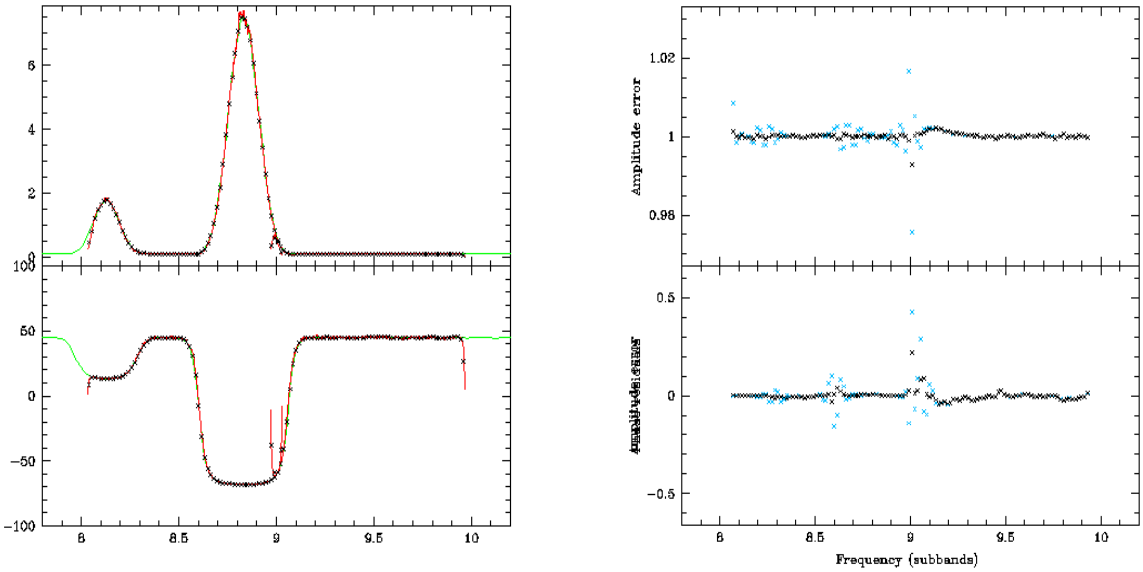

Fig. 3.3 : Stitching of two sub-bands (simulated data). Top panel for amplitude, bottom panel for phase. Horizontal scale is expressed in units of one sub-band. The original signal is shown in green. The computed cross-product spectrum is shown as black crosses. Red line represents the computed cross-product spectrum at much higher resolution. Bands overlap by 4 channels. On right panel, residuals are shown. Blue points represent data corrected only for static channel response , a_j, while black points include correction for the shift in channel position, m_j.

In principle, it is possible to recover $R(\nu_j)$ from R_j inverting the above relation, e.g. from $R(\nu_j)=(R_j - R'_j m_j)/a_j$, where R'_j is an estimate of the derivative of $R(\nu_j)$. In practice, these relations are always approximate and near the sub-band edges the first order approximation is no longer accurate enough. Because of all of these problems, in the presence of spectral features, the above relation gives erroneous values of $R(\nu_j)$ up to 10-20% and phase errors up to 10 degrees for the extreme channels. These errors are significantly reduced when the sub-bands overlap and by discarding the more troublesome external channels. With an overlap of 2 channels (97% usable bandwidth for 64 channels per sub-band) these errors are reduced to 3-5% in amplitude and 2-3 degrees in phase, while with 4 channels overlap (94% usable bandwidth), amplitude and phase errors of less than 1% and 0.3 degree are reached (Fig. 3.3).

In this chapter we have considered 64 channels per sub-band and two filter shapes (one optimized for no overlap and the other for an overlap of 1/16 of the sub-band bandwidth). Increasing the number of spectral channels, the situation is worsened in the non-overlap case, and improved using an overlap.

4. Implementation of 2-Stage Filters

4.1. Technology Specifications

A 2-stage digital filter prototype was designed with existing FPGAs to demonstrate the feasibility and functionality of the architecture described in Section 3 and selected by the European team for the 2GC. The products adopted in our study are the highest density FPGAs from Altera. This chip family, named Stratix, is based on 0.13µm CMOS technology. It provides several types of hardware resources. Logic Elements (LEs) correspond to the smallest granularity in the chip. They basically consist of a small memory used as a Look Up Table (LUT) and a register. Any logic function can be implemented cascading these LEs. The Embedded System Blocks (ESBs) form a memory tri-matrix. Configurable RAM blocks of 512 bits (M512), 4K bits (M4K) and 1M bits (M1M) are available to implement most of the storage functions. Digital Signal Processing Blocks (DSPBs) provide hard multipliers, accumulators, adders, registers and multiplexers which can be interconnected to form the DSP function core. Optimizing the filter system – namely minimizing the cost and the power consumption and achieving the speed requirement – requires optimum use of available resources during the design phase. Some details on resource usage for each module of the 2-stage filter system is given in Section 4.3. Further optimization in terms of power consumption and cost is feasible with Altera products by porting the FPGA design to HardCopy devices ; details are given in Section 4.6.

4.2. Design Flow

Functional description of the digital filter is performed using VHDL (except for some low hierarchical levels). Thus, the design work done for this prototype can easily be used for future 2GC developments because it is not related to a specific technology. In addition, VHDL allows us to create an efficient simulation environment. For each system module that we wish to simulate, a VHDL simulation file is created. This file implements the module as a Device Under Test (DUT), checks the reading of test vectors stored in ASCII files (input and reference output sample files) and automatically compares, one sample by one sample, the output generated by the VHDL model of the DUT to the reference output. The input and reference output samples are created by using a mathematical software (Interactive Data Language). Statistic study and spectral analysis of these samples can then be undertaken to check that they are correct. Results related to IDL modelizations of each system module are given in the next section. Input samples are generated from quantization of a gaussian noise signal. This has a double advantage. First, it is a good test vector for functional simulations because of its random nature. Second, it corresponds to the mere nature of the radio astronomical signal to be processed by the filtering system during the observations. Thus, realistic estimates of the power consumption can be obtained from the toggle rate when performing functional simulations.

4.3. Implementation, Optimization and Functional Simulation

The input signal considered here is a white gaussian noise signal with four Continuous Waves (CW_i) whose respective normalized frequencies are ν_1=0.397, ν_2=0.405, ν_3=0.41, ν_4=0.45 (see Fig. 4.1).

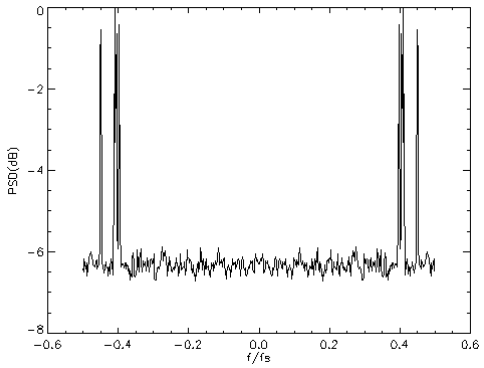

Fig. 4.1 : Input real signal corresponding to one 2GHz baseband. Horizontal scale is normalized to the 4GHz sampling frequency.

4.3.1. DDS and Mixer Device

Simulation

The input signal is translated in the frequency domain by -0.4 to place the sub-band of interest (around the CWs) at frequency zero. The complex output of the DDS is shown in Fig. 4.2 (left panel). The distributions of real and imaginary streams at the DDS and mixer device ouput (Fig. 4.2, right panel) are characteristic of cosine and sine distributions weighted by the 8 quantization levels of the 3-bit input samples.

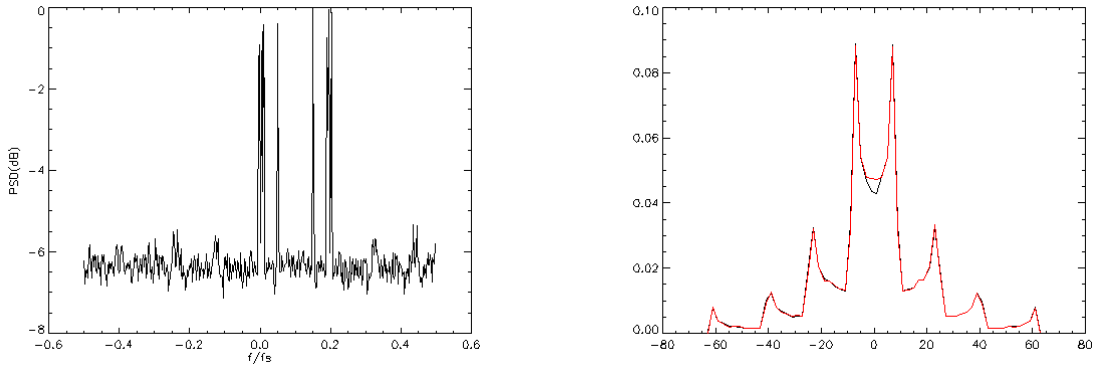

Fig. 4.2 : Translated input signal at the mixer output (left)
and complex output distribution (in percent) after the DDS & mixer stage.

Electronic Specifications

- 12 bits DDS accumulator
- Mixers implemented in LUTs (one LUT = 512*5 RAM bits)
- Use of sine and cosine symmetries, sign processing in external LEs
- Phase sub-increment calculated by DDS accumulator (see ref. [5])

4.3.2. First Filter Stage

Simulation

The complex output of the first filter stage is shown in Fig. 4.3 ; CW_4 falls in the first stage stopband and is thus filtered. The complex output band covers the normalized frequency range [-1/64, 1/64] of the input band [-1/2, 1/2] equivalent to [-2GHz, 2GHz]. This corresponds to a 125MHz sub-band.

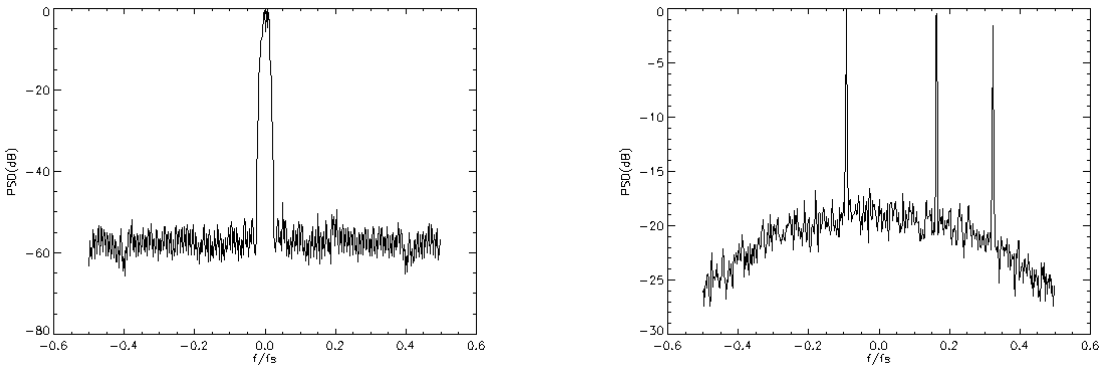

Fig. 4.3 : First filter stage output before decimation and requantization (left panel)
and after decimation and requantization (right panel).

Electronic Specifications

- Input : 2 streams of 32 parallel 6 bit samples at 125MHz
- 2 parallel identical FIR filters processing real and imaginary streams
- 128 taps, 8-bit encoded, symmetrical and frozen
- Addition of symmetrical samples
- Multiplications implemented in LEs
- Full scale adder tree
- Output : 2 streams of 8 bits at 125MHz

Optimization

Symmetry of the filter tap sequence is used to halve the number of implemented multiplications. We also take advantage of the fact that the taps are frozen for this first stage ; they do not change with the filtering mode. The multiplications are described as multiplications of the signal by constants (the tap weights). Optimization is obtained considering the specific value of each tap and implementing the multiplications by additions and 2^n multiplications (equivalent to bit shifting).

4.3.3. Second Filter Stage

Simulation

Fig. 4.4 shows the complex output of the second filter stage for filtering mode 1, where a 62.5MHz sub-band is synthesized. CW_3 falls in the second stage stopband and is thus filtered. The output band roughly represents the normalized frequency range [-1/4, 1/4] of the input band and corresponds to a 62.5MHz sub-band initially centered at 0.4 x fs = 1.6GHz.

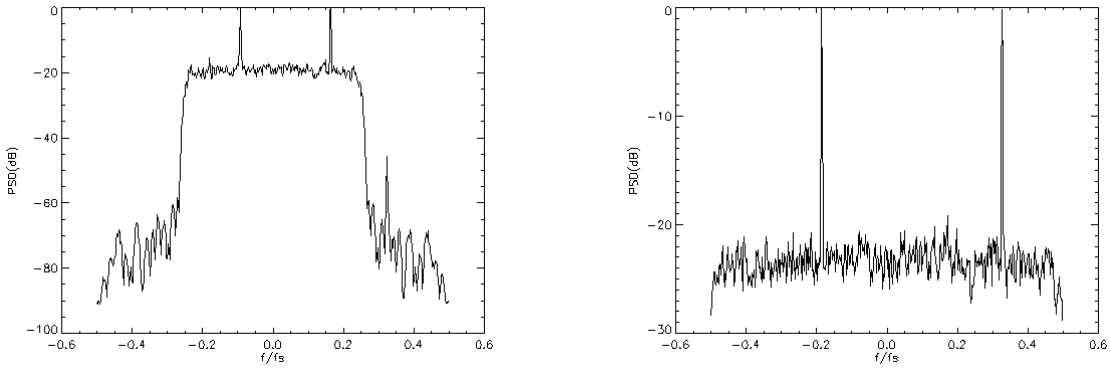

Fig. 4.4 : Second filter stage output before decimation (left panel) and after decimation (right panel).

Electronic Specifications

- Input : 2 streams of 8 bit samples at 125MHz
- 2 parallel identical FIR filters processing real and imaginary streams
- 64 taps (respectively 128, 256, 512 taps for sub-band narrowing down to 31.25MHz, 15.63MHz and 7.81MHz), 9-bit encoded, symmetrical and stored in RAMs
- Addition of symmetrical samples
- Tap recirculation
- Multiplications implemented in DSPBs
- Full scale adder tree
- Output : 2 streams of 9 bits at 62.5MHz (/2, /4, /8 for sub-band narrowing)

Optimization

Symmetry of the filter tap sequence is used to halve the number of multiplications. The number of multiplications is also diminished by implementation of multipliers working at an upper rate (125MHz) than required by the output sample rate (62.5MHz for filtering mode 1). This technique is referred as tap recirculation. Thus, a given multiplier is used for 2 different taps and the corresponding products are calculated on 2 successive clock cycles (for filtering mode 1). The number of physical multipliers is 64/2/2=16. For sub-band narrowing the tap recirculation technique has been preferred to the possibility of cascading resources implemented for several filtering systems. In order to narrow a sub-band more taps are required (respectively 128, 256 and 512 taps for the filtering modes 2, 3 and 4). However, the output data rate decreases in the same proportion ; thus, the same multiplier operating at the original clock rate can compute respectively 4, 8 or 16 consecutive taps. With tap recirculation the number of physical multipliers remains the same at the expense of slightly more complex architectures for the shift register and the adder tree.

4.3.4. Complex to Real Converter and Requantization Stage

Simulation

The real and imaginary streams are multiplexed as explained in Section 3.1. As expected, the spectrum of the real output signal is symmetrical (Fig. 4.5). The real output signal conveys the same information as that related to the complex signal at the second stage filter output.

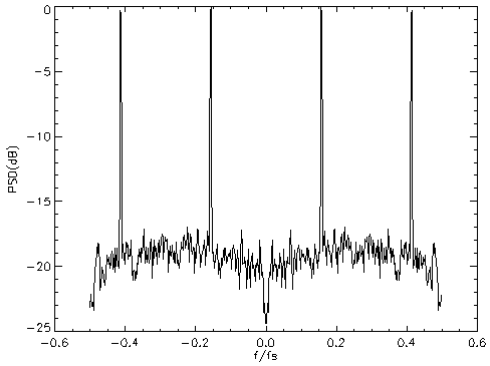

Fig. 4.5 : Signal output after complex to real conversion and 2-bit requantization.

Electronic Specifications

- Input : 2 streams of 9 bits at 62.5MHz (/2, /4, /8 for sub-band narrowing)
- Requantization implemented in one LUT (512*2 RAM bits)
- Output : 1 stream of 2 bits at 125MHz (/2, /4, /8 for sub-band narrowing)

4.4. Resources and Power Consumption

The detailed resource usage given in Table 4.1 corresponds to the design for filtering mode 1 using Stratix FPGAs. Main specifications of the Stratix family and the number of 2-stage filters which can be implemented in a single chip are specified in Table 4.2. The functional simulation of one full 2-stage filter at 125MHz gives a power consumption of nearly 2W.

	LE[1]	M512[1]	M4K[1]	DSPB[1]
DDS & mixer	1926	0	64	0
1st filter stage	6866	0	0	0
2nd filter stage	3542	8	0	4
Converter & requantization stage	49	0	1	0
Overall	12383	8	65	4

Table 4.1 Resource usage summary for one 2-stage filter

Device	LE	M512	M4K	DSPB	Number of 2-stage filters
EP1S10	10,570	94	60	6	~1
EP1S20	18,460	194	82	10	1
EP1S25	25,660	224	138	10	2
EP1S30	32,470	295	171	12	2, ~3
EP1S40	41,250	384	183	14	2, ~3
EP1S60	57,120	574	292	18	4
EP1S80	79,040	767	364	22	5

Table 4.2 Overview of the available resources for the Stratix family

4.5. FPGA Solution

The high performance digital filter bank proposed in this work requires implementation of our design in the most recent generation FPGA family. With Altera products this is achieved with the Stratix family. This could also be achieved with Xilinx FPGAs. However, we have adopted Altera products for two main reasons :

(i) possible migration of a FPGA-proven design to a 'HardCopy' device without custom ASIC design (see section below)

(ii) excellent contact with Altera-Europe team.

[1] cf. Section 4.1 : Technology Specifications

In the previous section, we show that up to five 2-stage filters could be integrated in a single Stratix chip. The optimum solution in terms of power consumption, board complexity and cost is discussed in Sections 5 and 6. With the FPGA solution each filtering mode corresponds to a different configuration of the FPGA. We stress that only part of the second filter stage has to be changed to achieve the different filtering modes. Thus, the required resources and the power consumption do not change significantly.

It is interesting to note that a study on effects of radiation on Altera's Stratix products has been undertaken for us in the frame of this work [7]. Component and system soft error rates have been predicted.

4.6. HardCopy Solution

For production above around 1000 units, Altera proposes a migration from a FPGA-proven design to HardCopy devices. HardCopy devices are mask-programmed devices created from a direct mapping of the FPGA architecture. Only top interconnection layers are specific to the design. This procedure leads to 70% die size reduction ; all logic cells required for FPGA reconfiguration and some M1M RAM blocks are removed. This greatly decreases the unit costs, improves the performances, and lowers the power consumption (40% better) compared to the original FPGA. Altera builds HardCopy devices based on customer provided FPGA design files and constraints files (see Fig. 4.6). Thus the FPGA-proven netlist is preserved and consistency with the original design is ensured. This migration process minimizes risks and accelerates the production.

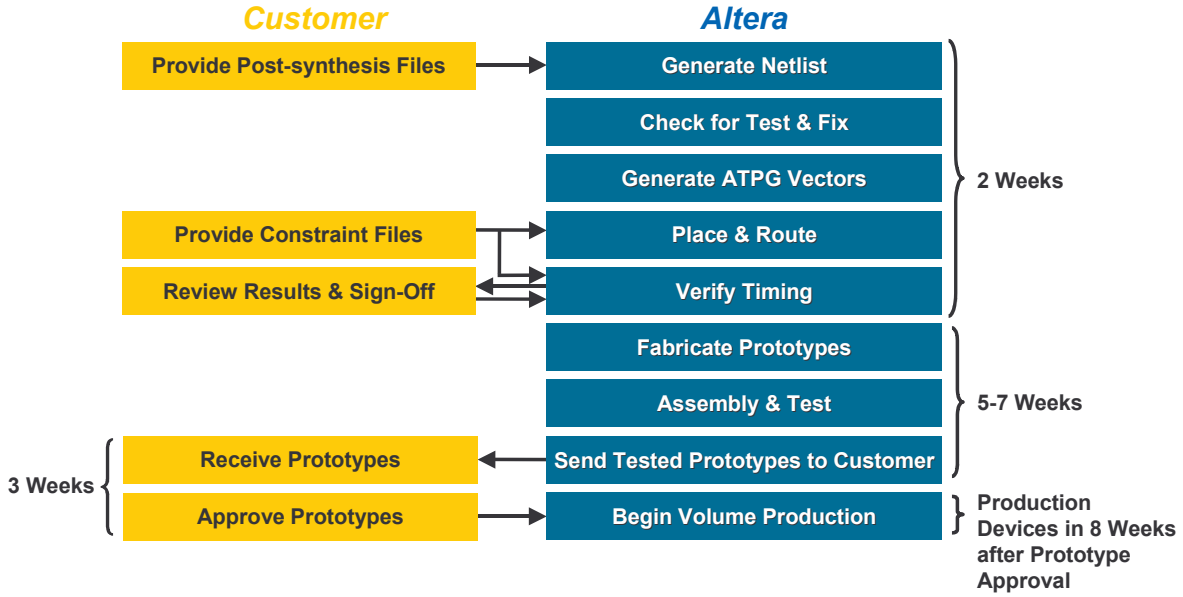

Fig. 4.6 HardCopy migration flow

With the HardCopy solution we have no longer access to the FPGA reconfiguration flexibility. Thus, a "universal" second stage filter (providing all filtering modes) has to be designed. A preliminary "universal" design has been investigated ; it would require more resources.

5. Interfacing the Change to the Baseline Correlator

5.1. Physical Dimensions

From the current resource estimations we conclude that we can implement two digital filters in one chip (see Table 4.2). Hence, the filter board requires 16 chips to implement 32 sub-bands for each 2GHz baseband. To this end, the EP1S25 or EP1S30 chips from the Altera Stratix family could be used. The package we plan to use is a Fineline Ball Grid Array (FBGA) with 780 pins. The size of each chip is 29x29 mm^2. A sketch of the proposed board is shown in Fig. 5.1.

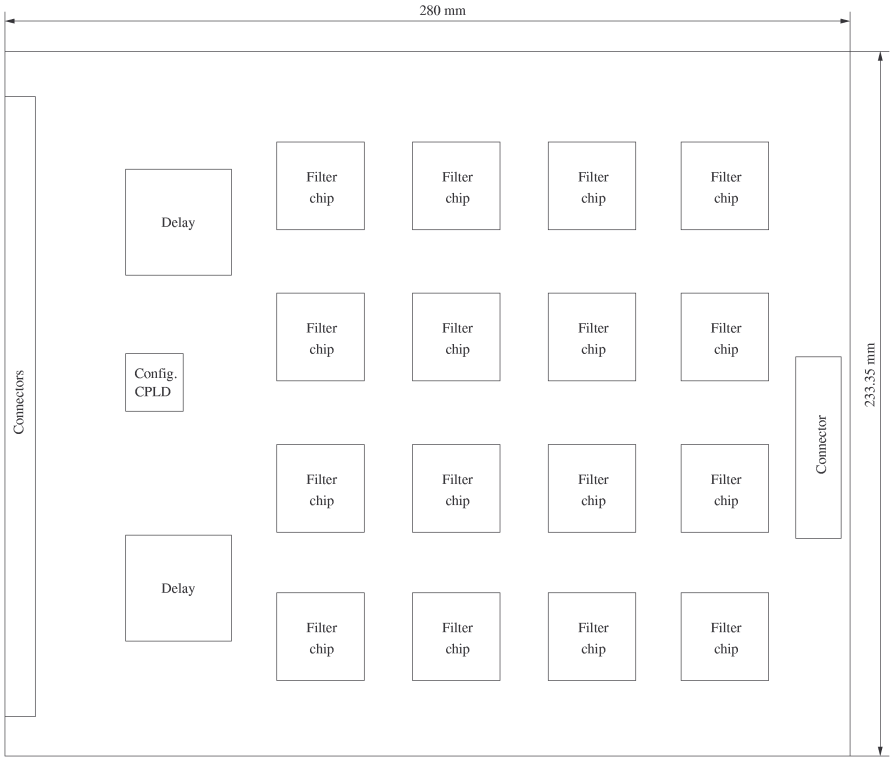

Fig. 5.1 : Sketch of the filter board

Other configurations with 3, 4 or 5 filters per chip (respectively EP1S40, EP1S60, EP1S80 would be required) and thus less chips per board would be possible using the Stratix family. However, this would result in higher power dissipation in localized areas of the board implying more difficult heat extraction. The 2 filters per chip solution is also preferred because of cost considerations (see Section 6).

5.2. Power Dissipation

The supply voltages required on the new filter card are 1.5V (core voltage of the FPGAs) and 3.3V. In addition, 1.8V is required for the Xilinx delay chips which are currently used in the baseline correlator filter board ; the possibility to migrate these chips to a 1.5V Xilinx family would ease power distribution.

The power dissipation per filter chip is based from the results obtained after implementation of a complete filter (DDS & mixer, first and second stages and complex to real conversion & requantization stage) as described in Section 4. With 2 filters implemented in one device the total power dissipated at 125MHz is 3.9W. Note that the power dissipation per filter does not change according to the adopted device. Because the dynamic requantization stage (see digital total power in Fig. 3.1) is not accounted for in our power estimate, we take 2W per filter as typical.

Component	Power per Component (W)		Number of Components	Total Power (W)	
	FPGA	HardCopy		FPGA	HardCopy
Filter Chips	4	2.4	16	64	38.4
Delay Chips	2		2	4	
Glue Logic	2		1	2	
Total				70	44.4

Table 5.1. Power dissipation for one filter board with two filters implemented in one chip

When 32 boards per rack are installed the total power per rack required for the boards is about 2.2kW. Most of the power dissipation comes from the filter chips and there are two possibilities to improve the dissipation :

(i) still improve the filter design

(ii) adopt the HardCopy device option to save up to about 40% power.

With filters implemented in HardCopy devices we expect a total power of 44W per board and 1.4kW for one rack.

5.3. Configuring the Altera FPGAs

Programming and interfacing the filter board FPGAs is performed by a CPLD chip (config. CPLD of Fig. 5.1.). This chip (referred as CPLD2 in the Baseline Correlator filter board documents) is programmed with the standard design used for the original filter card. Altera FPGAs are programmed at startup time using a RAM based mechanism very similar to that used by Xilinx FPGAs, and the same interface can be used to program both FPGA families on the same board. The main difference is in the polarity of the control signal nPROG generated by the chip. Therefore both true and complemented signal must be provided with a very minor modification in the design. All other differences are easily handled by choosing the appropriate control signals for the FPGAs.

The CPLD2 chip manages the programming interface for runtime control of the FPGAs. The programming model uses two registers for each FPGA and indirect addressing allows for almost unlimited address space in each FPGA (e.g. for downloading a large number of tap coefficients). The programming model used for the Baseline Correlator filter board is thus adequate for the newly proposed digital filter board and no hardware modifications are required. Software changes are needed but the indirect addressing scheme used for the Baseline Correlator filter will be followed as much as possible for enhanced compatibility and software reuse.

6. Cost Estimate for the Change

Altera has provided us with detailed costing of their Stratix devices for both the FPGA and HardCopy solutions. This costing is based on commercially available products distributed through Altera. We have not included here any cost projection based on newer technology (90nm) although future lower costs are foreseen.

For the FPGA solution the cost depends on the speed grade (5, 6 and 7 ; 5 corresponding to the fastest device) and on the delivery date. For speed grade 5 and considering that only 2 filters are implemented in one 1S25 chip we need for the first correlator quadrant 2048 chips and the total cost is around 0.34M$ for delivery in 2005. The total for all other 3 quadrants corresponding to 6144 chips delivered beyond October 2006 is around 0.9M$. Implementing 4 filters in a bigger FPGA (1S60) significantly increases the total cost because the price per chip does not rise in proportion with offered resources.

Migration of the filter design to the HardCopy solution is relatively quick as described in Section 4.6. The HardCopy solution requires a minimum order of 1000 units which applies well to the ALMA case. With 2 filters in one HC1S25 the total price for all 4 correlator quadrants requiring 8192 chips is around 0.81M$ (including 160000$ NRE). These chips are available in 2004 and beyond ; the cost does not change for order requests placed in 2004, 2005 or 2006. The component cost per filter card is then 1592$ compared to 2484$ for the FPGA option above.

In addition to the cost of components we estimate that a total of about 900$ per card is required for glue logic, connectors and board fabrication and assembly. The latter item (fabrication and assembly) is 800$, an estimation based on the 2GC 'interconnect demonstrator'. We believe that this cost will be lower at the time of discussions on final quantity production.

We have also worked on the implementation of filtering modes 2, 3 and 4 synthesizing smaller sub-bands. These filtering modes can be implemented in the HardCopy solution (see Section 4.6) but they require more resources and 1S30 products are needed ; the total cost of components for all 4 quadrants goes up to 1.18M$ (instead of 0.81M$).

7. Conclusions

We have shown that replacing the present ALMA digital filter card with the 2-stage filter system is feasible within the constraints of the Baseline Correlator design. The 2-stage filter design can be implemented in commercially available FPGAs or HardCopy devices. In both cases, the power consumption and the physical dimensions of the 2-stage filter boards are consistent with the Baseline Correlator requirements. We favor the HardCopy solution because of lower power consumption and better cost.

The characteristics of the 2-stage filter design have been described rather extensively. The main advantages of the 2-stage design are overlapping sub-bands to minimize aliasing effects and independent sub-band tunability offering high astronomical observing flexibility and potential zooming capability. In addition, we stress that this design is similar to the basic digital filter bank proposed for the future 2GC.

Acknowledgements : *This work was accomplished with support from OASU, University of Bordeaux1 (France), Arcetri Observatory (Italy), ASTRON (The Netherlands) and ESO for the ALMA project. We thank P. Caïs and P. Nouel (University of Bordeaux1) for their support during this work, and J. Raoult and B. Dhers for their help during the design flow of the 2-stage FIR filter. We thank Altera/Arrow-France for their help in quoting the FPGA components and in informing us on the newest possibilities of Altera HardCopy devices. We thank R. Escoffier (NRAO) for sending several details about the baseline ALMA filter card.*

8. References

[1] A. Bos, "A System Design Study of the ALMA Second Generation Correlator", March 2003. http://almaedm.tuc.nrao.edu/forums/alma/dispatch.cgi/iptdocscorr/saveWS/correlator/docProfile/100025/3328455

[2] A. Baudry editor, "Workshop on ALMA Second Generation and Large Correlators", March 26 & 27, 2002. http://www.observ.u-bordeaux1.fr/public/alma/correlator

[3] A.W. Gunst, A. Bos, A. Baudry, B. Quertier, G. Comoretto, "ALMA 2GC System Requirements", March 2003. available soon on :
http://almaedm.tuc.nrao.edu/

[4] R. Escoffier, J. Webber, "Enhancing the Performances of the Baseline ALMA Correlator", ALMA Memo N°441, November 2002.
http://almaedm.tuc.nrao.edu/forums/alma/dispatch.cgi/almamemos/saveWS/memos/docProfile/100045/4557083

[5] G. Comoretto, "Design of a FIR filter using a FPGA", Arcetri Technical Report, November 2002. http://www.arcetri.astro.it/science/Radio/alma/Report_5a_2002.pdf

[6] G. Comoretto, "Phase effects in an hybrid correlator", Arcetri Technical Report, August 2003. http://www.arcetri.astro.it/~comore/phase.pdf

[7] Altera confidential memo, "Radiation Induced Soft Errors in Altera Products", August 2003.