National Radio Astronomy Observatory
520 Edgemont Road
Charlottesville, VA 22903-2475

John Webber
IR

ALMA Construction Project Book

Version 4.0, 2001-Feb-8

ALMA Construction Project Book

Version 4.0 - 2001-Feb-08

Table of Contents

ALMA Project Book: INTRODUCTION

Robert Brown and Richard Kurz
Last changed 2000-Sept-27

Revision History:
2000-Sep-27: First ALMA version.

The Atacama Large Millimeter/Submillimeter Array (ALMA) is a revolutionary instrument in its scientific concept, its engineering design, and its organization as a global scientific endeavor. ALMA will provide scientists with precise images of galaxies in formation seen as they were twelve billion years ago; it will reveal the chemical composition of heretofore unknown stars and planets still in their formative process; and it will provide an accurate census of the size and motion of the icy fragments left over from the formation of our own solar system that are now orbiting beyond the planet Neptune. These science objectives, and many hundreds more, are made possible owing to the design concept of ALMA that combines the imaging clarity of detail provided by a 64-antenna interferometric array together with the brightness sensitivity of a single dish antenna.

The challenges of engineering the unique ALMA telescope begin with the need for the telescope to operate in the thin, dry air found only at elevations high in the Earth's atmosphere where the *light* at millimeter and submillimeter wavelengths from cosmic sources penetrates to the ground. ALMA will be sited in the Altiplano of northern Chile at an elevation of 5000 meters (16,500 feet) above sea level. The ALMA site is the highest, permanent, astronomical observing site in the world. On this remote site the 64 12-meter diameter ALMA antennas will each operate superconducting receivers that are cryogenically cooled to less than 4 degrees above absolute zero. The signals from these receivers are digitized and transmitted to a central processing facility where they are combined and processed at a rate of 1.6×10^{16} operations per second. As an engineering project, ALMA is a concert of 64 precisely-tuned mechanical structures each weighing more than 50 tons, superconducting electronics cryogenically cooled, and optical transmission of terabit data rates--all operating together, continuously, on a site more than 3 miles high in the Andes mountains.

ALMA is a joint endeavor of nations and science institutes worldwide. The cost and burden of building and operating ALMA will be shared among the participants. This cooperation brings to the Project a

broad base of experienced people and resources. Properly used, this breadth of experience has the potential to reduce risk in many areas. Here the challenge is to manage the combined resources in a way that empowers the participants and effectively coordinates their efforts.

The ALMA Project Book is the key used to unlock much of the solution to the scientific, engineering and organizational challenges facing the Project. For ALMA, the Project Book is the description of the science requirements, the technical specifications, the schedule on which tasks are to be accomplished, and the task responsibilities. Where one task interacts with another either in design or integration, the interface requirements are specified.

The Project Book is the controlled configuration for the Project. Specifications in the Project Book--technical specifications, interface specifications or schedule—are controlled by ALMA System Engineering. Changes cannot be made to the configuration without the process specified by System Engineering, through the Control Board, being followed and approval gained.

The Project Book is the fundamental reference for what is, and is not, in the Project. As decisions are made and implemented by System Engineering, or by the Control Board in the case of changes, those decisions will be incorporated into the Project Book. The Project Book is kept electronically and is always available on-line for reference by the Project and interested others. A revision history is included to aid change tracking. Maintenance of the Project Book is the responsibility of ALMA System Engineering.

The ALMA Project Book serves to cement the scientific, technical and organizational aspects of the Project together. For the geographically distributed ALMA Project, the Project Book serves as a crucial facility to aid communication within the Project and to anchor decisions. Although the Project Book is a living document and will evolve, it is at all times the current, and complete, Project configuration to which all participants in the Project are working to achieve.

ALMA Project Book, Chapter 2

ALMA SCIENCE REQUIREMENTS

Robert Brown
Last revised 2000-April-25

Revision History
1998-08-08: First version
1998-11-23: Major revision: Table 2.1 added
1999-04-26: Minor Corrections
2000-04-25: *Table 2.1 revised*

Summary

U.S.: In the U.S. the scientific capabilities required in the ALMA were refined in community science workshops sponsored by the NRAO throughout the decade of the 1980s and confirmed by the September 1995 ALMA Science Workshop held in Tucson, AZ. Five reports were written following the Tucson Workshop that summarize the science goals in the following categories:

1. Cosmology and Extragalactic
2. Star Formation and Stellar Evolution
3. Galactic Molecular Clouds and Astrochemistry
4. Solar System
5. Sun and Stellar

These reports are available on the ALMA www pages. While these different scientific areas emphasize different capabilities, they all require precision imaging over the millimeter and sub-millimeter wavelength bands and over resolutions from arcseconds to less than a tenth of an arcsecond.

Europe: In Europe the scientific requirements for the LSA were similarly discussed and summarized in documents:

1. Science at high Z or the youth of the Universe
2. Planetary Formation or the youth of the Solar System

The arguments were summarized at a 1995 meeting in Garching, at which the motivation was written. The LSA science requirements were issued, and in the 1998 *April LSA/MMA Feasibility Study* options for combining the arrays were explored. At a meeting in 1999 February in Tucson, the science requirements were joined.

The two projects merged into the ALMA project in 1999 June, resulting in a merger of the similar science needs. The capabilities of the merged array were discussed at the 1999 October Workshop "Science with the Atacama Large Millimeter Array." The proceedings of this conference will be published. In the meantime, unedited versions of the contributions to the Proceedings of that conference may be obtained at the ALMA Conference raw material dropbox. The science requirements and the technical specifications that derive from this material are summarized in Table 2.1.

Table 2.1 ALMA Science Flowdown to Technical Specifications

Science Requirement and Examples	Technical Requirements Needed to Achieve
1. High Fidelity Imaging • Imaging spatial structure within galactic disks; • Imaging chemical structure within molecular clouds; • Imaging protostars in star formation regions	• Reconfigurable Array • Robust Instantaneous uv-coverage, $N_{ant} > 60$ • Precision Pointing, 6% of the HPBW • Antenna Surface Accuracy RMS = 20 microns • Primary Beam Deviations < 7% • Total Power and Interferometric Capability • Precise (1%) Amplitude Calibration • Precise Instrumental Phase Calibration (<10 degrees rms) • Precise atmospheric phase calibration (<15 degrees rms) with compensation using both fast switching and water vapor radiometry
2. Precise Imaging at 0."1 Resolution • Ability to discriminate galaxies in deep images; • Imaging protoplanets orbiting protostars; • Imaging nuclear kinematics	• Interferometric baselines longer than 3 km • Precise Instrumental Phase Calibration (<10 degrees rms) • Precise atmospheric phase calibration (<15 degrees rms) with compensation using both fast switching and water vapor radiometry
3. Routine Sub-milliJansky Continuum Sensitivity • To enable imaging of the dust continuum emission from cosmologically-distant galaxies; • To enable imaging of protostars and protoplanets throughout the Milky Way • To enable astrometric observations of solar system minor planets and Kuiper-belt objects	• Array site with median atmospheric transparency < 0.05 at 225 GHz • Quantum-limited SIS receivers • Antennas with warm spillover <5K, and aperture blockage <3% • Antennas of aperture efficiency > 75% • Wide correlated IF bandwidth, 16 GHz • Dual polarization receivers • Array collecting area, $ND^2 > 7000$ m^2
4. Routine Milli-Kelvin Spectral Sensitivity • Spectroscopic probes of protostellar kinematics • Spectroscopic chemical analysis of protostars, protoplanetary systems and galactic nuclei • Spectroscopic studies of galactic disks and spiral structure kinematics	• Array site with median atmospheric transparency < 0.05 at 225 GHz • Quantum-limited SIS receivers • Antennas with warm spillover < 5 K, aperture blockage <3% • Antennas with aperture efficiency > 0.75 • Wide correlated IF bandwidth, 16 GHz • Dual polarization receivers • Array collecting area, $ND^2 > 7000$ m^2 • Array collecting length, ND > 700 m

Science Requirement and Examples	Technical Requirements Needed to Achieve
5. Wideband Frequency Coverage • Spectroscopic imaging of redshifted lines from cosmologically-distant galaxies • To enable comparative astrochemical studies of protostars, protoplanets and molecular clouds • To enable quantitative astrophysics of gas temperature, density and excitation	• Receiver bandwidths matched to the width of the atmospheric windows • Tunable local oscillator matched to the bandwidth of the receivers • Cryogenic capacity > 1 W at 4 K
6. Wide Field Imaging, Mosaicking • Imaging galactic disks • Imaging the astrophysical context of star formation regions • Imaging surveys of large angular regions • Searches for dusty and luminous protogalaxies • Searches for minor planets in the solar system • Solar astrophysics	• Compact array configuration, filling factor > 0.5 • Instantaneous uv-coverage that fills more than half the uv-cells, $N_{ant} > 60$ • Precision pointing, 6% of HPBW • Antenna surface accuracy 20 microns • Total power and interferometric capability • Precise amplitude calibration, 1% • Precise Instrumental Phase Calibration (<10 degrees rms) • Correlator dump time 10 msec • Capability to handle data rates > 100 Mbyte/sec
7. Submillimeter Receiving System • Measurement of the spectral energy distribution of high redshift galaxies • Chemical spectroscopy using CI and atomic hydrides • Determination of the CII and NII abundance in galaxies as a function of cosmological epoch	• Array site with median atmospheric transparency < 0.05 at 225 GHz • Quantum-limited SIS receivers • Antennas with warm spillover < 5 K, aperture blockage <3% • Antennas with aperture efficiency > 0.75 • Precise Instrumental Phase Calibration (<10 degrees rms) • Precise atmospheric phase calibration (<15 degrees rms) with compensation using both fast switching and water vapor radiometry
8. Full Polarization Capability • Measurement of the magnetic field direction from polarized emission of dust • Measurement of the magnetic field strength from molecular Zeeman-effect observations • Measurement of the magnetic field structure in solar active regions	• Measure all Stokes parameters simultaneously • Cross correlate to determine Stokes V • Calibration of linear gains to <1%
9. System Flexibility • To enable VLBI observations • To enable pulsar observations • For differential astrometry • For solar astronomy	• Ability to phase the array for VLBI • Sum port on the correlator for external processing • Sub-arraying, 4 subarrays simultaneously • Optics designed for solar observations

II. General Requirements

1. Frequency Coverage

ALMA needs eventually to cover all the available atmospheric windows between about 30 and 900 GHz. An initial implementation of the four frequency bands of highest scientific priority was specified by the ALMA Science Advisory Committee. These requirements are summarized in the Frequency Bands whitepaper. This requires an outstanding site as discussed in the "Recommended Site for the Millimeter Array" document.

2. Spectral Line and Continuum

The ALMA must operate as both a sensitive spectral line and continuum array. This implies using the widest continuum bandwidth practical from the point of view of the IF and the correlator. This appears now to be 8 GHz per IF; however, a 4 GHz per IF bandwidth may be accommodated should the wider bandwidth compromise receiver performance. A flexible correlator as described in the MMA Correlator Whitepaper will process a total of 16 GHz from each antenna. See also Chapter 10 of this Project Book and the ALMA Correlator Home Page.

3. Sensitivity

The array must maximize both point source and surface brightness sensitivity. For antennas with the same overall properties, this requires maximizing the different quantities, nD^2 (for point source sensitivity), nD (for surface brightness sensitivity in a sparsely filled array) and $n^{½}$ (for a tightly packed array or total power mode). See MMA Memo 177, MMA Memo 243, MMA Memo 273.

4. High Resolution

Given the expected brightness and size of sources considered in the science documents, this implies array configurations capable of providing precision, high fidelity imaging at 0."1 angular resolution. Such configurations will have an extent of 4 km. In addition, a very extended configuration of 12 km diameter is needed for imaging at ~10 milli-arcseconds. This requirement demands the ALMA be adequately phase stable both internally and in the presence of atmospheric phase fluctuations. This will be discussed further in Section 3.

5. Large Source Imaging

On the other end of the angular scale, a further requirement exists for imaging objects both close to and bigger than the primary beam. This requirement has several implications. First, 3)+ 4) + 5) require that the array be reconfigurable into configurations optimized for the resolution and sensitivity required by each experiment. Second, the array must be able to make large mosaicked images (multiple pointings) to image regions on the order or larger than the primary beam. Third, the array must have a sensitive, stable total power system so that spatial frequencies smaller than are available in interferometer mode can be measured. Since the primary beams at the highest frequencies for antennas of 12 m diameter are < 10 arcseconds, modes 2 and 3 should be very common, perhaps being the vast majority of all observations with the array. A goal of the array is to produce images of similar size and resolution to those produced from optical telescopes, to facilitate direct comparison. Such a goal entails mosaicking.

6. High Fidelity Imaging

Especially in the modes discussed in 5), a significant fraction of the experiments and some the most important require high fidelity imaging. That is, the signal-to-noise must be high enough and the uv-coverage complete enough that even for complex sources errors in pointing and calibration will not degrade the scientific usefulness of the experiment.

Such problems require 1) excellent pointing, 2) high quality amplitude calibration and 3) accurate phase calibration; these topics will be discussed quantitatively in Section III.

7. Polarization

Observations of both linear and circular polarization of lines and continuum emission are a significant part of the ALMA science program. At centimeter and longer wavelengths interferometers produce linear polarization by correlating the opposite circular polarizations from different antennas, that is R with L and L with R. However, it appears technically difficult to do this at millimeter wavelengths across the broad bands needed for with the ALMA. Thus it seems best to observe in the more natural linear polarization with the ALMA. This means we crosscorrelate to calculate the V stokes parameter; we get I, Q and U from linear combinations of the two linear correlations. This requires both linears be present all the time and that either their relative gains remain very stable and/or we have the necessary internal calibration signals to measure their changes. (see MMA Memo 208 , Cotton, 1998 and ASAC Report on Polarization).

8. Solar Observations

Requirements for observing the sun are discussed in the Sun and Stars science document and by Bastian et al, 1998.

9. VLBI

The highest resolution with the ALMA will be obtained from VLBI observations using the ALMA as a single element. The requirements for this are discussed by Claussen and Ulvestad, 1998.

10. Pulsar/High Speed

Pulsar observations will require a gating mode with the correlator as well as a sum port which can be attached to specialized external recording equipment. This latter capability should also be available for other high speed phenomena such as solar or stellar flares.

III. Implications

The requirements summarized above imply the need for the array capabilities summarized here.

1. Phase Stability

As the observing frequency increases into the submillimeter the electrical path length through the atmosphere and through the electronics must be increasingly stable in order to enable the ALMA to produce high fidelity images.

- Internal Phase Stability. The electronics systems must be stable enough that they do not degrade the imaging relative to those path length fluctuations caused by atmospheric effects. For example, at 900 GHz the instrumental phase must be less than 10 degrees. A system to monitor and compensate for electrical path length changes in the instrument is necessary.

- Atmospheric Phase Stability. Atmospheric path length changes must be measured and corrected to preserve the capability for high fidelity imaging. Techniques to be developed to accomplish this include fast switching of the array antennas and radiometric techniques for measuring and correcting atmospheric phase distortion.

- *Fast Switching.* In this mode the antennas are rapidly cycled between a nearby calibrator source and the program source before the atmosphere can change significantly. The method has proven to be effective (MMA Memo 139, MMA Memo 173); it requires the antennas to have the capability to move between source and calibrator that are separated by less than 2 degrees on the sky on a cycle time of less than 10 seconds.

- *Radiometric Phase Correction.* Since water vapor in the atmosphere is responsible for both temporal changes in the sky opacity and changes in the electrical path length (the phase), measurements of the changing sky brightness can be used to infer changes in the atmospheric phase distortion. The techniques require stable radiometry and are best employed using either the 22 GHz atmospheric water line or the 183 GHz water line. The expected efficacy of the techniques, and the precision required by the ALMA, are discussed in MMA Memo 210. This has been updated in an ASAC White Paper.

2. Amplitude Stability

The capability to measure and maintain amplitude stability of the ALMA at the level of one percent is needed to combine imaging information from one array configuration to another reliably and permit accurate comparison of line strengths to determine such physical parameters as the excitation temperature of interstellar clouds or material in galactic nuclei. This will require use of an external calibration system such as discussed in MMA Memo 225. The specifications can be found in an ASAC White Paper and in ALMA Memo 289.

3. Integration Times

The fastest integration time needed by the ALMA will be driven as much by the need to perform total power continuum observations and fast on-the-fly mosaicking as it will be by the need to measure time variability in astronomical sources. This issue is evaluated quantitatively in MMA Memo 192.

4. Contingency Scheduling

This is an operational issue. The ALMA will need to be scheduled to allow the most demanding submillimeter observations, and mosaicking observations in the most compact configuration, to be done in conditions of favorable transparency and low prevailing wind. To accomplish this the array will need to be scheduled in near real time.

5. Data Flow

This is another operational issue. The astronomer will benefit by the ability to see his or her data in near real time. Most observations requiring longer than a few hours will be scheduled such that they are made over several source transits so little or no data is taken at extreme hour angles where the low elevation will compromise the system noise. This provides the opportunity for the astronomer to refine his or her observational techniques as the observations are in progress. The design requirement is for real-time imaging and for the capability for those images to be transmitted from the Chile site to the astronomer in the U.S. or elsewhere in a timely way. This requirement and its implications are explored in MMA Memo #164.

ALMA Project Book, Chapter 2.2

ALMA ENGINEERING SPECIFICATIONS

Darrel Emerson & Jaap Baars
Last revised 2000-Dec-12

Revision History
2000-12-01: First near-complete version

Introduction

In this chapter we give a summary of the engineering specifications of all aspects of the ALMA project. The material is drawn from other chapters of this Project Book, but is collected here for convenience. This chapter is intended to be completely consistent with all other chapters of the book. If discrepancies are found, please notify the editors DTE or JWMB as soon as possible.

Tables of specifications

Table 2.1 ALMA Science Flowdown to Technical Specifications

Science Requirement and Examples	Technical Requirements Needed to Achieve
1. High Fidelity Imaging • Imaging spatial structure within galactic disks; • Imaging chemical structure within molecular clouds; ○ Imaging protostars in star formation regions	• Reconfigurable Array • Robust Instantaneous uv-coverage, $N_{ant} > 60$ • Precision Pointing, 6% of the HPBW • Antenna Surface Accuracy RMS = 20 microns • Primary Beam Deviations < 7% • Total Power and Interferometric Capability • Precise (1%) Amplitude Calibration • Precise Instrumental Phase Calibration (<10 degrees rms) • Precise atmospheric phase calibration (<15 degrees rms) with compensation using both fast switching and water vapor radiometry
2. Precise Imaging at 0.1" Resolution • Ability to discriminate galaxies in deep images; • Imaging protoplanets orbiting protostars; • Imaging nuclear kinematics	• Interferometric baselines longer than 3 km • Precise Instrumental Phase Calibration (<10 degrees rms) • Precise atmospheric phase calibration (<15 degrees rms) with compensation using both fast switching and water vapor radiometry

3. Routine Sub-milliJansky Continuum Sensitivity • To enable imaging of the dust continuum emission from cosmologically-distant galaxies • To enable imaging of protostars and protoplanets throughout the Milky Way • To enable astrometric observations of solar system minor planets and Kuiper-belt objects	• Array site with median atmospheric transparency < 0.05 at 225 GHz • Quantum-limited SIS receivers • Antennas with warm spillover <5K, and aperture blockage <3% • Antennas of aperture efficiency > 75% • Wide correlated IF bandwidth, 16 GHz • Dual polarization receivers • Array collecting area, $ND^2 > 7000\ m^2$
4. Routine Milli-Kelvin Spectral Sensitivity • Spectroscopic probes of protostellar kinematics • Spectroscopic chemical analysis of protostars, protoplanetary systems and galactic nuclei • Spectroscopic studies of galactic disks and spiral structure kinematics	• Array site with median atmospheric transparency < 0.05 at 225 GHz • Quantum-limited SIS receivers • Antennas with warm spillover < 5 K, aperture blockage <3% • Antennas with aperture efficiency > 0.75 • Wide correlated IF bandwidth, 16 GHz • Dual polarization receivers • Array collecting area, $ND^2 > 7000\ m^2$ • Array collecting length, ND > 700 m
5. Wideband Frequency Coverage • Spectroscopic imaging of redshifted lines from cosmologically-distant galaxies • To enable comparative astrochemical studies of protostars, protoplanets and molecular clouds • To enable quantitative astrophysics of gas temperature, density and excitation	• Receiver bandwidths matched to the width of the atmospheric windows • Tunable local oscillator matched to the bandwidth of the receivers • Cryogenic capacity > 1 W at 4 K
6. Wide Field Imaging, Mosaicking • Imaging galactic disks • Imaging the astrophysical context of star formation regions • Imaging surveys of large angular regions • Searches for dusty and luminous protogalaxies • Searches for minor planets in the solar system • Solar astrophysics	• Compact array configuration, filling factor > 0.5 • Instantaneous uv-coverage that fills more than half the uv-cells, $N_{ant} > 60$ • Precision pointing, 6% of HPBW • Antenna surface accuracy 20 microns • Total power and interferometric capability • Precise amplitude calibration, 1% • Precise Instrumental Phase Calibration (<10 degrees rms) • Correlator dump time 10 msec • Capability to handle data rates > 100 Mbyte/sec

7. Submillimeter Receiving System • Measurement of the spectral energy distribution of high redshift galaxies • Chemical spectroscopy using CI and atomic hydrides • Determination of the CII and NII abundance in galaxies as a function of cosmological epoch	• Array site with median atmospheric transparency < 0.05 at 225 GHz • Quantum-limited SIS receivers • Antennas with warm spillover < 5 K, aperture blockage <3% • Antennas with aperture efficiency > 0.75 • Precise Instrumental Phase Calibration (<10 degrees rms) • Precise atmospheric phase calibration (<15 degrees rms) with compensation using both fast switching and water vapor radiometry
8. Full Polarization Capability • Measurement of the magnetic field direction from polarized emission of dust • Measurement of the magnetic field strength from molecular Zeeman-effect observations • Measurement of the magnetic field structure in solar active regions	• Measure all Stokes parameters simultaneously • Cross correlate to determine Stokes V • Calibration of linear gains to <1%
9. System Flexibility • To enable VLBI observations • To enable pulsar observations • For differential astrometry • For solar astronomy	• Ability to phase the array for VLBI • Sum port on the correlator for external processing • Sub-arraying, 4 subarrays simultaneously • Optics designed for solar observations

Calibration Requirements

The precision imaging to be attained by the ALMA will be achieved through accurate calibration. The types of calibration are summarized in Table 3.1.

Table 3.1 ALMA Calibration Requirements.

Pointing	0.6" absolute
Primary Beam	2-3%
Baseline Determination	0.1 mm
Flux Calibration	1% absolute flux accuracy goal
Phase Calibration	0.15 radian at 230 GHz
Bandpass Calibration	10000:1 to 100000:1
Polarization Calibration	10000:1
Single Antenna Calibration	Employed

Antenna Specifications

The ALMA radiotelescope is currently planned to consist of a goal of 64 antennas, each of 12 m diameter. In this chapter we outline the general requirements for the antennas and the detailed specifications can be in the contract for the prototype antenna (NRAO, 2000) and in the Interface Control Documents (ICD, 2000) which are part of the contract. The principal requirements for the antennas are shown in Table 4.1.

Table 4.1 ALMA antenna principal performance requirements.

Configuration	Elevation-over-azimuth mount, Cassegrain focus
Frequency range	30 GHz to 950 GHz
Precision performance conditions	Nightime: wind 9 m/s Daytime: wind 6 m/s and sun from any angle
Reflector surface accuracy	20 microns rms, goal; 25 microns rms, spec
Pointing accuracy, rms	0.6 arcsec (offset, 2 deg in position and 15 min time), 2.0 arcsec (absolute)
Fast switching (settle to 3 arcsec pointing)	Move 1.5 deg in position in 1.5 seconds
Phase stability	15 microns rms
Close packing	1.25 dish diameters (15.0 m) between azimuth axis
Solar observing	Allowed
Transportability	Transportable on a rubber-tired vehicle

The antennas will be designed and built by one or more commercial companies. Prototype antennas are being built for the US and European ALMA partners by Vertex Antenna Systems LLC (Santa Clara, CA) and European Industrial Engineering (EIE) (Mestre, Italy) respectively.

Receiver Specifications

The document *Specifications for the ALMA Front End Assembly* (latest version) contains the detailed specifications. Portions of this have been approved by the AEC. The main specifications are:

- Frequency coverage: from 31.3 to 950 GHz in 10 bands (see Table 5.1)
- simultaneous reception of two orthogonal polarizations
- receiver noise between 6 and 10 times $h\nu/k$ over 80% of the band, with a goal of achieving 3 to 8 times $h\nu/k$, depending on the band
- IF bandwidth 8 GHz total per polarization
- observations at one frequency at a time (no dual frequency obervations)
- inclusion of a water vapour radiometer using the 183 GHz line for phase correction.

For details, see the full *Specifications for the ALMA Front End Assembly* (latest version).

Table .1 – Frequency bands for ALMA

Band	from (GHz)	to (GHz)
1	31.3	45
2	67	90
3	89*	116
4	125	163
5	163	211
6	211	275
7	275	370

8	385	500
9	602	720
10	787	950

* change to 84 GHz has been proposed

LO specifications

The LO subsystem also forms part of the array master clock, in cooperation with a computer of the monitor-control subsystem. It does this by providing an interface to an external time scale (currently GPS) and by measuring the difference between external time and array time. Measures of time larger than 48 msec are obtained in the MC system by integration. Further details are given in the LO chapter.

Table 1: Specification Summary

Item	Specification	Goal (if different)
Frequency Range, 1st LO	1st LO: 27.3 to 938 GHz (see Table 2) 2nd LO: 8-10 and 12-14 GHz	
Output Power	1st LO: band dependent (see Table 3) 2nd LO: +10dBm ea. to 2 converters.	100 μW
Sideband Noise, 1st LO	10 K/μW	3 K/μW
Amplitude Stability, 1st LO	.03% <1s; 3% between adjustments	.01%; 1%
Phase Noise (>1 Hz)	63 fsec (18.9 μm)	31.4 fsec (9.4 μm)
Phase Drift (<1 Hz)	29.2 fsec (8.8 μm)	6.9 fsec (2.1 μm)
Tuning step size, maximum	On the sky: 250 MHz SIS mixer 1st LO: 500 MHz	
Subarrays with independent tunability	TBD (3 or more)	5
Simultaneous different sky frequencies	1 per subarray	
Time for frequency change, maximum	Within .03% (freq switching): 10 msec Otherwise: 1.5 sec	1 msec 1.0 sec
Repeatability	1. Phase-unambiguous synthesis 2. Stability specs apply across frequency changes.	

LO Output Power

The local oscillator must provide adequate mixer drive power for both HFET and SIS based receivers. A conventional balanced mixer used in a millimeter-wave HFET front-end requires approximately 5 mW of LO power. However, 20 mW may be required if a sideband-separating mixer follows the low noise HFET amplifier.

Chapter 7, Table 3: First LO Power Requirements

ALMA Receiver Band	LO Tuning Range [GHz]	Type of Receiver Front-End	Number of SIS Junctions	Minimum Required Mixer Power	Required Power at Input of -20 dB Coupler of SIS Mixer	Required Power at LO port of an Image -Reject & Balanced Mixer	LO Power Specification of 50% Over Worst-Case
1	27-33	HFET	---	5 mW	---	10 mW	15 mW
2	71-94	HFET	---	5 mW	---	10 mW	15 mW

3a	101-104	HFET	--	5 mW	---	10 mW	15 mW
3b	101-104	SIS	4	0.10 μW	10 μW	0.40 μW	15 μW
4	137-151	SIS	4	0.15 μW	15 μW	0.60 μW	23 μW
5	175-199	SIS	4	0.26 μW	26 μW	1.06 μW	39 μW
6	223-263	SIS	4	0.46 μW	46 μW	1.84 μW	69 μW
7	287-358	SIS	2	0.21 μW	21 μW	0.84 μW	32 μW
8	397-488	SIS	2	0.40 μW	40 μW	---	60 μW
9	614-708	SIS	2	0.42 μW	42 μW	---	63 μW
10	799-938	SIS	1	0.37μW	36 μW	0.73 μW	54 μW

In the worst-case scenario where only single-ended, two-port SIS mixers are used, a waveguide or quasi-optical LO coupler, having a coupling factor of -20 dB, will be required to combine the LO and RF signals appropriately. The LO power required at the input of the coupler is also given in Table 3. However, if a balanced mixer can be utilized, the LO power is supplied via a separate LO port on the mixer thus rendering the coupler unnecessary. Column #7 in Table 3 lists the power requirements for a balanced mixer configuration that is both image separating and balanced. The last column is a suggested *specification* per RF band based upon a 50 percent overhead for the worst-case conditions. The LO power *goal* will be 100 μW per band to ensure adequate power to overcome losses within the mixer block.

The Downconverter

In each ALMA antenna there will be two Downconverter modules, one for each polarization, and the two inputs to each module will carry upper and lower side-band signals. A block diagram of the Downconverter is shown in Figure 9.1.1 and the specifications are in Table 9.1.1 The input and output noise power spectral power distribution will be nominally flat over the passband as given in the specifications. The Downconverter will take the wideband 4 - 12 GHz input signals received from the front end subsystem and produce four output signal channels each with a passband of 2 - 4 GHz suitable for bandpass sampling at by the digitizers, which are clocked at 4 GS/s

Table 9.1.1 Specifications for Downconverter

DOWNCONVERTER MODULE

SPECIFICATIONS for ALMA CONSTRUCTION

Reference: Block Diagram, Document # ALMA06002KX0002

*** indicates interfaces**

Number of modules	142 (2 x 64 antennas plus 14 spares)
***Inputs from front end**	
Number of inputs per module	Two: USB, LSB (upper and lower sidebands)
Frequency range, nominal	4-12 GHz or 4-8 GHz
Power level within any 2 GHz bandwidth when the antenna temperature is 290K	-40 +/-3 dBm, less loss of coax and connectors between front end outputs and module inputs (3m of phase stabilized Andrew FSJ1P-50A ¼ inch diameter, attenuation = 2.4 dB @ 12GHz)
Variation of power spectral density vs. frequency (flatness)	<+/-1.5 dB across the nominal frequency range

Headroom when the antenna temperature is 290K (see definition)	>20 dB
***Inputs from Second LO (LO2)**	
Number of inputs per module	Four (A, B, C, D), independently tunable
Frequency range	8.0-14.0 GHz nominal; frequency LO2 > frequency input
Power level	+13 +/-1 dBm
Power level of spurious frequencies	<-70 dBc, except <-40 dBc for 2nd harmonic
***Outputs to digitizers**	
Number of outputs per module	Four (A, B, C, D)
Frequency range	2 - 4 GHz nominal
Power level	-TBD +/-TBD dBm plus loss of coax and connectors between output and input to digitizer module
Headroom when the antenna temperature is 290K	>20 dB
LO2 spurious and leakage at outputs	<(power level -40 dB) for all combinations of frequencies of LO2-A, -B, -C, -D

Throughput from front end inputs to outputs to digitizers	
Input S_{11} reflection magnitude 4 – 12 GHz	<-20 dB (VSWR < 1.22) to minimize spectral ripples
Input noise figure 4 – 12 GHz	< 10 dB (2 610K); SP_{DC} < -164 dBm/Hz << SP_{FE} = -133 dBm/Hz
Image rejection	>20 dB
Filter, 4-12 GHz nominal bandpass for total power detection	passband <4.0 GHz and >12.0 GHz at -1 dB, max ripple +/-0.5 dB; stopband 3.5 GHz and 12.5 GHz at < -20 dB, 0-3.0 GHz and 13.0-18 GHz at < -40 dB

Filter, bandpass image reject	(may be revised after re-analysis of spurious mixer responses)
4-8 GHz nominal	passband <4.1 GHz and >8.4 GHz at -1 dB, max ripple +/-0.5 dB; stopband 4.0 GHz and 8.6 GHz at < -10dB, 0-3.0 GHz and 10-18 GHz at < -40 dB
8-12 GHz nominal	passband <7.6 GHz and >12.0 GHz at -1 dB, max ripple +/-0.5 dB; stopband 7.4 GHz and 12.4 GHz at < -10 dB, 0-6.0 GHz and 14-18 GHz at < -40 dB
Filter, outputs A, B, C, D (may be revised after re-analysis of mixer and digitizer spurious responses)	passband <2.1 GHz and >3.9 GHz at -1 dB, max ripple +/- 0.5 dB; stopband 0-2.0 GHz and 4.0-12 GHz at < -20 dB
Passband amplitude ripple	<1.0 dB peak-peak
Passband deviation from linear phase	<40 degree peak-peak
Gain stability	<0.1 dB peak-peak over 1 minute, <0.5 dB peak-peak over 60 minutes
Phase/delay stability	<10 degree peak-peak over 1 minute, <40 degree peak-peak over 60 minutes
Headroom[1] when the antenna temperature is 290K	>20 dB
Crosstalk (inverse of isolation) among any input and any unconnected output	>40 dB rejection
Attenuators in input path 4-12 GHz	
Steps	1 +/-0.3 dB
Range	>30 dB
Phase variation vs. attenuation	<20 degree peak-peak over attenuation range 0-20 dB
Deviation from linear phase vs. frequency 4-12GHz	<20 degree peak-peak over attenuation range 0-20 dB

Attenuators in output path 2 - 4 GHz	
Steps, nominal	0.25 +/-0.15 dB over attenuation range 0-20 dB
Range, nominal	>30 dB

Phase variation vs. attenuation	<10 degree peak-peak over attenuation range 0-20 dB
Deviation from linear phase vs. input frequency 4-12 GHz	<10 degree peak-peak over attenuation range 0-20 dB
Matching among all downconverters	
Amplitude vs. frequency	TBD
Phase vs. frequency	TBD
Total power detectors (TPD)	
Input path 4-12 GHz	
Number	two, one for each input channel
Response vs. input frequency at any LO2 frequency	< 2 dB peak-peak.
Output path 2 - 4 GHz	
Number	four, one for each output channel A, B, C, D
Response vs. input frequency at any LO2 frequency	< 1.5 dB peak-peak.
Linearity	<1 % deviation from square law over range -6 dB to +13 dB relative to antenna temperature = 290 K
Monotonic resolution of digitizer, minimum	16 bits for 13 dB headroom above antenna temperature = 290 K
*Readout	2 millisec integrations and dumps to MC-AMBTP card via serial or parallel interface
*Offset calibration	MC sets the input power to zero by either setting the preceding attenuator to >(20 dB + headroom) or by removing bias to the preceding amplifier
Stability of output relative to inputs from front end	<50 ppm in 1 second, <500 ppm in 60 seconds
*Interface to MC-AMBTP	dedicated total power data link to antenna bus master (ABM)
***MC control functions**	
Set levels of input total power detectors	1 byte for each of two attenuators
Set levels of output of each total power detector and input level of each output digitizer	1 byte for each of four attenuators

Set to zero all inputs to total power detectors	1 byte to remove bias to six amplifiers; or set all attenuators to maximum
Set all 3 matrix switches (select image filters for each output)	1 byte

***MC monitor functions**	
Total power detectors	3 bytes every 2 milliseconds for each of 6 detectors
Temperatures	2 bytes every 10 seconds for each of 8 locations
Supply voltages derived within module	2 bytes every 10 seconds for each of 8 voltages
***External power supply inputs**	+18 +/-0.5VDC @ <2.2A, -18 +/-0.5VDC @ <0.7A, +8 +/-0.3VDC @ <0.6A, +5 +/-0.1VDC @ <0.6A
Internal voltage regulators	
Output voltages @ amperes	+15 @ 2.2 (total of >1 regulator), -15 @ 0.6, +5 @ 0.6, -5 @ 0.1
Output regulation plus noise	0.01% peak-peak over time interval > 60 seconds
Timing generator	
*Inputs from Reference Receiver	25 MHz sine wave at 0 dBm; 20.833 Hz positive edge, 5V differential
Output for timing total power integration	500 Hz TTL pulses of >1 usec width synchronized to 20.833 Hz timing reference
Output for digitizer clock	TBD MHz to match digitizer; synchronized to 20.833 Hz timing reference
***Operational environment**	
Altitude	5000 meter (16,000feet)
Shock	Negligible
Vibration	TBD

Temperature of air flow past sides of module	Plenum temperature set 16 – 22 Celsius, variation < 2 C peak-peak
Air mass flow rate past sides of module	TBD
Specific heat of air flow	TBD
Packaging	
Module	3 to 6 width x 5U high x TBD depth standard module (ATNF) with extruded vertical heat fins on one side or both sides
*Multi-pin connector (power, MC-AMB, MC-TP)	One double density 100 pin D type [male]
*Coaxial connectors	12 OSP (M/A-COM) blind mating [male]

1. Define *headroom* as the dB ratio of *available power at 1% gain compression {P(-1%)}* to the *total system noise power {P}*. Typically, P(-1%) is 16 dB less than the available power at –1 dB gain compression and 26 dB less than the available power at third order intercept. -end-

---

The Digitizers

9.2.1 Introduction, Top Level Specifications

The analog-to-digital converters, or digitizers, installed in the antennas provide the flexibility required for the fiber optic transmission of the IF. Signal digital conversion is of course indispensable to the correlator in order to derive the correlation function as a function of digital lags for spectroscopy. The digitizers are thus crucial and single-point-failure elements in the system. The ALMA system incorporates 3-bit digitizers thus improving the overall sensitivity compared to the classical 2-bit case.

The goal specifications are given in Table 9.2.1

Input BW 2-4 GHz

Sample clock 4 GHz (250 ps)

Bit resolution 3 bits

Quantization levels 8

Aperture time ~ 50 ps

Jitter a few ps

Threshold indecision region a few mV

Output demultiplexing factor 1/32 (125 MHz system clock)

PLL Clock distribution 4 GHz, 125 MHz (system)

Fine delay command

Low power consumption

ALMA Correlator

This section describes the ALMA correlator. The design described here is for a lag correlator with a system clock rate of 125 MHz. The goals of Phase 1 are to produce paper designs and some simulations of all major correlator elements, including the correlator chip, and to fabricate and test prototype hardware. The goals of Phase 2 are to produce a prototype minimally populated correlator, deliver such a prototype for use in the test interferometer, and deliver the complete correlator to the ALMA site.

Table 10.1 ALMA Correlator Specifications

Item	Specification
Number of antennas	64
Number of baseband inputs per antenna	8
Maximum sampling rate per baseband input	4 GHz
Digitizing format	3 bit, 8 level or 4 bit, 16 level
Correlation format	2 bit, 4 level
Maximum baseline delay range	30 km
Hardware cross-correlators per baseline	1024 lags + 1024 leads
Autocorrelators per antenna	1024
Product pairs possible for polarization	HH, VV, HV, VH (for orthogonal H and V)

Table 10.3 Selected correlator modes

# of Digitizers	Bandwidth/ Digitizer	Cross-pol Products?	Channels/ Product	At 230 GHz, in velocity space: Range	Resolution km/s
8	2 GHz	Yes	64	9391	40.8
8	2 GHz	No	128	18783	20.4
8	1 GHz	No	256	9391	5.1
8	500 MHz	Yes	256	2348	2.5
8	250 MHz	No	1024	2348	0.32
4	2 GHz	Yes	128	4696	20.4
4	1 GHz	No	512	4696	2.5

4	500 MHz	Yes	512	1174	1.3
4	250 MHz	No	2048	1174	0.16
2	2 GHz	Yes	256	2348	10.2
2	1 GHz	No	1024	2348	1.3
2	500 MHz	Yes	1024	587	0.64
2	250 MHz	No	4096	587	0.08

ALMA Computing, principal requirements

Sustained data rate, science data	6 MB/s (Average) 60 MB/s (Peak sustained)
Image pipeline	First-look images produced automatically for standard observing.
Dynamic scheduling	Nearly automatic scheduling of the array, accounting for current weather and other conditions, to optimize the scientific throughput of the array.
Archiving	Networked archive of all ALMA raw science data and associated calibration data and derived data products.

Antenna configuration on the Chajnantor site

Table 15.1 Guidelines for Configuration Design

Main D&D Task	Design a set of configurations which allow for a range of angular resolution and sensitivity
Flexible design philosophy	Configurations must allow for graceful expansion through possible collaboration
Costing	Optimize for shared stations to minimize cost
Site placement	Choose specific locations for antenna placement on Chajnantor site

The table below outlines different designs up to 3 km maximum baseline. A larger, 12 km baseline will now be included as well; the beam size at 345 GHz will be approximately 0.013 arc seconds. The choice between a donut or a spiral configuration is being discussed.

Table 15.3 Specifications for the ALMA strawperson configurations.

Array	Minimum	Maximum	Array	Time for	Natural
	Baseline	Baseline	Style	FOC = 0.5	Beam at 345 GHz
	[m]	[m]		[hours]	[arcs]
A	30	3000	donut	10	0.050
B	24	1430	donut	2	0.101
C	18	680	donut	0.1	0.22
D	16	325	donut	0.1	0.47
E	16	150	filled	0	0.97

Table 15.4 Specifications for the compact configuration N-S elongations.

Array	Min. N-S	Elev. of first	Min. observing	Max. observing	N-S
	Distance	Shadowing	Elevation	Elevation	Elongation
E1	1.3 D	50 deg	40-45	90	1.2
E2	1.9 D	31 deg	30	50+	1.6
E3	3.0 D	19 deg	14	33+	2.9

ALMA Project Book, Chapter 2.3: System Block Diagram

ALMA SYSTEM BLOCK DIAGRAM

Larry D'Addario et al.
Last changed 2001-Feb-05

Revision History:
2001-02-02: Revision D

Introduction

The system level block diagrams presented here are intended to show the high level design of the telescope. They depict the main astronomical signal path and all levels of its processing; the generation and distribution of auxilliary signals, such as timing references; and the organization and breakdown of the equipment into its logical parts.

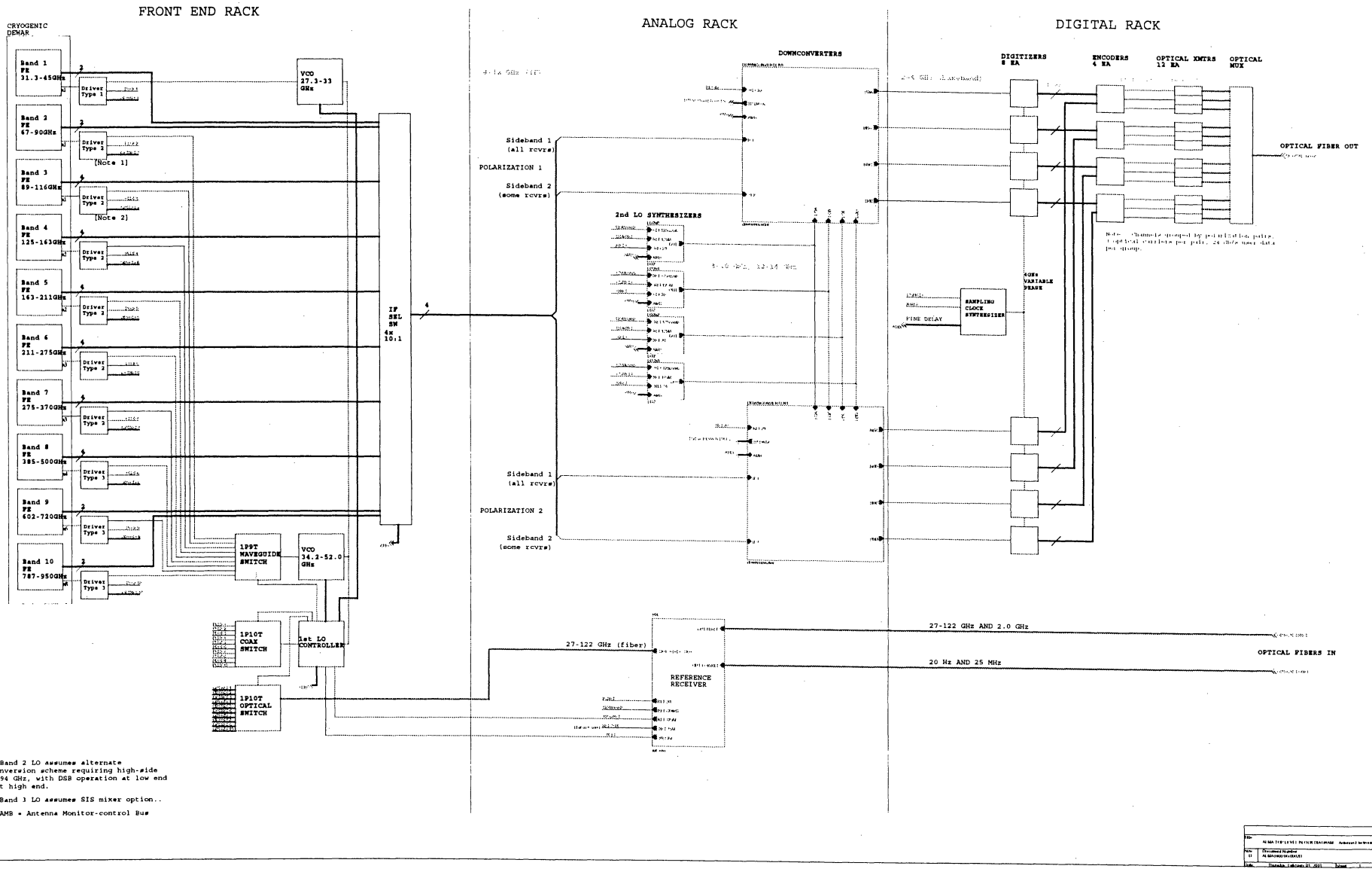
The diagrams do not necessarily show the packaging of the equipment -- that is, its physical breakdown into assemblies. The blocks are logical ones only.

The level of detail is necessarily limited, so as to depict the high level structure clearly. However, any other diagrams that show additional detail must be consistent with these. Any discrepencies between what is shown here and other engineering documents (e.g., those of subsystems or specific assemblies) must be reconciled. Proposals for designs that require deviating from what is shown here should be reviewed before being adopted.

Some important elements of the telescope are missing from the present versions of these diagrams, but they are intended to be added in future versions. These include most elements of the Monitor/Control system; the antenna and associated hardware supplied by contractors; power distribution and conversion; and certain calibration instruments.

See the ALMA System Block Diagrams

ALMA ANTENNA ELECTRONICS



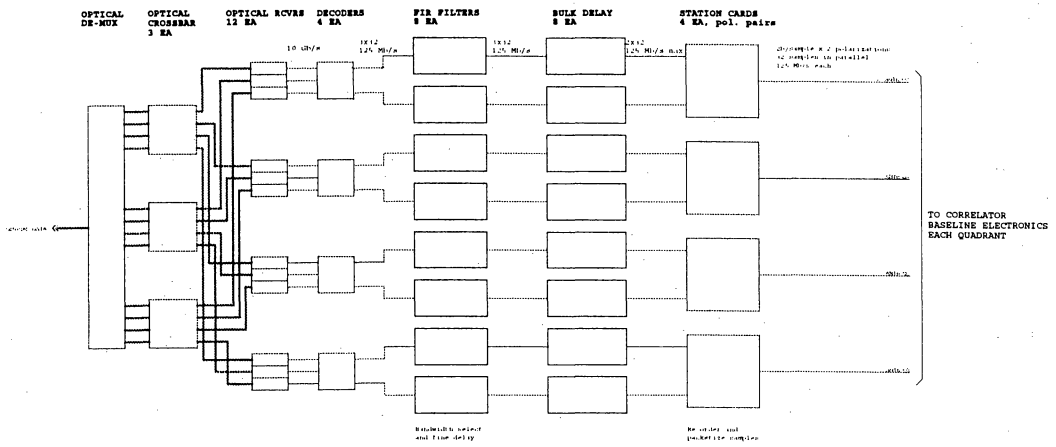
te 1: Band 2 LO assumes alternate ngle-conversion scheme requiring high-side at 79-94 GHz, with DSB operation at low end d LSB at high end.

te 2: Band 3 LO assumes SIS mixer option..

te 3: AMB = Antenna Monitor-control Bus

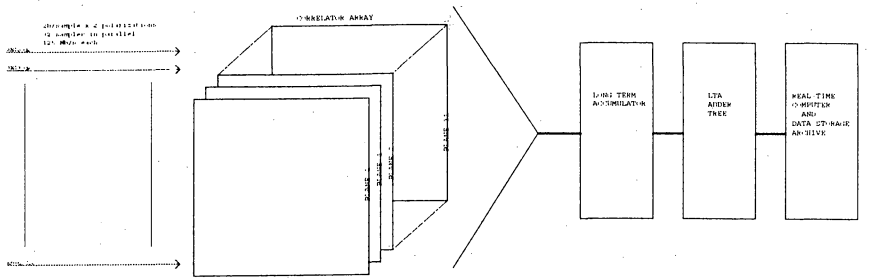
ALMA CENTRAL ELECTRONICS

CORRELATOR STATION ELECTRONICS

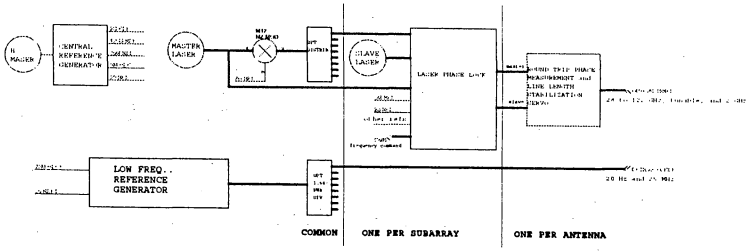


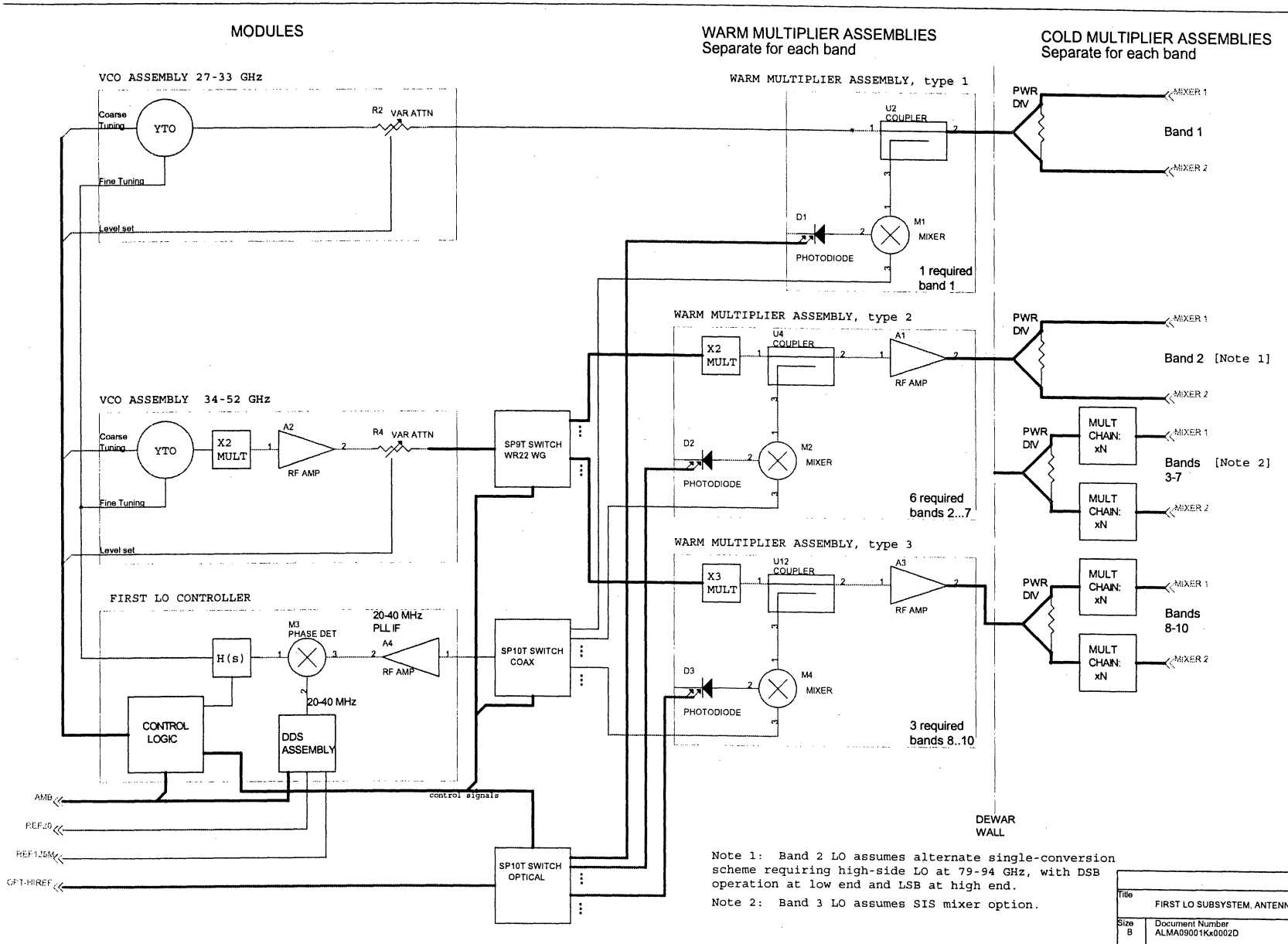
CORRELATOR BASELINE ELECTRONICS

(ONE OF FOUR QUADRANTS)



LOCAL OSCILLATORS





Band	Driver	N	Output
1	27.3-33	1	27.3-33
2	79-94	1	79-94
3	96-104	2	96-104
4	68.5-75.5	2	137-151
5	87.5-99.5	2	175-199
6	74.3-87.6	3	223-263
7	71.7-89.5	4	287-358
8	99.2-122	4	397-488
9	102.3-118	6	614-708
10	99.8-117.3	8	799-938

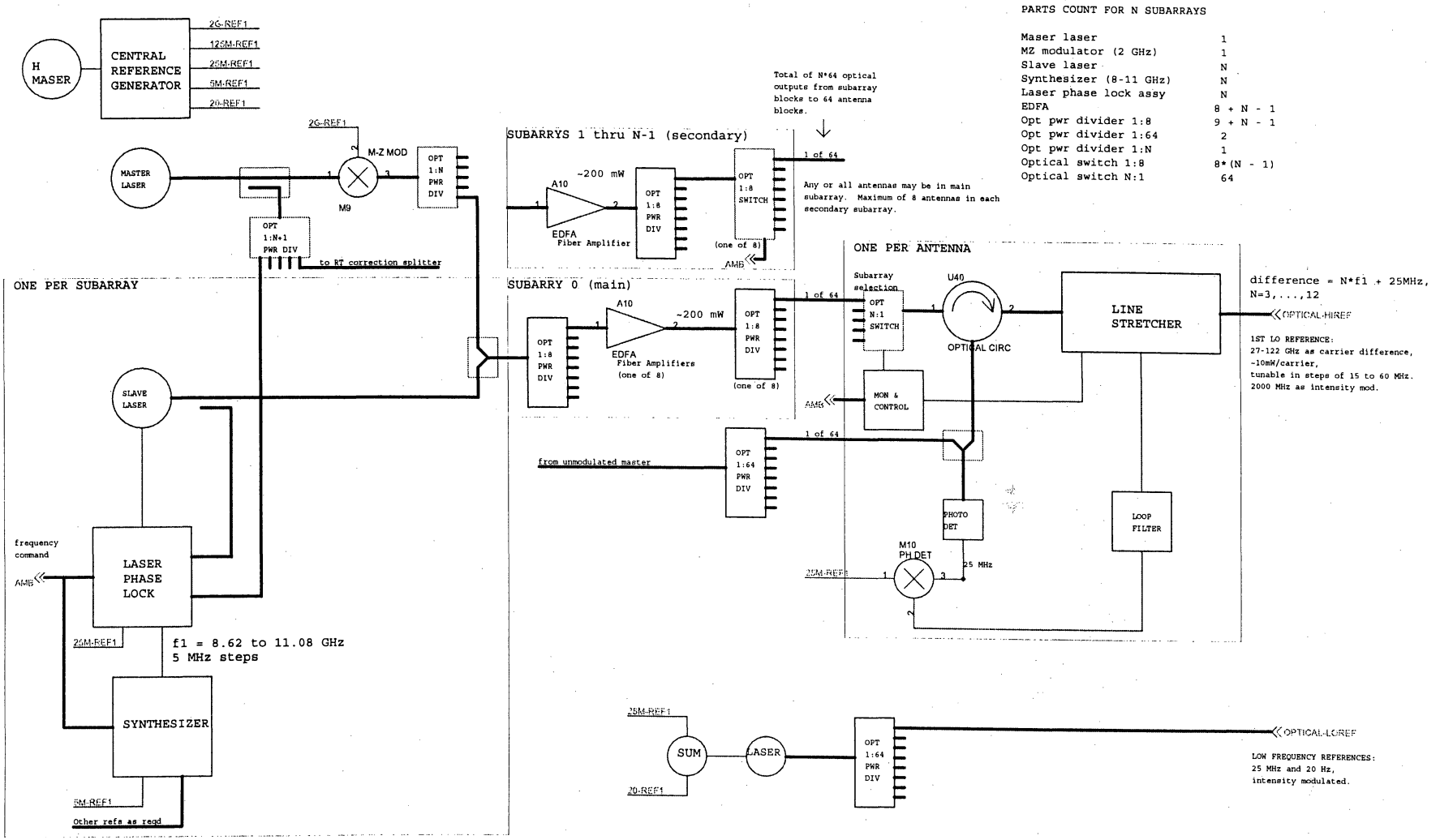
Note 1: Band 2 LO assumes alternate single-conversion scheme requiring high-side LO at 79-94 GHz, with DSB operation at low end and LSB at high end.

Note 2: Band 3 LO assumes SIS mixer option.

Title: FIRST LO SUBSYSTEM, ANTENNA PORTION

Size: B | Document Number: ALMA09001Kx0002D | Rev

Date: Thursday, February 01, 2001 | Sheet 3 of 10

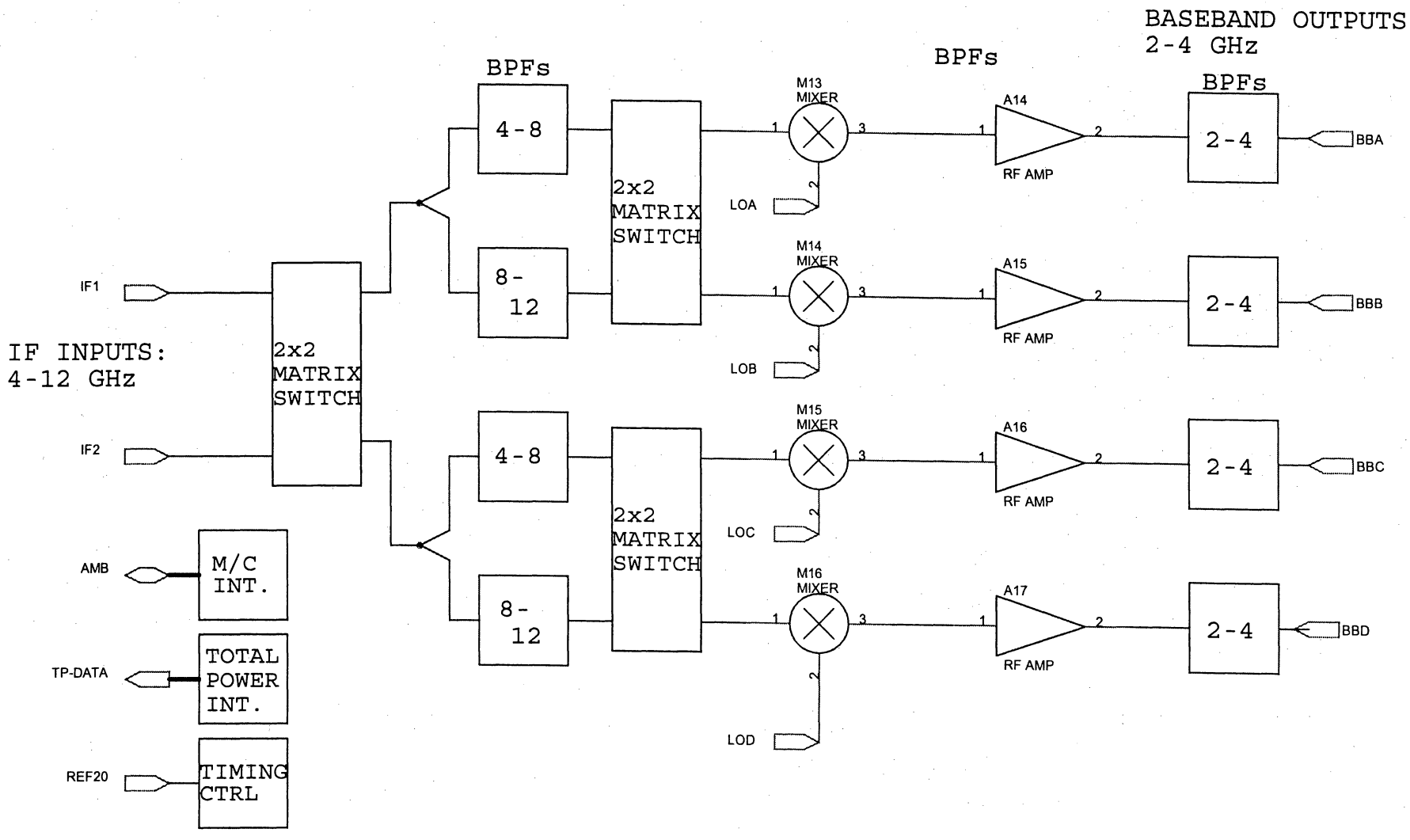


PARTS COUNT FOR N SUBARRAYS

Maser laser	1
MZ modulator (2 GHz)	1
Slave laser	N
Synthesizer (8-11 GHz)	N
Laser phase lock assy	N
EDFA	8 + N - 1
Opt pwr divider 1:8	9 + N - 1
Opt pwr divider 1:64	2
Opt pwr divider 1:N	1
Optical switch 1:8	8*(N - 1)
Optical switch N:1	64

LO DETAILS -- CENTRAL

Title: ALMA TOP LEVEL BLOCK DIAGRAM -- LOs, central portion		
Size: B	Document Number: ALMA09001Kx0002D	Rev
Date: Thursday, February 01, 2001	Sheet 4 of 10	

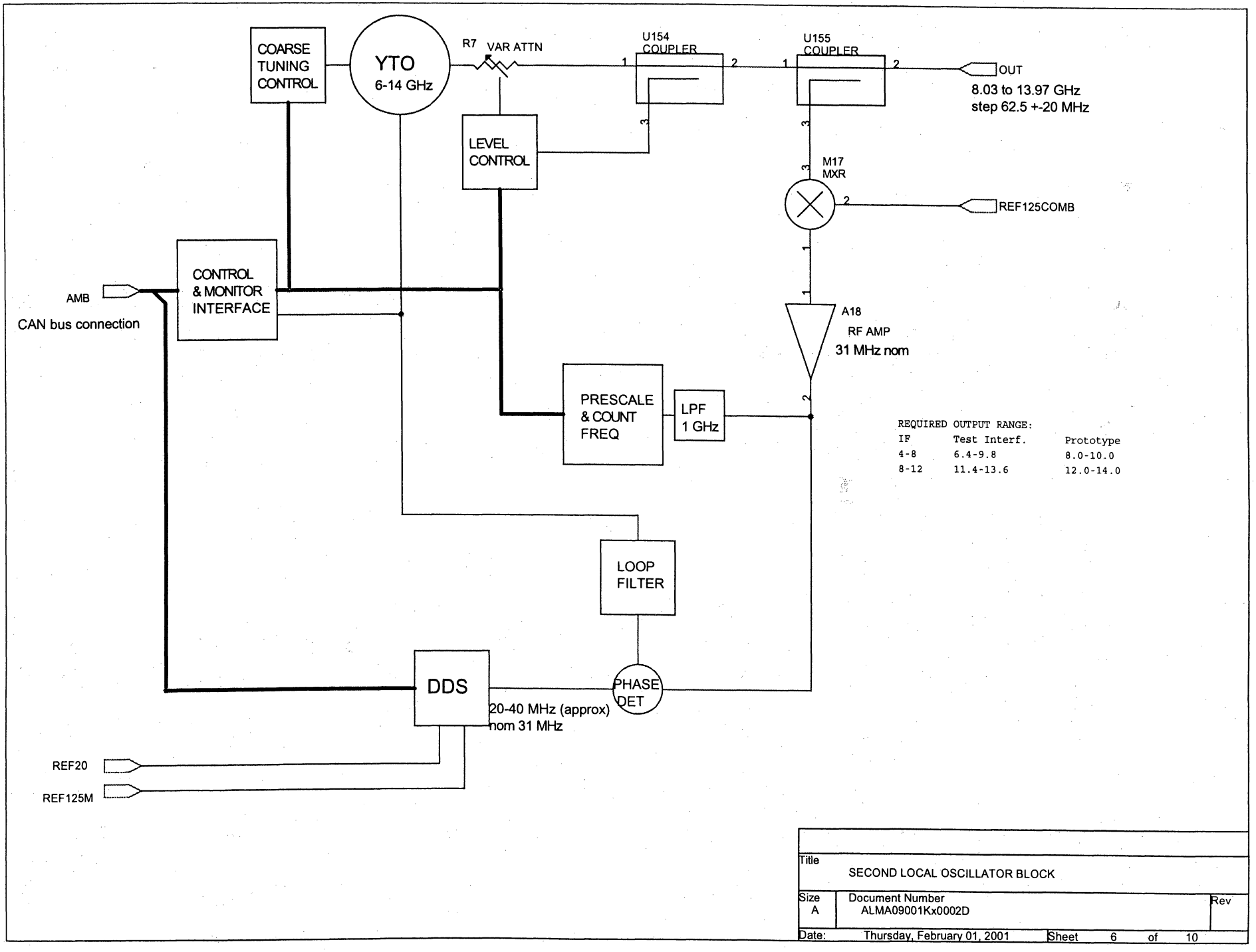


IMPORTANT DETAILS NOT SHOWN:

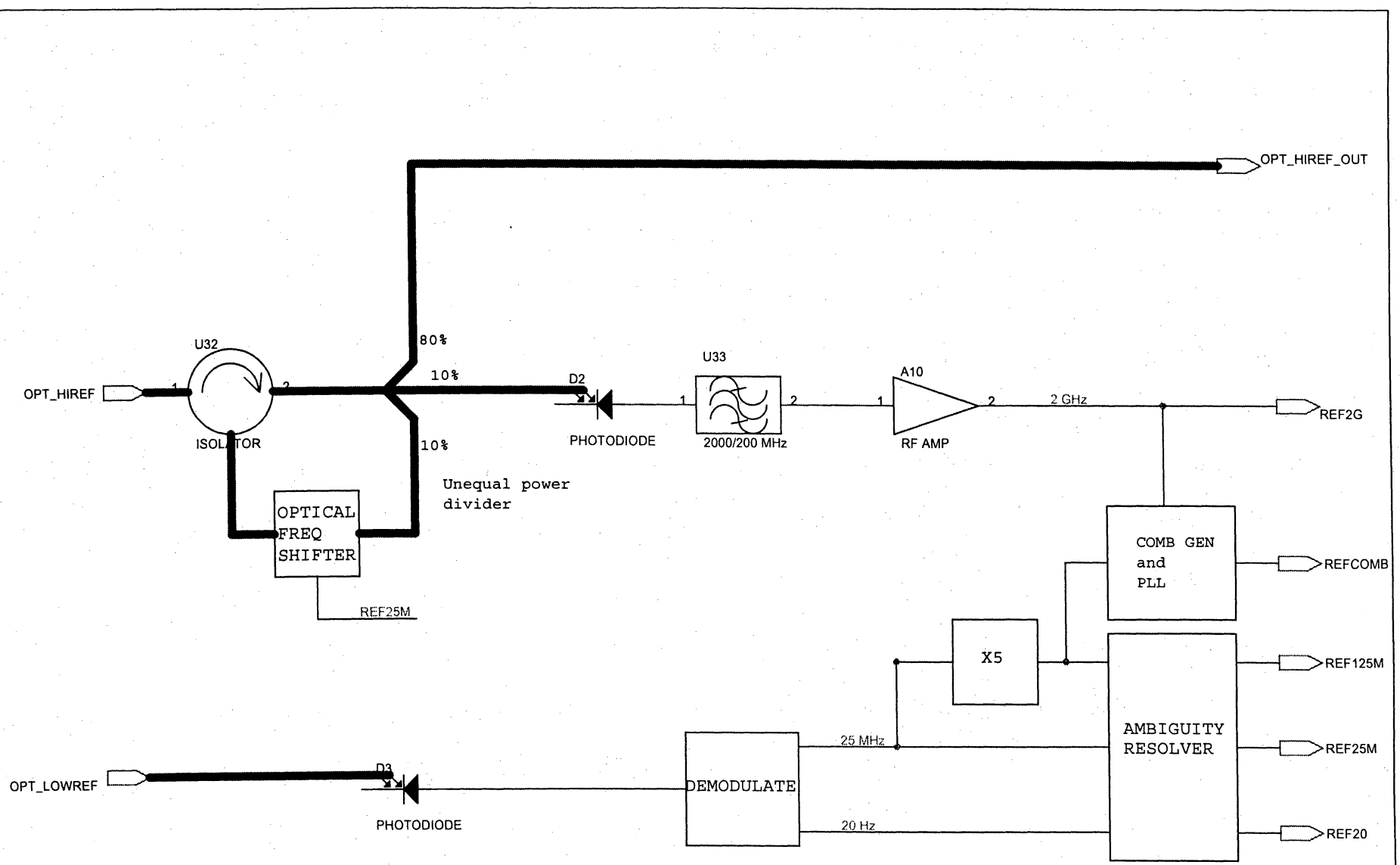
1. Square law detectors required for each IF input signal (4-12 GHz) and each BB output signal (2-4 GHz).
2. Gain must be adjustable via command from computer, resolution 1dB or less.

Test Interferometer Note: A special version of this module for use in the TI will differ only in the output bandpass filters. Two will cover 1.6-2.4 GHz and the others will cover 1.6-1.7 GHz, rather than 2-4 GHz..

Title: DOWNCONVERTER BLOCK		
Size: A	Document Number: ALMA09001Kx0002D	Rev
Date: Thursday, February 01, 2001	Sheet 5 of 10	



Title	SECOND LOCAL OSCILLATOR BLOCK	
Size A	Document Number ALMA09001Kx0002D	Rev
Date: Thursday, February 01, 2001	Sheet 6 of 10	



Ambiguity resolution is accomplished by capturing each signal on the next positive zero crossing of the next faster signal, using fast flip-flops. This transfers the phase stability of the fastest signal to the others. It requires an initial timing adjustment and then stability better than about 20% of the period of the next faster signal.

Title	ALMA LO: Reference Receiver Block	
Size A	Document Number ALMA09001KX0002D	Rev
Date: Thursday, February 01, 2001	Sheet 7 of 10	

ALMA Construction Project Book, Chapter 3.1: Calibration Issues

Calibration Issues for the ALMA

Mark Holdaway

Revised by Al Wootten & others

Last changed 2001-Feb-06

Revision History:
1998-Nov-03: Format modified to Project Book Standard
1998-Nov-11: Memo references and text brought up to date
1998-Nov-11: Memo references, text brought into ALMA standard
2001-Feb-06: Update on reference to nutator design, and some minor stylistic changes.

Summary

The precision imaging to be attained by the ALMA will be achieved through accurate calibration. The types of calibration are summarized in Table 3.1.

Table 3.1 ALMA Calibration Requirements.

Pointing	0.6" absolute
Primary Beam	2-3%
Baseline Determination	0.1 mm
Flux Calibration	1% absolute flux accuracy goal
Phase Calibration	0.15 radian at 230 GHz
Bandpass Calibration	10000:1 to 100000:1
Polarization Calibration	10000:1
Single Antenna Calibration	Employed

Calibration strategies will be developed and implemented on the OVRO and BIMA arrays. The principal goals to be achieved by the end of the ALMA Design and Development (D&D) Phase are to demonstrate radiometric phase correction at 22 GHz, including demonstration of correction algorithms at OVRO and BIMA; and to demonstrate improved instrumental amplitude calibration.

3.1.1.0 Introduction

Calibration is the process by which the astronomer converts electronic signals from the telescope into meaningful astronomical data. Calibration is crucial for the ALMA. As millimeter and submillimeter wavelength radiation will be adversely affected by the atmosphere and the electronic signal path in a variety of ways, and as the antennas will also be affected by the observing environment, we must understand how we will correct for or remove these various effects. The calibration strategy interacts heavily with the science requirements, the system design, the receivers' functioning, the antenna's physical behavior, the site conditions (*i. e.*, site characterization), the scheduling of the telescope, the real time computing, and the post-processing software which will take most of the burden of implementing the required calibration schemes.

Never before has a radio astronomical instrument been built with such a detailed understanding of the site and its impact on the telescope. With this knowledge in hand, we can optimize the full calibration strategy to produce the maximum scientific output for the ALMA.

We are convinced that all the calibrations that are required can be effectively achieved, but we have not yet made all the decisions as to how to achieve these calibrations. Furthermore, many of the options which we lay out for the various calibrations interact with each other, so we have a long way to go yet before we have a coherent picture of all calibration systems and their interdependencies. This chapter tries to express the astronomical requirements for the various calibrations which need to be performed, as well as the hardware and software requirements for the competing methods for performing these calibrations.

3.1.2.0 Pointing

3.1.2.1 Astronomical Requirement for Pointing

Cornwell, Holdaway, and Uson (1993) show a requirement of 1 arcsec for pointing accuracy, based on the requirements of good mosaicing image quality with 8m diameter primary aperture. Observations at the highest frequencies (900 GHz) will also require this pointing accuracy for even single pointing observations, and high frequency mosaics would often benefit from even better pointing. However, a large fraction of the ALMA observations, such as single pointing observations at frequencies up to 500 GHz, mosaic observations at 115 GHz or lower frequency, or low SNR mosaics at millimeter wavelengths will not require this precision pointing spec.

With an increase in D from the ALMA 8m to the ALMA 12m, the pointing spec should tighten to 0.6 arcsec. Holdaway (1997; Memo 178) performed a more detailed analysis, showing the effects of pointing errors ranging from totally random to totally systematic. While we do not divide up the 0.6 arcsec pointing error specification into various systematic and random terms here (but see Table 3.4.3-2 of the ALMA Antenna RFP), we note that the effects of any pointing error budget with various systematic and random terms could be translated to an estimated image quality. We do note here that random errors have less effect than systematic errors, but we also assume that systematic errors can be calibrated out.

With only minor exceptions, pointing calibration must be performed prior to astronomical observations or the data are useless. This also means that we cannot generally interpolate pointing solutions backwards

in time. This makes pointing calibration all the more crucial.

3.1.2.2 What Affects Pointing?

The antenna pointing will be affected by several slowly varying terms such as systematic imperfections of the antenna and the pad, gravitational forces, and thermal loading from the sun. Depending upon the strategy of the astronomical observations, much of these slowly varying effects can be removed by frequent offset pointing observations Some active corrections might also be incorporated into the antenna design. Thermistors have been employed on existing antennas to monitor solar thermal response. Lamb and Woody investigated the use of tiltmeters and a carbon fiber reinforced plastic (CFRP) pointing reference structure in their investigation of a 12m antenna design. Lugten (ALMA Memo 232) discussed implementation of a CFRP reference structure in detail, showing that use of such a structure might improve pointing performance.

In addition to these slowly varying systematic pointing errors, there will also be highly random pointing errors caused by wind loading and anomalous refraction. Measurements of the wind indicate that a great deal of the power in the wind is often in a constant speed and direction, so it is only the gusts about this mean speed and direction which will result in differential pointing errors between the calibration and the target source (see Holdaway 1996, Memo 159). The refractive pointing will usually not be a severe problem, but will sometimes limit the pointing (Holdaway, 1997, Memo 186; Butler, 1997, Memo 188; Holdaway and Woody, Memo 223, Lamb and Woody Memo 224). Since the refractive pointing is random on time scales of the antenna crossing time of the atmosphere (ie, 1 s), we mainly need to have statistically many different atmospheric instantiations for each of the five points of a pointing calibration to ensure that we are not applying an erroneous pointing position when we collect data on the target source.

Finally, at some level there will be a limit to the mechanical repeatability of the antenna pointing. At this level, we are left with completely random pointing errors which cannot be calibrated. If these purely random errors are too large, they will spoil the imaging characteristics of the ALMA and will not be correctable. If they are small to moderate in magnitude (ie, < 0.5 arcsec), we can tolerate them quite well as these random errors are the least damaging of any pointing errors.

3.1.2.3 Pointing Calibration Strategy

Our basic goal is to have systematic pointing errors which can be calibrated and removed completely so we are left with purely random pointing errors which are small enough not to bother us.

3.1.2.3.1 Pointing Model

The first step in removing the systematic pointing errors is to determine the systematic imperfections of the antenna and pad and the effects of gravity. Most radio telescopes periodically undergo a pointing routine which samples the sky with pointing measurements on about a hundred astronomical sources taken at night to minimize thermal and wind pointing errors. The ALMA will take about 60 minutes to perform 100 pointing calibrations across the sky. The results of these pointing measurements are then used to fit about 10 parameters in a pointing equation which accounts for various physical terms, such as misalignment of optical axes or four fold sag due to the antenna base being supported at four locations. Some experimenting will go into determining the optimal form for the ALMA pointing equation (see Mangum Memo 288).

The ALMA will probably rely heavily upon a lower frequency (e.g. 90 GHz or below) system for determining the pointing model. The wider beam at this low frequency and the high sensitivity and bright astronomical sources will facilitate pointing measurements, even after a reconfiguration. However, precise pointing offsets among the different frequencies will also need to be determined. It is our hope that the blind pointing after application of the pointing model is on the order of 2 or 3 arcsec rms. To achieve the precision pointing specification, frequent (i.e. every 30 minutes) offset pointing calibration will be required, to remove local deviations from the pointing model, or systematic slowly varying effects due to wind and/or thermal gradients.

3.1.2.3.2 Offset Interferometric Pointing on Quasars

Holdaway (1996; Memo 159) and Lucas (1997; Memo 189) have both demonstrated the feasibility of performing pointing calibration on quasars close to the target source; adequate SNR can be obtained with sufficient speed. The minimum calibrator flux, and hence the typical minimum distance to a pointing calibrator for pointing calibration, is a function of both the collecting area and the number of elements in the array.

A key question concerning the efficiency of these offset pointing calibrations: what are the differential pointing errors as a function of distance between cal and target sources? This depends upon the direction of wind, the sun angle, etc, and probably needs to be answered experimentally. Also, we must understand the stability of the pointing offsets among the different frequency bands, as the pointing calibration will usually be done at 30 or 90 GHz.

Both the pointing model solution and the offset interferometric pointing will require an extensive, up-to-date catalog of pointing calibrator sources, and the observing schedule program should allow for automation of the choice of a pointing calibrator and the pointing calibration strategy.

3.1.2.3.3 Infrared Pointing

The ALMA prototype antennas, and perhaps the production antennas will be outfitted to perform infrared or optical offset pointing. Offset interferometric pointing should work well enough. However, infrared pointing will increase the overall efficiency of the instrument, may improve the antennas' pointing and may help characterize mount components in the pointing equation. The infrared pointing will be largely immune to refractive pointing. Since our main strategy concerning refractive pointing is to minimize its random effect on the pointing measurements, the infrared pointing is not at a disadvantage because it is not affected.

3.1.2.3.4 Scheduling and Editing

A particular experiment's demands for precision pointing need to be combined with the current site conditions (phase stability for refractive pointing, wind and wind stability, and solar loading and variability) and qualitative rules of thumb for the success of pointing calibration and pointing stability in various conditions to determine when that experiment should be scheduled. Records of the pointing solutions, phase stability, wind fluctuations, and solar loading can be used to locate times when the pointing solutions are suspect, and the astronomical data during these times can be edited accordingly. If the pointing solutions show large drifts with time, but all antennas are behaving similarly, the mean array pointing position as a function of time can be corrected in post-processing.

3.1.2.3.5 Pointing Self-Calibration

Pointing self-calibration, an unimplemented algorithm, may not work well at all unless there is at least one bright point-like source within the target area. Once the antenna pointing offsets have been determined, it is simple and not too cpu-expensive to apply the mean array offset as a function of time to mosaic or single pointing data and use that in the imaging. However, if there is significant scatter among the antennas' pointing positions at each time, imaging wide field sources considering the correct pointing data may be prohibitively expensive (Holdaway, 1993; Memo 95).

3.1.3.0 Primary Beam Calibration

A detailed understanding of the primary beam will be required to image wide field astronomical objects. At lower frequencies, pointing errors will tend to limit mosaic image quality, but at high frequencies, surface errors resulting in primary beam distortions will dominate the errors in mosaic image reconstruction (Cornwell, Holdaway, and Uson, 1993). Early simulations (Holdaway, 1990; Memo 61) indicated that an understanding of the primary beam down to the 2-3% level is desirable. If this cannot be achieved, mosaic dynamic range will be limited by primary beam uncertainty more than pointing errors.

The desired primary beam information will result naturally from the low frequency holography campaigns. While the high frequency primary beams will not simply scale like the frequency, we can estimate them from a surface error model and the feed placement. However, we will probably want an independent measurement of the beam at several frequencies. At the highest frequencies where we expect the most problems with the primary beams, the primary beams will be hardest to measure due to decreased sensitivity and a lack of appropriate sources.

We may require primary beam models with different levels of complexity to achieve different scientific goals. The simplest primary beam model, which will suffice for low to intermediate dynamic range observations, will be a mean rotationally symmetric primary beam, measured out to the first or second sidelobe. The next level of complexity may be a mean 2-d (*i.e.*, non rotationally symmetric) beam. It is conceivable that we would someday require the use of different 2-d primary beams or voltage patterns for each antenna, for several different elevation angles, or even for both.

3.1.4.0 Baseline Determination

The relative positions of the antennas must be determined accurately so the geometrical delay can be correctly applied to the antenna voltages prior to correlation. Residual delays will result in phase errors which change across the observing band and differential phase errors between two different sources on the sky. For the ALMA with, *e.g.* an 8 GHz bandwidth, a reasonable limit of 1/3 radian phase difference across the band requires a baseline accuracy of about 1 mm. Requiring the differential phase error between two sources 5 degrees apart on the sky to be on the order of 5 degrees results in a baseline accuracy of about 0.1 mm. Atmospheric phase errors of more than 5 degrees would not severely impact the imaging, as these errors are random in time and among antennas. However, baseline errors will be partially systematic as they will be slowly varying in time, so we need to be more conservative with them than with the atmosphere.

3.1.4.1 Baseline Measurement Strategy

The baselines, or delays, may be measured by determining the delay on each baseline for on order of a hundred observations of point sources sampling the entire sky. Individual delays can be fit across the spectrum, as in VLBI. The complete set of delays is used to solve for the three dimensional locations of all antennas relative to a reference antenna. The observing strategy is similar to that for the pointing model determination, and should take about an hour to complete. Signal to noise is not an issue for 0.1 mm accuracy, and the 1 hour time scale is set more by the minimum time to sample many sources around the sky. Atmospheric phase fluctuations may affect the baseline delays, so ideally the observing conditions should be excellent. In poor conditions, the delays can probably still be determined based on the statistics of many differential measurements, as the atmosphere should tend towards a zero mean in differential measurements.

Of some concern is the time scale over which we can expect the antenna positions to remain fixed to within 0.1 mm. Permafrost has been anecdotally reported on the ALMA site, which enables an entire class of soil movements. We can probably expect some amount of soil creep, especially after earthquakes. We will gain experience concerning the frequency of baseline calibration once the ALMA begins to be operational at Chajnantor.

3.1.5.0 Flux Calibration

3.1.5.1 Astronomical Requirements for Flux Calibration

Yun et al. (ALMA Memo 211) have written a white paper on flux calibration which addresses most of the issues mentioned here. The primary scientific requirement for accurate absolute flux calibration is the comparison of astronomical images made at different frequencies. While the current 10% estimated absolute accuracies permit many qualitative conclusions, 1% absolute flux accuracy will really open up new quantitative scientific possibilities. This flux calibration accuracy must apply to both total power and interferometric modes.

3.1.5.2 What Affects Flux Calibration?

Changes in the receiver temperature, electronic gain drift, variable atmospheric opacity and emission, variable ground pickup, decorrelation (atmospheric and electronic) and gravitational deformation at high frequencies are some of the parameters affecting accurate flux calibration.

3.1.5.3 Flux Calibration Strategy

3.1.5.3.1 Instrumental Amplitude Calibration

The currently used ambient load chopper wheel method is accurate to about 5% (see Ulich & Haas (1976), Kutner (1978)). Bock et al. (1998; ALMA Memo 225) are investigating a two load system located behind the secondary and viewed through a hole in the secondary; see also ALMA Memo #318 by Jeff Mangum: "Amplitude Calibration at Millimeter and Sub-millimeter Wavelengths". This system could theoretically achieve 1% amplitude calibration of a well understood antenna/receiver system, but would rely upon accurate ancillary measurements of the decorrelation and the opacity. BIMA is prototyping this system, and is evaluating how well it works. An accurate estimate of the atmospheric opacity will be required for accurate flux calibration. At frequencies at which the atmosphere is opaque at the low elevation angles and partially transparent at high elevation angles, it will be possible to solve

for the sky temperature and the opacity with a sky tip. However, at frequencies at which the atmospheric transmission is excellent, a sky tip will only give the product of the opacity and the sky temperature, and the temperature must be assumed to calculate the opacity. Currently, atmospheric models are not sufficiently accurate to measure the opacity and sky temperature at a partially opaque frequency and accurately estimate the opacity at another frequency. The Water Vapor Radiometry (WVR) system should provide a measure of emission and absorption at its operational frequency of 183 Ghz. Actual measurement of the sky temperature and water vapor profiles via radiosonde are currently underway, along with modeling of the data. Continuous radiosonde monitoring, however, seems unwieldy and expensive. A more cost effective solution would be to float a tethered balloon over the site several times a day. The temperature, pressure, and water vapor information would also be useful for the radiometric phase correction schemes.

3.1.5.3.2 A Gain-Based Instrumental Amplitude Calibration

Larry D'Addario points out that tracking Tsys fluctuations may not be the easiest way to get good flux calibration. He points out that with a multi-bit correlator, we can measure the correlated power instead of the correlation coefficient; hence, we need to track the electronic gain from receiver to correlator. The electronic gain does not vary with atmospheric opacity, changing ground pickup, etc, and is therefore much easier to track, perhaps by injecting a broad-band signal through the system. Even so, we still need an accurate opacity measurement.

3.1.5.3.3 Astronomical Flux Calibration

While an instrumental amplitude calibration accurate to 1% is a desirable goal, it is unclear that we will be able to understand the antennas and the atmosphere well enough to achieve this ambitious goal, so it is important to find astronomical sources that could serve as flux standards. Currently, planets are used as astronomical flux standards for millimeter wavelength observations, but the estimated accuracy of the planets' fluxes is only about 10%, so new flux standards need to be found.

The ALMA will have the sensitivity to use much fainter sources as flux standards. At the highest frequencies (650 and 900 GHz windows), some stars will be bright enough to achieve the 1% flux calibration goal in a few minutes. A knowledge of their temperatures from optical data will determine the expected submillimeter flux, but this may be complicated by confusing emission from dust or even time variable gyrosynchrotron emission from sunspots or flares. Some efforts at stellar measurements are underway at BIMA to evaluate this possibility.

As the blackbody spectrum falls off very fast at lower frequencies, millimeter observations will not be able to use stars as astronomical flux standards. However, although still subject to the same blackbody spectral law, asteroids appear promising. Many are unresolved to many ALMA baselines, the bright ones are in the range of 50-1200 mJy, permitting fast detection at 1% accuracy, and they are fairly simple systems. One drawback is their non-uniform emission as they rotate and as they move with respect to the sun, which could change their flux by several percent. Observational tests are required on these prospective astronomical flux standards as the ALMA comes on line, and we can always hope for a less problematic class of sources for an accurate astronomical flux standard. See Yun et al. (1998) for a more detailed discussion of stars and asteroids for flux standards.

3.1.5.3.4 Phase Decorrelation

An uncorrected antenna based phase error of 10 degrees rms will result in a 3% decrease in the visibility

amplitude due to decorrelation. As the characteristics of the phase noise change, the amount of decorrelation will also change. The primary defense against decorrelation is to try to correct the phase as much as possible. However, when the phase cannot be fully corrected, we can estimate the magnitude of the decorrelation and correct the visibilities. Decorrelation could be estimated from:

- phase calibrator data (fast switching)
- independent phase monitor (atmospheric) PLUS injected LO signal (antenna mechanical & electronic)
- radiometric data (atmospheric) PLUS injected LO signal (antenna mechanical & electronic)

For atmospheric coherence times, see Holdaway 1997 (Memo 169).

3.1.5.3.5 Changes in Transparency

At millimeter wavelengths, the changes in atmospheric transparency will be very modest, under 1% over 10 minutes about 80% of the time. Since the same amount of water vapor results in much larger opacities in the submillimeter, the transparency fluctuations in the submillimeter over characteristic calibration time scales will be much larger, typically several percent during median stability conditions. Due to the lack of submillimeter calibration sources available for fast switching and the current uncertainty in the transmission models, we will probably need to perform frequent tipping measurements to solve for the transparency at the observed frequency.

3.1.5.3.6 Polarization Complications

As mentioned below under polarization calibration, if a linearly polarized calibration source is used to track changes in the amplitude gain or opacity, a telescope with linear feeds will produce parallel hand visibilities which are modulated by the linear polarization. The extra signal varies as a sinusoid of the parallactic angle, so the errors are systematic. This will not be a problem for the astronomical flux calibrators mentioned here, but would be a problem for quasars.

3.1.6.0 Phase Calibration

Phase errors limit resolution, limit the dynamic range of images, introduce artifacts, and reduce sensitivity by decorrelation. Without effective phase calibration, the maximum usable ALMA baseline would generally be about 300 m. Amplitude errors would limit image dynamic range and skew the flux scale.

3.1.6.1 Astronomical Requirements for Phase Calibration

The phase calibration working group report (Woody 1995; Memo 144) considered three cases at 230 GHz: high quality imaging with 8 deg phase errors, median conditions with 19 deg phase errors, and poor imaging with 48 deg phase errors. The phase errors have a budget which includes the atmosphere, the antenna, and the electronics.

3.1.6.2 What Affects the Phase?

At millimeter wavelengths, the main atmospheric constituent which causes phase errors is inhomogeneously distributed water vapor. Up to about 300 GHz, atmospheric water vapor is very nearly non-dispersive. Above 300, water vapor can be quite dispersive, especially near the water vapor lines in

the atmosphere. Submillimeter wavelength observations will need to account for this dispersion if the phase is being calibrated indirectly (*i.e.*, scaled from a lower frequency or determined by scaling the differential water vapor column as determined by water vapor radiometry).

The dry air results in a major contribution to the absolute phase. If there are appreciable temporal or spatial fluctuations in temperature or pressure in the dry air above the array, phase fluctuations will result. Furthermore, the absolute dry air phase depends upon the observing elevation angle and the topographical elevation, which will change from one source to another. It is believed that the dry phase is non-dispersive at millimeter wavelengths.

Any change in the distance between the subreflector and the feed will cause phase errors.

The stability of the LO and other electronics will also influence the phase.

3.1.6.3 Phase Calibration Strategies

3.1.6.3.1 Fast Switching

If a calibrator is sufficiently close and the telescope is sufficiently fast, fast switching between a calibrator source and a target source can effectively stop the atmospheric, electronic, and antenna phase fluctuations. If fast switching is used as the phase calibration method, it makes minimum requirements on the system sensitivity, the slew speed and settle down time of the antennas, and the online and data taking systems. Fast switching has been studied extensively (ALMA Memos 84, 123, 126, 139, 173, 174, 221, 262), and we are fairly confident that it will work for the ALMA.

3.1.6.3.1.1 Sensitivity Requirements

The basic criteria for fast switching to work is that the phase calibration source needs to be detected with sufficient SNR and the target source be observed for some amount of time within the coherence time and distance of the atmosphere. This translates into a requirement that there be sufficiently many calibrator sources which are sufficiently bright (Memo 123), and a requirement on the sensitivity of the array. In practice, this means that the calibrator source will typically be within a degree of the target source, the calibrator will usually be detected in less than a second, and the entire cycle time will be about 10 s, though the details vary with observing frequency. Spectral line observations will need to use wide bandwidth continuum observations of the calibrator.

With the current sensitivity of the ALMA and our understanding of the quasar source counts and their dependence on frequency, we will not always be able to perform fast switching calibration at the target frequency, but often we will get a higher SNR phase solution by observing the calibrator at a low frequency (like 30 or 90 GHz) and scaling the solution up to the target frequency.

3.1.6.3.1.2 Scaling the Phase to High Frequency Observations

The falling source counts and sensitivity at high frequency will often require fast switching to observe calibrators at low frequencies and scale the phases up to the observing frequency of the target source. This requires a much more accurate phase solution at the lower frequency. Since the dry atmosphere is non-dispersive, this extrapolation basically relies upon the wet differential delay to be non-dispersive as well. In the submillimeter, the wet differential delay is dispersive, which will either limit the effectiveness of fast switching or require more complications in the fast switching observing strategy, such as less frequent multi-frequency calibrator observations to help separate out the non-water vapor

phase contributions.

On longer timescales, set by the phase stability of the electronics, it is necessary to measuring the instrumental phase offset between different frequencies, by observing a single source at both wavebands. See section 7.0.6 of the Local Oscillators chapter of this Project Book for a discussion of this point.

3.1.6.3.1.2 Requirements on Antenna Movements

The antenna movement requirement is currently a slew of 1.5 degrees and settle down to 3 arcsec pointing in 1.5 seconds.

3.1.6.3.1.3 Requirements on Antenna and Electronics Stability

At the very least, the antenna needs to be mechanically stable to within a small fraction of a wavelength (ie, 5-10 degrees at the target frequency) over a calibration cycle time, even when the antenna is moved by a few degrees on the sky. Similarly, the electronics need to be equally stable over the calibration cycle time. However, if we are to succeed in the submillimeter, the antenna and electronics need to be stable over much longer times, such as the cycle time between the multifrequency observations required to separate the wet and dry phase errors. These performance requirements are specified in the Anntena Request for Proposal.

3.1.6.3.1.4 Requirements on Computing

The on-line system needs to control the antennas gracefully enough to move them quickly without exciting the lowest resonant frequency. Also, the quanta of integration time and scan length need to be sufficiently small so as not to restrict the integration time spent on the target source and calibrator or the time spent between sources. Flexibility at the 100-200 ms level is desirable. Fast switching data can be calibrated with existing software, but some extensions in spatial-temporal interpolation will be useful.

3.1.6.3.1.5 Sensitivity Loss from Fast Switching

Fast switching will reduce the sensitivity of observations due to time lost observing the calibrator and moving the antennas, and due to decorrelation from residual phase errors. Both effects can be reduced by observing in the best conditions, which often result in very low residual phase errors at a minimum expense in time lost to the calibration process. However, not all projects can be observed during the best phase conditions. ALMA Memo 174 concludes that fast switching will generally result in less than a 20% decrease in sensitivity for the phase conditions at the Chajnantor site.

3.1.6.3.1.6 Interaction with Scheduling

During poor phase stability conditions, fast switching won't work at the high frequencies. Also, a given target field may have a dearth of calibrator sources, requiring that the field be observed during better phase conditions than the average field. For reasons like these, dynamical scheduling is absolutely required to optimize the utility of the ALMA. We envision one or more phase stability monitors providing real time information to the array control center, and contributing to observing decisions - *e.g.*:

- what project should run on the telescope?
- do the present conditions permit the current project to continue?
- what is the optimal calibrator for the current project in the current atmospheric conditions and hour angle?

3.1.6.3.1.7 Calibrator Survey and Maintenance of a Calibrator Database

The quasars which will form the bulk of the fast switching calibrators will be highly variable at millimeter wavelengths, and a quick survey of a few square degree region about the target source will sometimes be required. The ALMA has the sensitivity to perform a blind search for calibration sources in a few minutes. Surveys directed with lower frequency source catalogs will be even faster. Whenever a potential calibrator is observed, the source information will need to go into a comprehensive calibrator database, which can also be used for choosing an appropriate calibrator.

3.1.6.3.2 Radiometric Phase Correction

The most promising alternative to fast switching is radiometric phase correction (ALMA Memo 209, ALMA Memo 210: 'Radiometric Correction white paper', Weidner 1998 Ph. D. Thesis, Woody and Marvel 1998, ALMA Memo 252). Radiometric phase correction utilizes the variable emission caused by inhomogeneously distributed atmospheric water vapor to determine the phase fluctuations caused by water vapor. While water vapor is not the only source of phase errors, it is the dominant source of short time scale phase fluctuations. This method has had several early successes, but the correlation between the radiometric fluctuations and the interferometrically measured phase fluctuations changes with time, and there are some times when the method does not work well at all.

The current plan for radiometric phase correction is that the 183 GHz water vapor line will be monitored on 1 s or better intervals through a water vapor radiometry (WVR) system. The partial saturation of this line, even in the driest conditions on Chajnantor, initially seemed problematic, but Lay (Memo 209) indicates the unique line shape helps to discriminate between water vapor and errors like spillover, water droplets, temperature fluctuations, height fluctuations, and gain fluctuations. A total of 16 channels each of 500 MHz bandwidth would permit good discrimination between the water vapor and these errors. The ALMA Science Advisory Committee has noted that the benefits of a cooled system in terms of stability and noise probably outweigh the costs. Stable, sensitive WVRs may also contribute to amplitude calibration via measurements of emission and absorption (e.g. ALMA Memo 300). When the PWV column w is under 4 mm, residual antenna based rms path errors of under $10(1+w)$ microns should be achievable on 1 s timescale over a period of 5 minutes and elevation change of 1 degree. Larger water vapor columns preclude high frequency observations, so the larger phase errors associated with high opacity conditions may not be critical.

In the ALMA Design and Development (D&D) Phase there exists an instrument present at Chajnantor to investigate but not implement radiometric correction at 183 Ghz (ALMA Memo 271). At existing sites this line will be saturated much of the time (e.g. Memos 237 and 238 estimate this for Kitt Peak and the VLA site). Hence MDC partners OVRO and BIMA will build and demonstrate 22 GHz radiometric phase correction systems. This will include construction and deployment of hardware and development of algorithms for application of the correction to astronomical data. The CSO/JCMT interferometer operated a 183 GHz phase correction radiometer at Mauna Kea (Memo 252). ESO has duplicated this system at Chajnantor for operation with the 12 GHz interferometers at the site. Reduction of data from this system will help the project to decide how to implement the 183 GHz water vapor spectrometer on theALMA: do we use a standalone cooled or uncooled system, a dedicated radiometer in the receiver dewar, or do we simply use the 183 GHz astronomical receiver as a water vapor spectrometer? See Memo 271 for a report on this system, and Appendix E of the Report of the ALMA Scientific Advisory Committee March 2000 Meeting.

3.1.6.3.4 Calibration of the Electronic and Antenna Phase with an Injected Signal

Radiometric phase correction will only correct for those phase fluctuations which are caused by water vapor, and will not correct for any phase errors caused by variations in the dry atmospheric delay, mechanical instabilities in the antenna, or instabilities in the electronics. Therefore, radiometric phase correction requires some supporting observations or calibration technique to remove phase errors caused by these other sources.

It should be possible to periodically inject a stable signal, perhaps derived from the LO, into the feed to calibrate the electronic contributions to the phase errors. If the calibration signal is injected from the subreflector, then this calibration system will also track the most important mechanical phase drifts of the antenna. If the calibration signal is derived from the LO, and the LO itself has phase instabilities, they will either cancel or be doubled, depending upon the relative phase of the LO and the injected signal. In fact, by alternating the relative parity of the injected signal and the LO, we can solve for both phase errors in the LO and in the rest of the electronics and the antenna up to the subreflector. So, between a reference signal injected at the subreflector and radiometric phase correction, only fluctuations in the dry atmosphere will be unaccounted for.

The on-line system would need to control the details of the injected signal. Information about the injected signal would need to be recorded with the data, and an option for determining and correcting for the electronic phase errors in real time should exist.

The injected signal calibration scheme is an area of research for the design and development phase of the project and will be developed in coordination with the LO system.

3.1.6.3.5 Paired Array Phase Correction

It is possible to use some of the antennas to observe a calibrator and the rest of the antennas to observe the target source. At this time, no special plans are being made for this "paired array" phase calibration technique. Specifically, the array is not being designed in a way that closely pairs antennas to optimize paired array calibration. In the smaller arrays, the configurations will naturally permit paired array calibration. See ALMA Memo 262 for results of VLA tests.

3.1.7.0 Bandpass Calibration

In rough terms, the dynamic range of a single spectral channel which is limited by errors in continuum subtraction caused by bandpass errors will be

$DR=(S_{line}\ \mathrm{sqrt}(N))/(S_{cont} s_{BP})$

where S_{line} is the strength of the line, N is the number of antennas, S_{cont} is the continuum strength of the target source, and s_{BP} is the rms error in the bandpass. For the spectral line observations to be limited by thermal noise and not by bandpass errors, and assuming the bandpass errors are themselves due to thermal noise in the observations of the bandpass calibrator, we have the condition that $N\ \mathrm{sqrt}(t_{cal})/\mathrm{sqrt}(t_{line}) > S_{cont}/S_{cal}$ where S_{cal} is the flux of the bandpass calibrator.

3.1.7.1 Scientific Requirements

A majority of the spectral line observations made with theALMA will probably have no problem meeting the condition of being limited by thermal noise before they are limited by bandpass calibration. Observations of weak spectral lines in the host galaxy of a bright flat spectrum quasar will be about as demanding on the bandpass calibration as these experiments currently are for the VLA or the AT. However, observations of spectral lines on planets in the solar system will be extremely demanding.

For example, the continuum brightness temperature of a planet might be 200 K, and the thermal noise for a bandwidth of 20 m/s and a 1 hour observation at 200 GHz with a T_{sys} = 40 K would be about 0.006 K, requiring a lot of time on a very bright source free of spectral emission to provide a good bandpass calibration. Furthermore, a large fraction of the ALMA's targets will be galactic, and it will be difficult to find bandpass calibrators which are not affected by galactic emission or absorption given the large bandwidth of ALMA.

3.1.7.3 Bandpass Calibration Strategy

We will always have the capability of calibrating the bandpass in the current manner, by observing a strong source for a sufficiently long time. However, injection of a strong noise signal which is flat over the observed frequency range would be an ideal solution to both the planet problem and to the galactic confusing line problem. The injection can be made directly into the feed, or can be part of the ancillary LO system which has been proposed to be injected from the subreflector. The injected noise will need to be much stronger than the 400 K signal at a few percent coupling mentioned in the flux calibration section.

At millimeter and submillimeter wavelengths, the atmosphere will also contribute to the bandpass for wide bandwidth observations, so we must either perform an independent determination of the bandpass astronomically to solve for the atmospheric bandpass component, or we would measure the precipitable water vapor from opacity measurements made at a fiducial frequency and determine the atmospheric contribution to the bandpass through the use of an atmospheric transmission model. Currently, the atmospheric transmission models are probably not good enough for this sort of work, but the ALMA would provide enough data for an ad hoc model or to improve the theoretical models. We will be concerned with changes in the atmospheric component of the bandpass on reasonably short time scales and among the different antennas.

There is an implicit specification placed on the system design that the electronic bandpass be either stable or that it vary linearly with time to something like 10000:1 to 100000:1. If the bandpass changes are mainly linear, we can remove them through interpolation if we calibrate often enough.

3.1.8.0 Polarization Calibration

Linearly polarized feeds have a wider usable bandwidth than circularly polarized feeds, and the ALMA will most likely use linear polarization in order to get complete coverage of all millimeter wavelength atmospheric windows with a reasonable number of receivers. Sault et al. (1991) and Cotton (1998, Memo 208) have both treated the problem of polarization calibration with linear feeds in detail (see also Appendix C to the Report of the ALMA Science Advisory Committee March 2000 Meeting. The main details that we must be concerned with here are that linear polarization leaks into the "parallel hand" visibilities, and that it is not easy to distinguish circular polarization from the instrumental polarization

terms.

Because the linear polarization is entangled with the total intensity in the parallel hand visibilities, there are times when all four cross correlations per baseline will need to be performed, which will probably result in halving the bandwidth and cutting the sensitivity by root two. We consider several cases which could come up with the ALMA to demonstrate when we may need to consider all four cross correlations and when we may use approximations to make use of just the two parallel hand cross correlations:

- Amplitude calibration is performed instrumentally and phase calibration is performed on a quasar (or a combination of radiometric plus a quasar). The quasars will generally be a few percent linearly polarized, but may be as much as 10-20% polarized, and hence Stokes Q and U will influence the parallel hand visibilities. These sources have almost no circular polarization. For a point source, the calibrator's linear polarization will not affect the phase, only the amplitude. If the amplitude calibration is performed instrumentally, as in the scheme of Bock et al., there is no problem with a polarized calibrator and linear feeds. We further consider two subcases:
  - Total intensity imaging with no polarization in the target source. Many millimeter spectral line sources will have little or no linear polarization. Nothing special needs to take place, as the parallel hands will basically contain Stokes I.
  - Total intensity imaging with appreciable linear polarization in the target source. The linear polarization in the target source will corrupt the parallel hand visibilities in a systematic way. However, when the XX and YY visibilities are added together, the linear polarization corruptions cancel out. This is acceptable for low to moderate dynamic range total intensity observations, but may not be sufficient for high dynamic range total intensity observations, as residual gain errors will limit the cancellation of the linear polarization and adding the XX and YY correlations results in a condition in which gain errors no longer close, limiting the use of self-calibration. High dynamic range total intensity imaging of a source with appreciable linear polarization may require full polarization calibration and imaging.
  - Polarization imaging. A bright calibration source must be observed to determine the instrumental polarization leakage or "D" terms. If the calibrator has known (or zero) linear polarization and no circular polarization, the D terms can be determined in a single snapshot. If the calibrator has unknown linear polarization, the calibrator must be observed through sufficient parallactic angle coverage to permit separation of the calibrator and the D terms. Application of the D terms will permit the polarization imaging.
- Amplitude calibration is performed astronomically. If the amplitude calibrator is not polarized, there is no problem. If it is linearly polarized, then the parallel hand visibilities will vary systematically with parallactic angle, the XX and YY visibilities varying in opposite senses. There are several options:
  - For total intensity observations of a target source at low to moderate SNR, the array-wide XX and YY gain ratios can be determined and corrected for.
  - High SNR total intensity observations will require accounting for the different parallactic angles of each antenna, which will result in imperfect cancellation when using the array-wide gain ratios. In this case, the full polarization calibration will need to be performed, even if there is no interest in polarization.

In all cases in which the cross hand visibilities are explicitly used, the X-Y phase offset must be monitored for each antenna. As there is no simple way to determine the X-Y phase offset astronomically,

the ALMA could inject a tone into the feeds, as the AT does. Cotton (1998) points out that it is difficult to generate a millimeter RF tone, and that injecting an IF tone further downstream in the electronics is simpler, though not as good instrumentally. On the other hand, we could derive an RF signal from the LO and inject it into the feeds for the X-Y phase calibration.

The choice of a flux calibrator may also interact with the polarization calibration. Unresolved asteroids which are not azimuthally symmetric will have some time dependent linear polarization, which will complicate the flux calibration. If stars are used for a flux standard, they may display some circular polarization, which would require that another source be used for the D term calibration.

As stated above, the full polarization calibration requires good coverage in parallactic angle to separate the constant instrumental polarization (D term) signal from the sinusoidally varying astronomical polarization signal. This causes some concern since the ALMA is envisioned to be predominantly a near-transit instrument with real time imaging capability. If instrumental polarization calibration is required for many observations, it may be prudent to keep a database of the instrumental polarization solutions at the various frequencies and bandwidths and rely upon that whenever possible. Unlike the VLA, the ATNF compact array shows essentially no time variability in the instrumental polarization (less than 1:10000 over 12 hours, with variations of 0.1% over months). Given the constraints of the ALMA, time constant instrumental polarization may be a good design goal for the feeds, but not a strict requirement.

One way around the complication of good parallactic angle coverage is to use sources of known polarization (ie, unpolarized sources). Holdaway, Carilli, and Owen (1992, VLA Scientific Memo 163) have demonstrated that it is possible to solve for the instrumental polarization for a single snapshot, (ie, a single parallactic angle) if the source polarization is known in advance. So, it would be beneficial to ALMA observing to identify bright, compact sources with known polarization or no polarization for use as polarization calibrators.

3.1.9.0 Special Single Dish Calibration Issues

The ALMA differs from any other aperture synthesis array in that, from the outset, the instrument will support no-compromise single-dish observing modes in addition to the more usual interferometric modes. Some of the issues are discussed in ALMA Memo 108 (*"Single Dish Observing and Calibration Modes"*, D.T. Emerson, P.R. Jewell). Receiver stability and other issues are addressed in ALMA Memo 289, and in Appendix D of the Report of the ALMA Scientific Advisory Committee March 2000 Meeting.

Because single-dish observing is in total power, albeit it switched against, for example, blank sky, there are extraordinary demands on instrumental gain stability. In addition, the extra, variable emission from the sky comes in directly, and tends to mask the much weaker (by perhaps 4 orders of magnitude) astronomical emission. This is in contrast to interferometry, which of course by the use of cross-correlation rather than self-correlation, is relatively immune to these factors.

Astronomical sensitivity calibration in single-dish mode has to be on a dish-by-dish basis; calibration sources need to be detectable with adequate signal-to-noise ratio by one single dish of the array. This is again in contrast to interferometric astronomical calibration measurements, in which the large collecting area of the entire array can contribute to the signal-to-noise ratio achievable in calibrating individual

dishes of the array.

Polarization calibration of single dish observations has its own problems. At mm-waves, polarization measurements are conventionally made with a "widget" in front of the receiver feed. This "widget" introduces changes in the polarization response of the receiver - for example a rotating grid and screen combination can continuously rotate the incident plane of linear polarization. The astronomical polarization is then detected by synchronous changes in total power intensity through the receiver as the sense of polarization changes.

The ALMA may indeed have to provide such "widgets" for each of the antennas. However, the complexity and potential unreliability of such a device could be avoided if it were shown possible to measure polarization reliably, in single-dish mode, by cross-correlation of the signal from orthogonally polarized feeds. Tests of the feasibility of this techniques are planned by early 2000.

3.1.9.1 Atmospheric Emission Cancellation

The emission from the atmosphere is much stronger than the emission from most astronomical sources, and, even worse, the atmospheric emission is variable as well. The variable part of the emission is mainly due to inhomogeneously distributed water vapor, which also causes the phase fluctuations. Since we have excellent statistics of the phase stability on the Chajnantor site, we can infer the severity of the variable atmospheric emission at any desired frequency by using a transmission model or FTS measurements.

For an interferometer, the atmospheric emission above two different antennas is not correlated, so it does not affect the visibilities. In total power continuum observations, the variable atmospheric emission is a major problem which requires some sort of switching on the sky. The total power spectral line case is much less demanding, as large atmospheric fluctuations can be tolerated, considering the much smaller channel widths and much higher thermal noise and the possibility of fitting an average baseline to each spectrum. The spectral line data will have secondary effects, such as the bandpass changing in response to the changing atmospheric load. However, the spectral line observations are much easier than the continuum case, so if we can beat the atmosphere for continuum observations, the spectral line observations will be no problem. The detailed treatment of this problem is presented in an upcoming ALMA Memo (Holdaway, Lugten, and Freund, 2000). ALMA Memo 300 presents a novel method for calibration to remove atmospheric emission which makes use of the very wide bandwidths of ALMA.

3.1.9.1.1 Beam Switching

Traditionally, beam switching by nutating subreflector has been used to remove the variable atmospheric emission. Our study indicates that most beam switching is non-optimal. For any given observation, we would like to be roughly equally limited by thermal SNR and by the residual variable atmospheric emission. If the noise is dominated by the variable atmospheric emission, we need to switch faster. The faster we switch, the better the atmospheric cancellation, but the lower the duty cycle, so the thermal noise will increase. Furthermore, the distance of the throw also needs to be considered. In general, it is optimal to have the smallest throw which gets completely off source. However, in an unstable atmosphere, multiple short throws are better. Hence, the detailed use of a nutating subreflector needs to be fine tuned to match the atmospheric conditions and the observing frequency. As with fast switching, we hope that the observer does not have to perform the calculations to find the optimal switching strategy; the observer should provide high level guidelines, and the program which performs the micro-scheduling should calculate the optimal switching strategy for the current atmospheric conditions.

The nutator design (see specifications in paragraph 4.2.9 of the Antenna chapter of this ALMA Construction Project Book) for the ALMA prototype antenna allows for a maximum throw of about 1.5 arcmin for symmetric beam throwing. Maximum nutating frequencies of about 10 Hz are planned. If it is affordable, nutators with higher peak acceleration and larger maximum throws would be desirable for the production antennas. The two beams should be as similar as possible to reduce the level of systematic errors in beam switching.

The analysis of the On-The-Fly technique for total power continuum observations indicates that it will be as good or better than beam switching in all situations. However, there is considerable risk involved in relying on the On-The-Fly method to cancel all atmospheric fluctuations. For this reason, it is generally agreed that the prototype antennas need to have nutating subreflectors. Currently, this is planned according to the Payne design referred to above.

3.1.9.1.2 On-The-Fly

In On-The-Fly (OTF) observing, the antennas scan quickly across a source at constant elevation angle, using the off-source regions on other side of the source region to define the sky emission. Very large sources will need to be pieced together at some SNR expense. The OTF technique promises to be quite effective at removing the atmospheric emission for three reasons:

- each Nyquist sample on the sky is observed for a very short time, so the system noise is large and a larger amount of sky fluctuation noise is tolerable. (The large number of Nyquist samples observed in each scan compensates for this large noise per Nyquist sample.)
- since more time is spent observing the OFF than an individual ON Nyquist sample, the atmosphere is well determined, unlike beam switching where we are differencing two noisy numbers.
- since the OFF's are observed over a range of time, we can remove a second order polynomial trend in the atmospheric emission time series, which greatly reduces the residual sky emission fluctuations.

For sources which are about one beam across, the OTF observing strategy works about as well as beam switching. For larger sources, OTF wins because of the relative increase in the SNR of the atmospheric determination and because multiple throws begin to degrade the beam switching SNR.

Because the entire antenna is moving, many systematic errors which plague beam switching (such as differences in the shapes and gains of the ON and OFF beams) are eliminated. However, it takes much more energy to move the entire antenna, and there is more risk in general with an observing strategy that attempts to move the entire antenna.

3.1.9.1.2.1 Controlling Antenna Movements

OTF will work only if we can slew and reaccelerate the antenna quickly without exciting the lowest resonant frequency of the antennas. An initial analysis of this problem has been performed by Holdaway, Lugten, and Freund (2000). Using a Gaussian acceleration profile and an error function velocity profile, they predict the antennas will be able to turn around from one scan direction to the other in about 0.2 s without appreciably exciting the lowest resonant frequency. This acceleration profile is a good one, but probably not an optimal one, so further work could help optimize the profiles for both OTF antenna motion and fast switching antenna motion.

In order not to excite the antenna motions, the acceleration must be very smoothly varying. This will put

strong constraints on both the control system and on the servo system.

3.1.9.1.2.2 Maximum Velocity and Acceleration

OTF simulations of sources of various sizes indicate that the optimal slew velocity varies linearly with source size. For a maximum interesting source size of 1 deg, a maximum slew rate of about 0.5 deg/s is required. This requires a maximum antenna angular acceleration of about 12 deg/s/s. Since the profile is Gaussian, we do not require this maximum acceleration for very long. These maximum velocities and accelerations are for an antenna with lowest resonant frequency of 6 Hz. An antenna which was less stiff could not utilize such large accelerations and velocities in OTF observing. A stiffer antenna would permit faster turnarounds, requiring larger accelerations and velocities. However, the 6 Hz antenna is effectively beating the atmosphere already, so not much is gained from a stiffer antenna.

3.1.9.1.2.3 Reading Out Encoders, Dump Time

OTF requires that we know where the antenna is for each Nyquist beam. At the 0.5 deg/s maximum slew rate, observing at 850 GHz with a half beam size of 0.001 deg will require that we dump the data and know where the antenna is every 2 ms. We don't need to make the antenna go to any precise place at any precise time, we just need to know where the antenna was at a precise time. We may not need to read the encoders every 2 ms; if the antenna position changes smoothly over time scales of 10 ms, we can read out the encoders more coarsely and interpolate. We do not require that the encoders be accurate to within the pointing specification of 0.6 arcsec.

3.1.9.1.2.4 Antenna Motions Don't Need to be Synchronized

Since we are only talking about total power OTF here, we need not synchronize all the antennas in their dance across the sky. The antennas could be staggered to permit a more constant utilization of electrical power.

3.1.9.1.2.5 1/f Noise

In addition to atmospheric brightness fluctuations, beam switching and OTF will remove a portion of the receivers' 1/f noise. From the optimizations we have performed, we can set specifications on the 1/f noise for each observing frequency. Even though the beam switching is performing the switching faster than OTF, the integration time spent on each ON is often larger than the integration time spent per Nyquist sample of an OTF observation, so OTF and beam switching are similar in their ability to switch out 1/f noise. If these specifications cannot be met, we must reoptimize the OTF observing strategy, which would result in moving more quickly to accomplish faster switching and less time or more white noise per Nyquist sample on the source. This would favor both higher maximum accelerations and a stiffer antenna.

Freq	Beam Size Source		0.5 deg Source	
[GHz]	noise	break frequency	noise	break frequency
	[Jy]	[Hz]	[Jy]	[Hz]
90	0.047	1.2	0.081	0.34
230	0.088	1.2	0.25	0.29

345	0.14	1.2	0.47	0.29
650	0.33	1.3	1.6	0.34

Table 1: For continuum (8 GHz bandwidth per polarization) OTF observations, what noise level must the 1/f noise be below, and at what frequency, for 1/f noise to have essentially no effect on OTF observations' sensitivity?

3.1.10.0 Solar Calibration

For solar observing, some type of attenuating "widget" in front of the receiver may be required, to reduce the necessary dynamic range of the receiver and backend electronics. Some special calibration scheme needs to be thought out specifically for calibration of solar observing. This is under study.

3.1.11.0 Editing

Both the on-line and post-processing software should provide for carrying various monitor data through the system and allowing easy editing of astronomical data based on the monitor data, on a per time or per antenna basis.

References

Bock, D., J. Welch, M. Flemming, and D. Thornton, 1998, "Radiometer Calibration at the Cassegrain Secondary Mirror" ALMA Memo 225, 1998.

Carilli, C., Lay, O. and Sutton, E., Radiometric Phase Correction White Paper. 1998.

Cornwell, Holdaway, and Uson, , 1993, ``Radio-interferometric imaging of very large objects: implications for array design", A&A; 271, 697-713.

Cotton, 1998, ALMA Memo 208, 1998.

Holdaway 1997, ALMA Memo 169, 1998.

Holdaway, Carilli, and Owen (1992, VLA Scientific Memo 163)

Kutner, M.L., (1978). Ap.Letters,**19**,81.

Lay, O. 1998. ALMA Memo 209, 1998.

Mangum, J. ``Amplitude Calibration at Millimeter and submillimeter Wavelengths" 2000.

Marvel, K. and Woody, D. 1998 BAAS 192, 8103.

Sault, R. J., Killeen, N.E.B., Kesteven, M.J. 1991, ``AT polarisation calibration", ATNF Technical Document Series 39.3015.

Ulich, B.L. & Haas, R.W. (1976). Ap.J.S.,**30**,81.

Wiedner, M. 1998 Ph. D. Thesis, Cambridge University.

Yun, M., Mangum, J., Bastian, T., Holdaway, M. and Welch, J., Amplitude Calibration White Paper 1998.

ALMA Construction Project Book, Chapter 3 Section 2.

Calibration: Hardware Schemes

John Payne
Andrea Vaccari
Darrel Emerson
Jeff Mangum
Last modified 2001-02-02

Revision History
1998-07-16: First version, for original MMA project
1998-11-12: Minor updates
1999-04-13: Major Revision (Photonic Calibration System)
2000-12-13: Minor changes for ALMA context
2001-02-02: Minor changes: added reference to transparent vane option

---

Summary

In this section, hardware solutions to the problem of calibrating the ALMA amplitude and phase are described. Both solutions use the blocked area in the center of the subreflector as the source of radiation from either a two-temperature load or a coherent signal source. A simple mirror mechanism is used to select between the two systems. The coherent source may be made phase stable through a round-trip measurement scheme so raising the possibility of continuous phase measurement and correction as well as providing a valuable trouble shooting tool.

The details are still (February 2001) under discussion between the international partners. This chapter is taken with minor changes from Chapter 3.2 of the original MMA project book. As calibration plans for the international ALMA project become more certain, the content will be updated. Although the 2-temperature load scheme presented here is the baseline plan, alternative chopped-load designs are still actively being considered; for example a semi-transparent vane. If further study indicates this to be an advantageous option, the concept may be tested on the Test Interferometer in 2002-2003.

3.2.1 Introduction

The previous section describes a number of schemes for calibrating the ALMA amplitude and phase. This section outlines two specific hardware schemes which can help to calibrate the instrumental phase and amplitude of ALMA. Other instrumental calibration schemes, such as round-trip phase calibration for the local oscillator, and AGC/total power monitors in the I.F. chain, are described elsewhere.

The two calibration schemes outlined here are:

Absolute amplitude calibration, which represents the conversion of total power signals from the radio astronomy receivers to calibrated antenna temperatures, and **Relative amplitude and phase calibration,** with an artificial coherent calibration signal suitable for interferometric calibration

3.2.2 Absolute Amplitude Calibration

It has been suggested by Jack Welch and others, and is currently the subject of a joint MDC development between BIMA and NRAO. It gives an absolutely calibrated signal of a few K at the receiver, over the complete frequency range covered by the ALMA. This system has been described in MMA Memo 225 by Bock, Welch, Flemming and Thornton.

The essence of this technique is that a black body radiator is placed at the center of the subreflector of each antenna, within the unused area of subreflector matching the central blockage of the antenna. In this way there is no effect on antenna sensitivity.

Within this central part of the subreflector, a plane mirror switches between two (or more) loads of different temperatures. The two loads have very precisely controlled and calibrated temperatures. The total power output of the receiver is sampled synchronously with the mirror switching between the two calibrated loads.

The added switched receiver noise is, to a first approximation, equal to the difference in temperature of the two hot loads, multiplied by the beam solid angle of the absorbers at the subreflector seen from the receiver feed, divided by the beam solid angle of the subreflector - assuming the receiver feed itself is matched to the angle subtended by the receiver. This ratio will be reasonably constant with frequency, but at a given frequency can be calibrated precisely by measurements of the feed antenn a pattern.

For more details see the memo by Bock, Welch, Flemming and Thornton. The joint development with BIMA will show, on a timescale of a few months, how well the technique can be expected to work in practice.

3.2.3 Relative phase and amplitude calibration

In the debugging stage of the ALMA, there will be a need for a generic test signal that can be used to debug the entire electronic system of a given antenna or antenna pair, from front-end to correlator. When the antenna surface and pointing are sufficiently reliable, astronomical sources can be used for this purpose, but having an independent, artificially generated signal that is not dependent on antenna performance will be invaluable in checking out and maintaining the system.

If the calibration signal can be made coherent at all individual antennas, it opens up the possibility of

calibrating the entire receiver system, front-ends, back-ends and correlator, amplitude and phase as a function of frequency, in a way independent of antenna tracking, pointing, or efficiency performance. The calibration system should be sufficiently stable that it can be used as a secondary calibration system, with only occasional cross-calibration with astronomical sources.

3.2.4 The Photonic Calibration System

3.2.4.1 Introduction

The photonic calibration system has a broad-band, radiating antenna situated at the center of the subreflector, where no extra antenna blockage is introduced. At the feed of the broadband antenna, there is an uncooled photomixer device. A single optica l fiber, carrying laser signals generated at a central laboratory or control room, feeds this photomixer. In the simplest form, the optical signal would come from two lasers, whose difference frequency corresponds to the telescope observing frequency, and which is phase-locked to the telescope frequency standard. The equipment required to do this would be nearly identical to that being developed for the photonic laser local oscillator system. Only one pair of lasers would be required for the entire array; the combined laser output would be split optically N ways (where N is the number of antennas) and routed via N independent fibers to each antenna.

In slight variants of this scheme, either a single laser signal, or the dual laser system tuned to the required mm-wave difference frequency, could be modulated. The modulation might take the form of a regular comb spectrum, simulating broadband noise. This becomes quite analogous to the pulse cal system developed for the VLBA, and could be used for checking the relative amplitude and phase response over the entire interferometric IF passband. The modulation might also be a truly random, or a pseudo-ra ndom digitally generated sequence, which would also provide a broad-band coherent test signal. This random or pseudo-random noise needs to be coherent at each antenna, so timing considerations, within a fraction of the reciprocal bandwidth, are important.

Naturally this injected signal needs to be stable, both in amplitude and phase. It may require round-trip delay compensation of some type, and perhaps an AGC system to keep the signal amplitude constant. However, attention to the stability of this cali bration signal may relax the technical requirements elsewhere in the system.

Most of the development for this coherent photonic calibration scheme is already being undertaken in the context of the photonic local oscillator development. The calibration scheme should in principle be much simpler, because several orders of magnitu de lower radiated mm-wave power is required. The main additional development needed is that of the broad-band radiating antenna, to be sited at the subreflector, fed by the signal from the photomixer.

3.2.4.2 Requirements

The first thing we did was evaluate the minimum power required at the receiver. The total noise power (P_N) for a receiver with system temperature (T_{sys}) of 100 K and observation bandwidth ($\Delta\nu$) of 2 GHz is

given by

$$P_N = kT_{sys}\Delta\nu = 2.76\ 10^{-12}\ W = -85.6\ dBm$$

We want the received calibration power to be at least 1% of the total noise power: $2.76\ 10^{-14}$ W = -105.6 dBm.

Another requirement is that the injected signal should be linearly polarized at 45 deg with respect to the OMT or polarization grids inside the receiver so that the power is equally distributed into the two channels.

As this signal should work as a reference we need its characteristics to be as constant as possible in the whole spectral range we are going to use it: 100 GHz - 1 THz.

3.2.4.3 Review of antennas for radiating the signal

The best way to launch the signal towards the receiver is use an antenna placed in the blind spot of the subreflector.

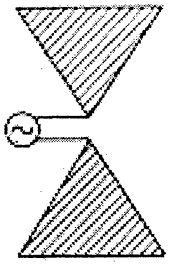
As we look at the requirements, we need an antenna with the following characteristics:

- very broad bandwidth
- linear polarization
- relatively small dimension (the transmitting system should fit in the system sketched in Fig. 4 in MMA Memo 225: *"Radiometer Calibration at the Cassegrain Secondary Mirror."*)
- possibly high directivity (this depends on how much power we can drain from the diode that will drive the antenna)

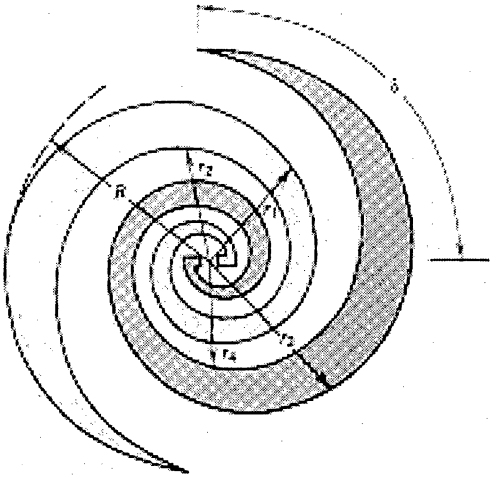
There are few antennas that fit these requirements. We focused our attention on the *self-similar and self-complementary planar antennas.*

The self-similarity is the geometric property of invariance under a uniform expansion or reduction of size. This property implies the absence of any characteristic length scale so that the geometry is entirely defined in term of angles. This guarantees the broad bandwidth.

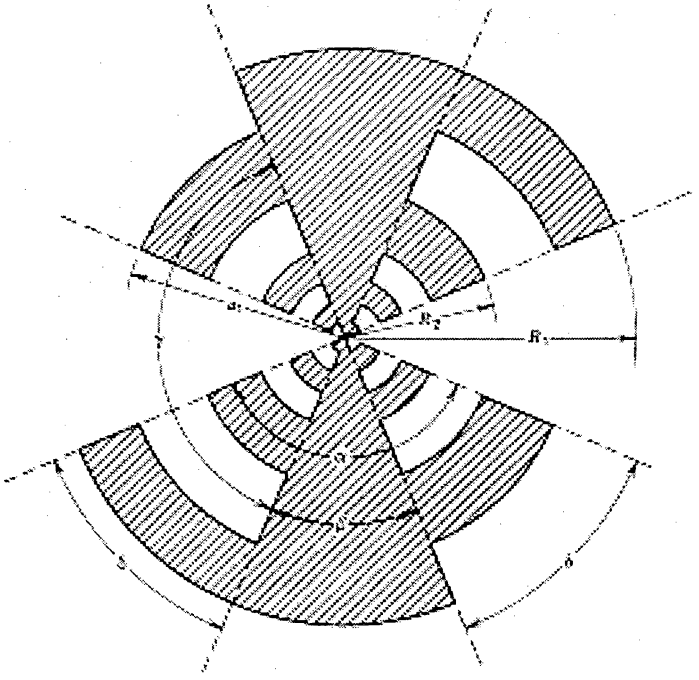
There are several type of self-complementary and self-similar planar antennas: Bow-Tie, Equiangular Spiral and Log-Periodic.



Bow-Tie[4]



Equiangular[4]



Log-Periodic[4]

The Bow-Tie is the simplest, however its pattern is double-peaked off-axis[5] and therefore is not suitable for coupling to gaussian or other commonly encountered beams.

The Equiangular is circularly polarized[1].

The beam pattern, the impedance and the rotation of the linearly polarized emitted signal of the Log-Periodic antenna are exactly periodic with the logarithm of frequency the period being given by the ratio of two successive teeth (R_1/R_2). We decide to study with more attention this antenna because its characteristics are closer to our requirement's than the others'.

The self-complementary is the geometric property of invariance (with a rotation) under an interchange of metallized and non-metallized regions. The Booker's relation in free space for complementary antennas[1] states that the impedances (Z_1, Z_2) of any pair of complementary antennas are related to the free space

impedance ($Z_0 = 120\pi\ \Omega$) by

$$Z_1 Z_2 = (Z_0/2)^2 = (189\ \Omega)^2$$

As a result self-complementary antennas in free space have constant real impedance of 189 Ω at all frequency.

The purpose of our feed antenna is to couple power from a device that is much smaller than a wavelength into a wave in free space. For the frequency range we are interested in, the linear dimensions of these antennas are so small[2] that the b est way to build them is by lithography. This manufacturing process has also the advantage to allow for an easier coupling to the emitting device since usually it is built by the same process and, by slight changes to the design, it is possible to build e mitting device altogether with the antenna.

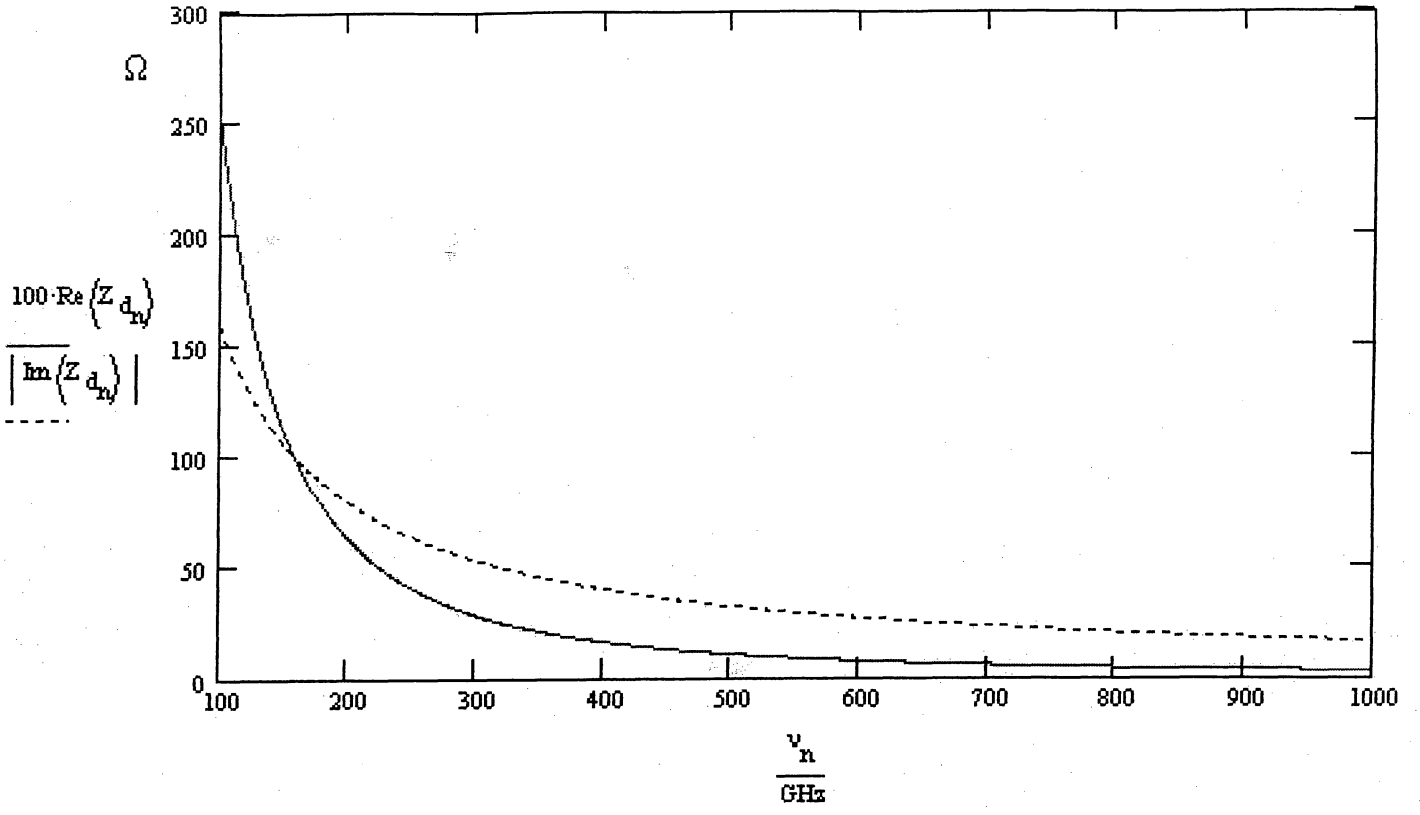
This manufacturing procedure implies a dielectric substrate over which the antenna will be deposited.

An antenna at a dielectric/air interface will preferentially couple into the dielectric substrate. In a first order the phase velocity and impedance are both reduced by[3] $\varepsilon_{eff} = [(\varepsilon + 1)/2]^{1/2}$ it follows that for a self-complementary antenna the impedance will be $Z = Z_0/\varepsilon_{eff}^{1/2}$ (for example 74 Ω for Si o GaAs and 114 W for crystalline quartz)[3].

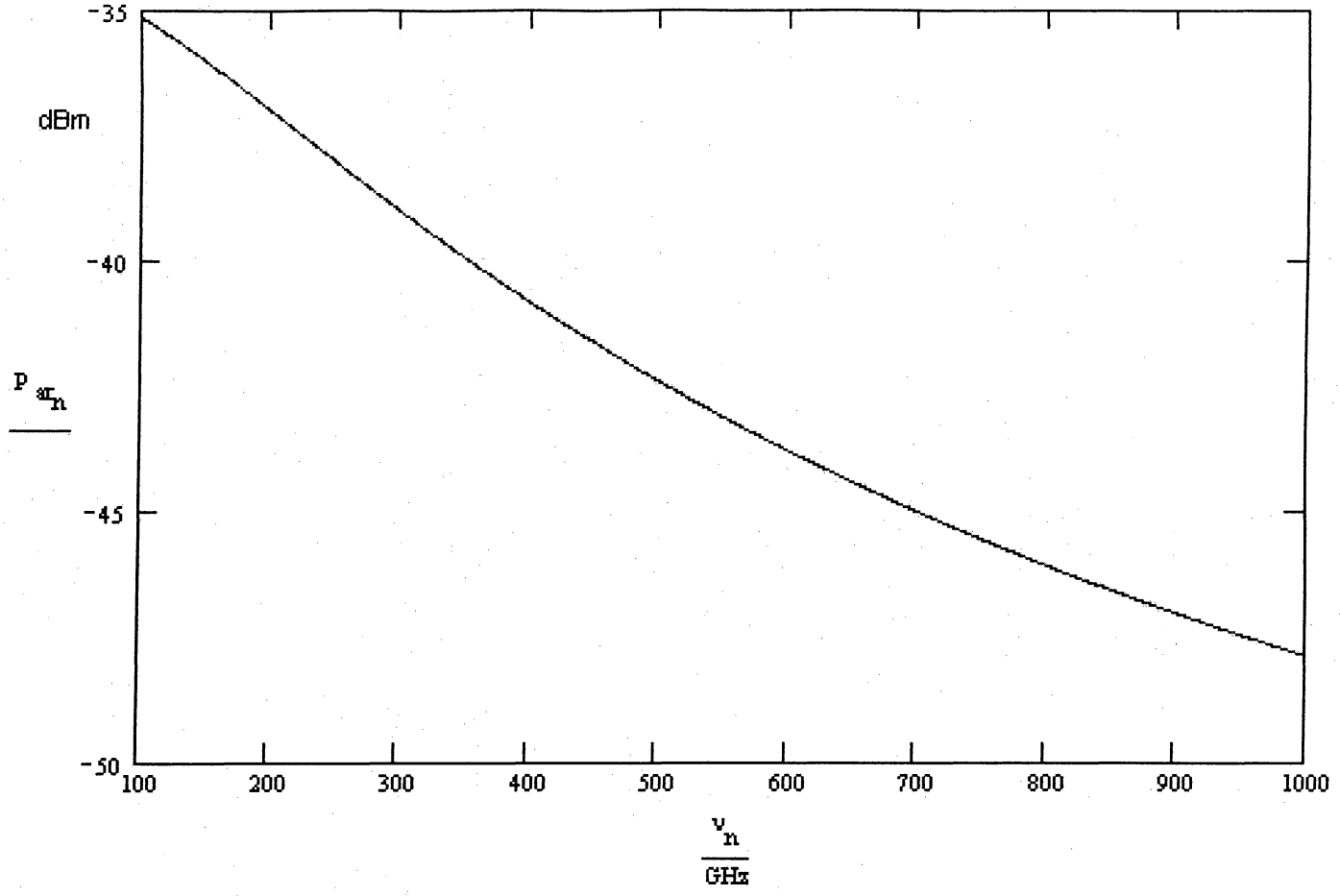
Another problem araising from the coupling in the substrate is that, in case the substrate geometry is simply a plane-parallel slab, any ray radiated from the antenna at angles exceeding the critical angle will be totally internal reflected by at the b ack surface of the substrate. There are two main ways to get rid of this problem. The first is to deposit the antenna on a very thin substrate (for example a free-standing membrane of silicon oxynitride and shape the holder as a horn to use all the power from the antenna). The second solution is to shape the substrate in such a way that it acts like a lens on the emitted wave. There are several shape that can be used to increase the directivity of the beam pattern: Hemisphere, Hyperhemisphere, Cartesian O val, Dielectric Filled Parabola. We are still investigating whether and which of these shaped dielectric substrate is suitable to our purposes.

3.2.4.4 Diode-Antenna matching

These impedance values have been calculated using 10 kΩ for the diode resistance and 10^{-2} pF for the diode capacitance and making the assumption that the diode behaves like an ideal current generator in parallel wi th the resistance and the capacitance. The impedance of the diode should change with the frequency approximately as in the following figure.



The real impedance has been multiplied by a factor 100. Using this evaluation for the impedance values and assuming that the antenna has a real impedance of 74 Ω we found the following results for the diode-antenna coupling.



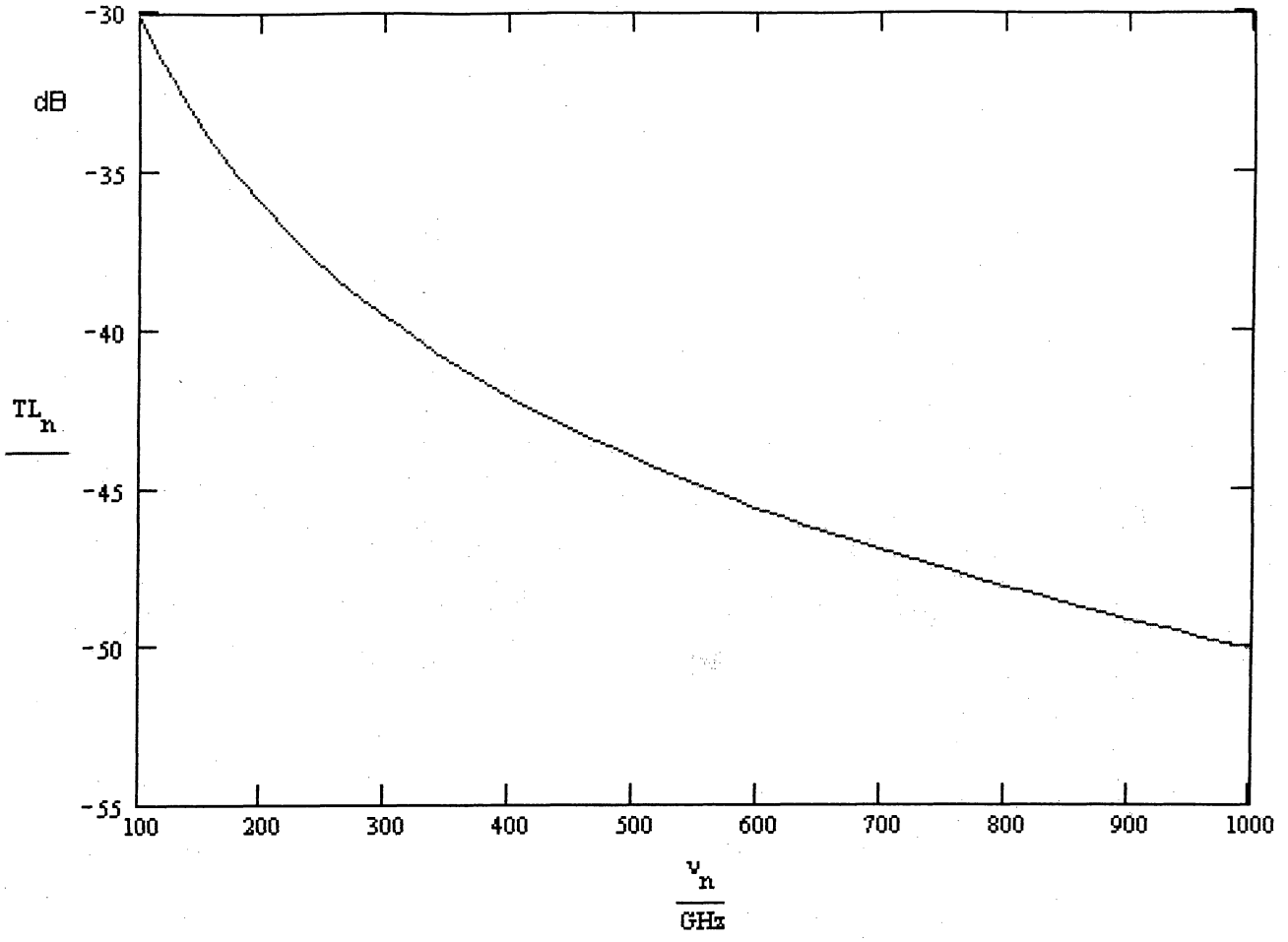
This represents the power delivered by the antenna to free space assuming that the diode is driven with a current of 100 μA.

3.2.4.5 Antenna-Receiver matching

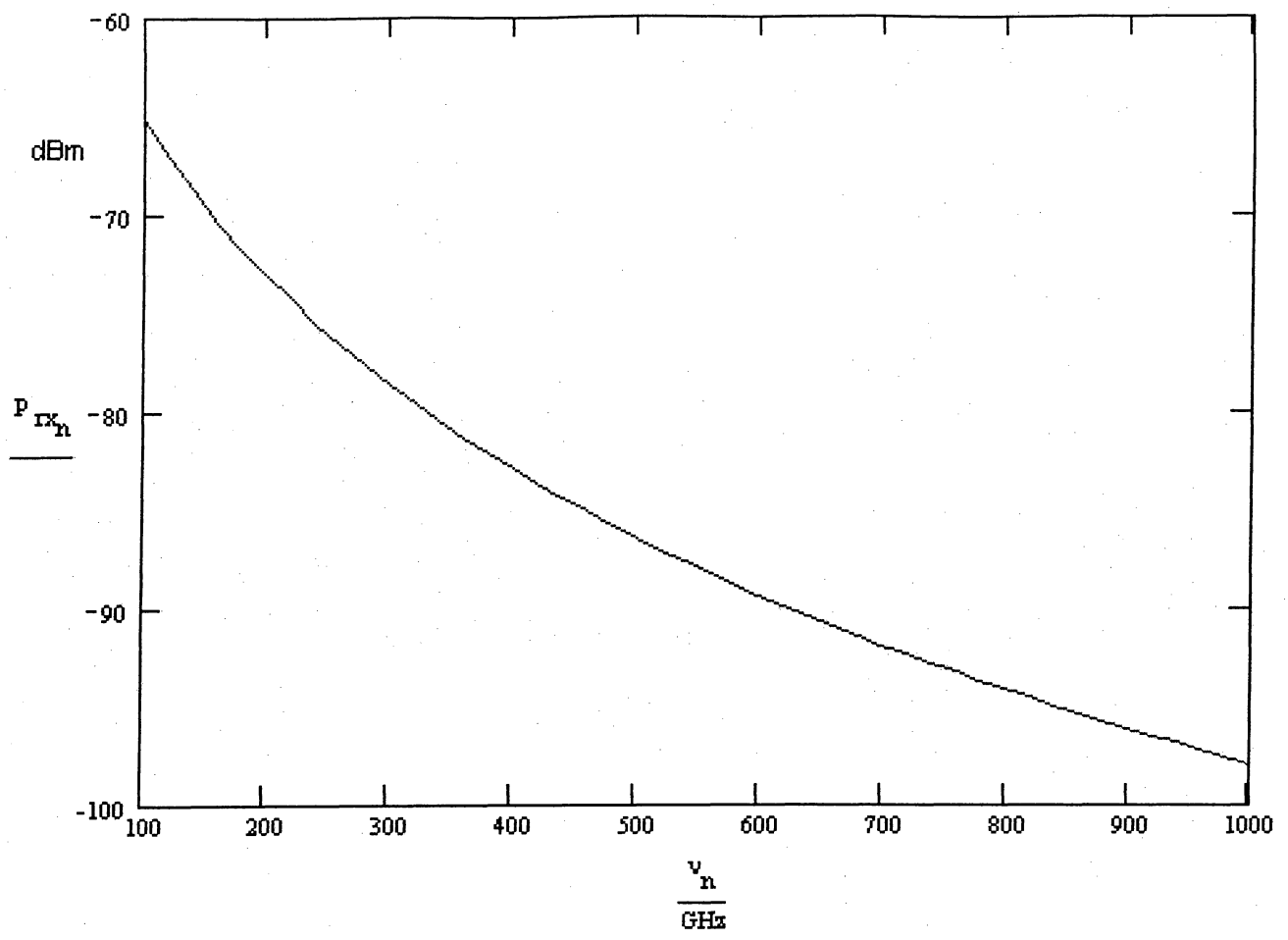
Let us consider now the efficiency of the coupling between the small antenna and the horn of the receiver. During the operation of the telescope several feed-horns will be changed to match the different frequencies so that the subreflector will always be fully illuminated, we can assume the beam pattern to be a gaussian with -11 dB point at the edge of the subreflector (3.58 deg[6]).

Also assume that the beam pattern of the planar antenna is a gaussian with -3 dB point at 20 deg[7]. As we saw we can increase the directivity by shaping the dielectric substrate.

Using these parameters we can evaluate the transmission loss of the antenna-receiver system and obtain the following result



Now if we consider the result achieved in the previous paragraph we can say that the total power available at the receiver is



The minimum power is -105.6 dBm and this result shows that with this configuration we can satisfy the power request in the whole spectral range of interest..

3.2.4.7 References

[1] Victor H. Rumsey, *Frequency Independent Antennas*, Academic Press

[2] D. Chouvaev et al. - "Normal Metal Hot-Electron Microbolometer with Andreev Mirrors for THz Space Applications" - Ninth International Symposium on Space Terahertz Technology

[3] E.N. Grossman - "Lithographic Antennas fo Submillimiter and Infrared Frequencies" - IEEE Int. Symp. on Electromag. Compat. - 14-18 Aug. 1995, pp. 102-107

[4] Warren L. Stutzman, Gary A. Thiele - "Antenna Theory and Design" - Wiley & Sons

[5] G.V. Elftheriades et al - "A 20 dB Quasi-integrated Horn Antenna" - IEEE Micro. and Guided Wave Lett., Vol. 2, pp. 73-75 (1992)

[6] MMA Memo 246: James W. Lamb - "Optimized Optical Layout for MMA 12-m Antennas"

[7] M.A. Tarasov et al. - "Quasioptical Josephson direct detectors for mm-wave spectrum analysis" - Proc. of the Int. Conf. on mm- and submm-waves and Applications (SPIE), 10-12 Jan. 1994, pp. 220-228

3.2.5 Combined Calibration System

The incoherent calibration scheme described earlier switches between blackbody radiators of different temperatures using a mirror. In principle, by allowing an extra position on this mirror, the radiated calibration signal can be switched between the incoherent blackbody loads and the coherent radiator. At a later stage in the development, when the feasibility of both coherent and incoherent calibration schemes has been demonstrated, the combination of the different calibration radiators into one packa ge will receive attention, as will studies of how to achieve the necessary amplitude and phase stability.

3.2.6 Work to be done

The photonic calibration builds on work already in progress in the context of the photonic local oscillator scheme, with the exception of the broadband antenna. Some simple design work is required now (e.g. exactly how much coherent power needs to be radiated from the subreflector, with what requirements on amplitude and phase stability?) but the bulk of the effort can be expected fairly late in the ALMA development phase.

Reference:

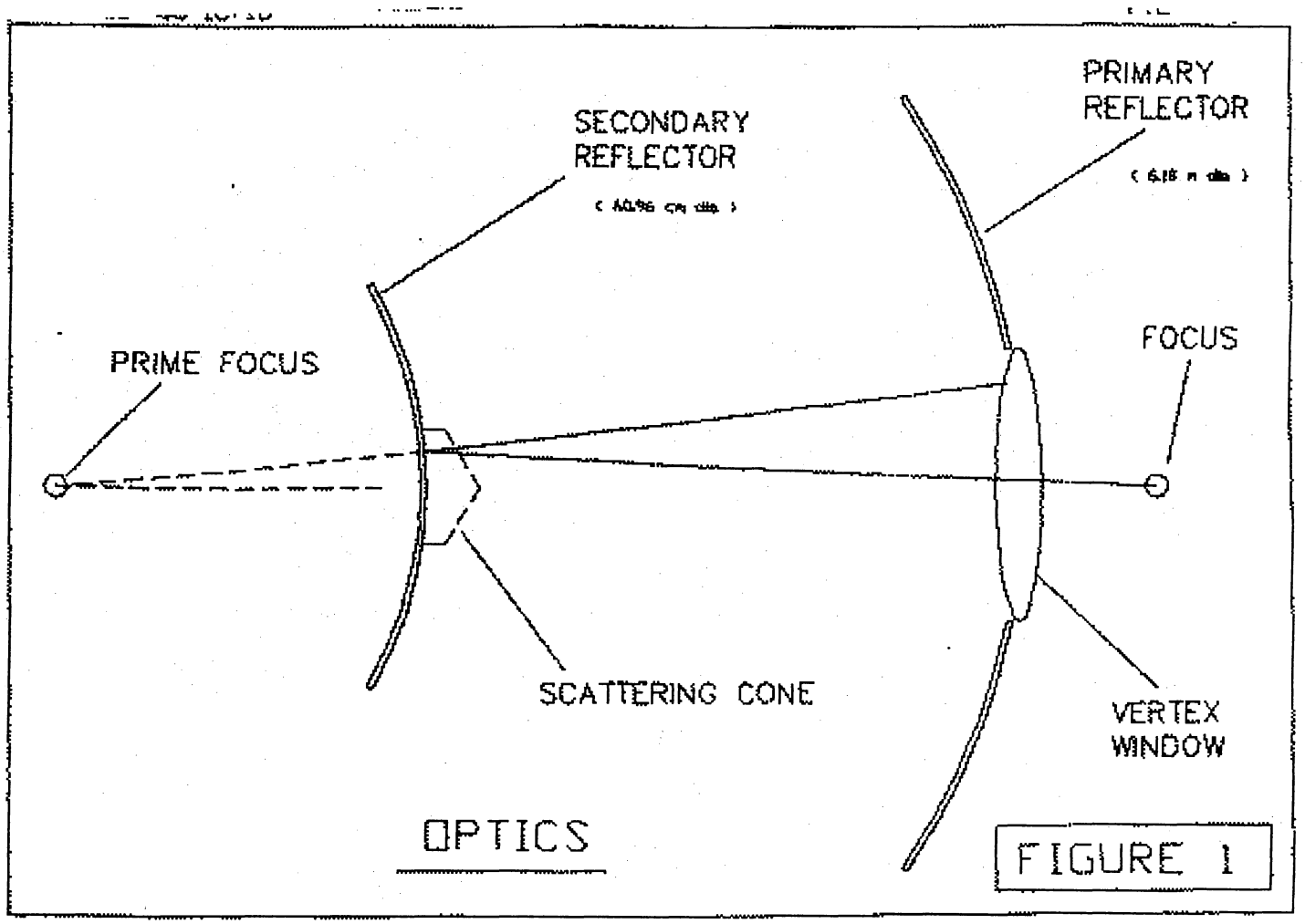
D. Bock, J. Welch, M. Flemming and D. Thornton, MMA Memo 225: *"Radiometer Calibration at the Cassegrain Secondary Mirror."* (See also the Appendix below.)

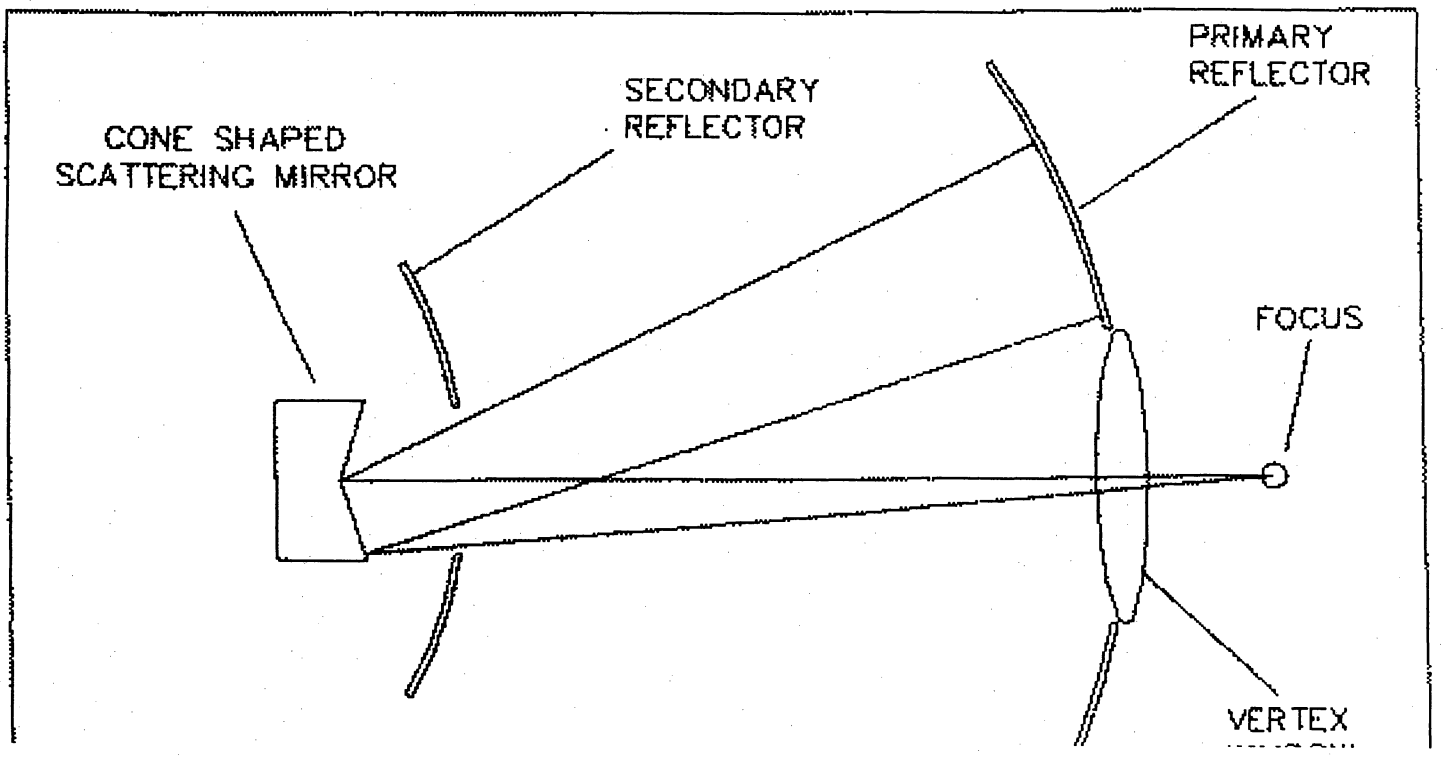
APPENDIX

The following figures are from the MMA Memo 225 by Bock, Welch, Fleming and Thornton, *"Radiometer Calibration at the Cassegrain Secondary Mirror."*

Figure 1 shows the general Cassegrain optics, which normally has a scattering cone covering the central

part of the subreflector to direct unwanted rays on to cold sky. **Figure 2** shows a scattering mirror behind the subreflector, giving mu ch the same effect.





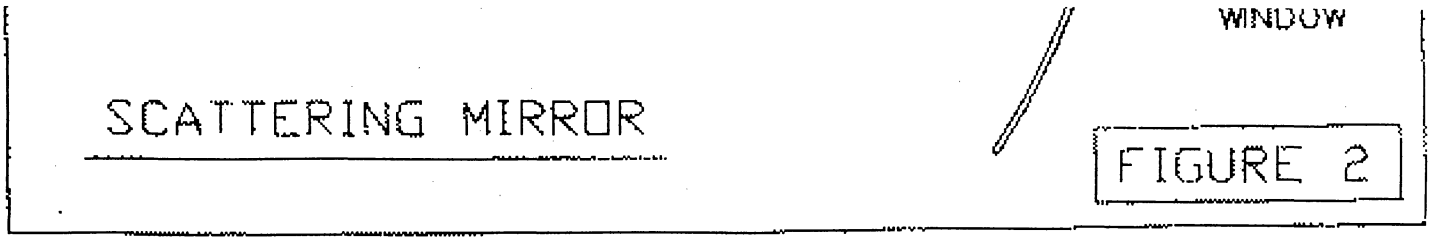
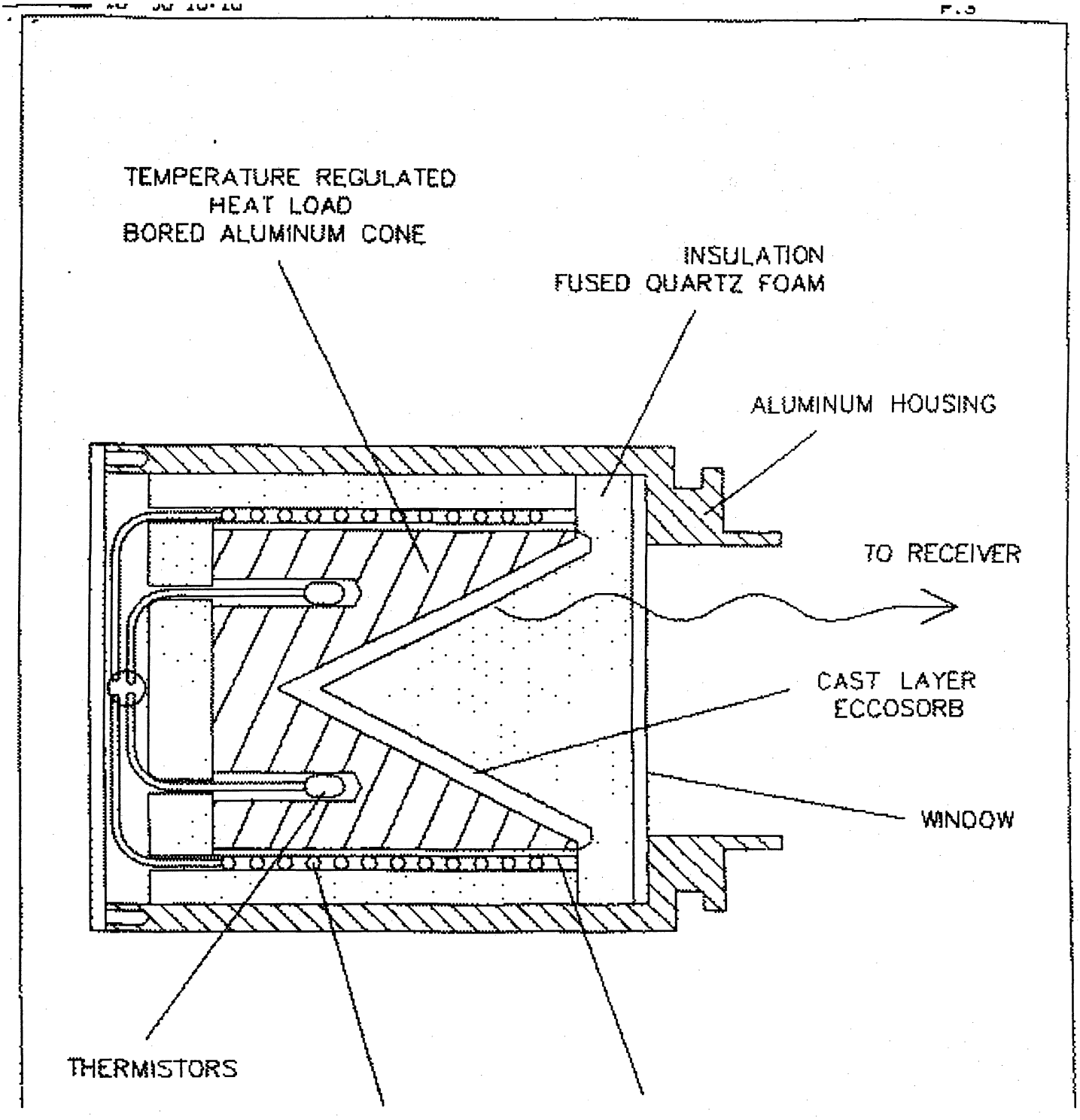


Figure 3 shows the absorbing black body load that would be placed behind the central hole of the subreflector



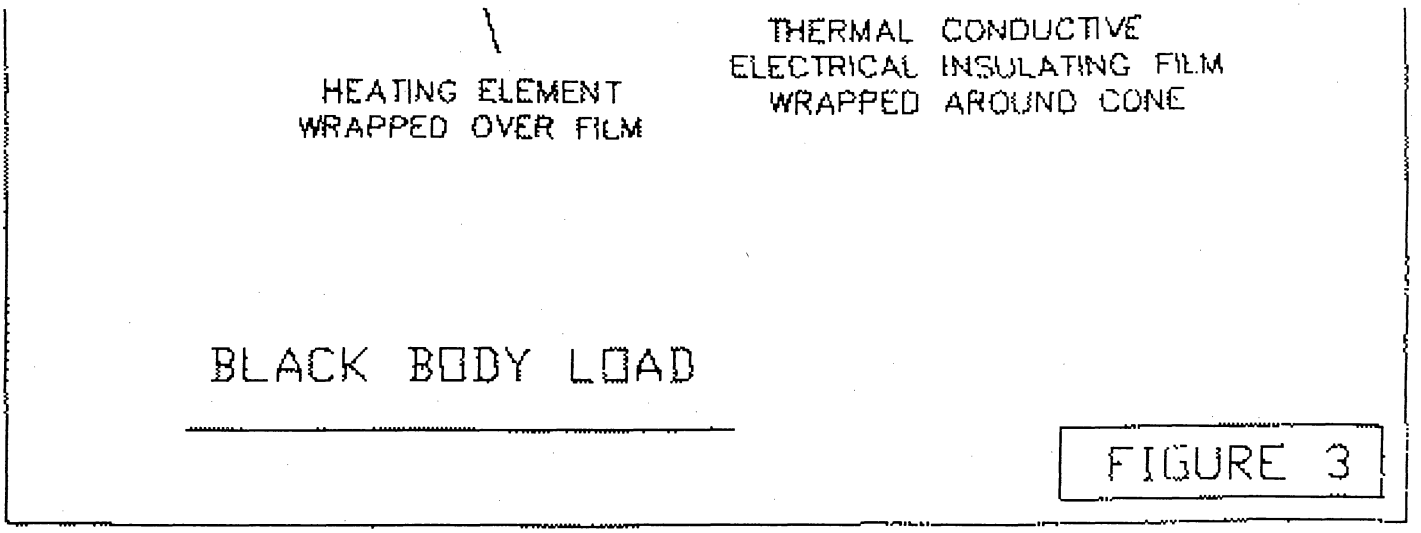
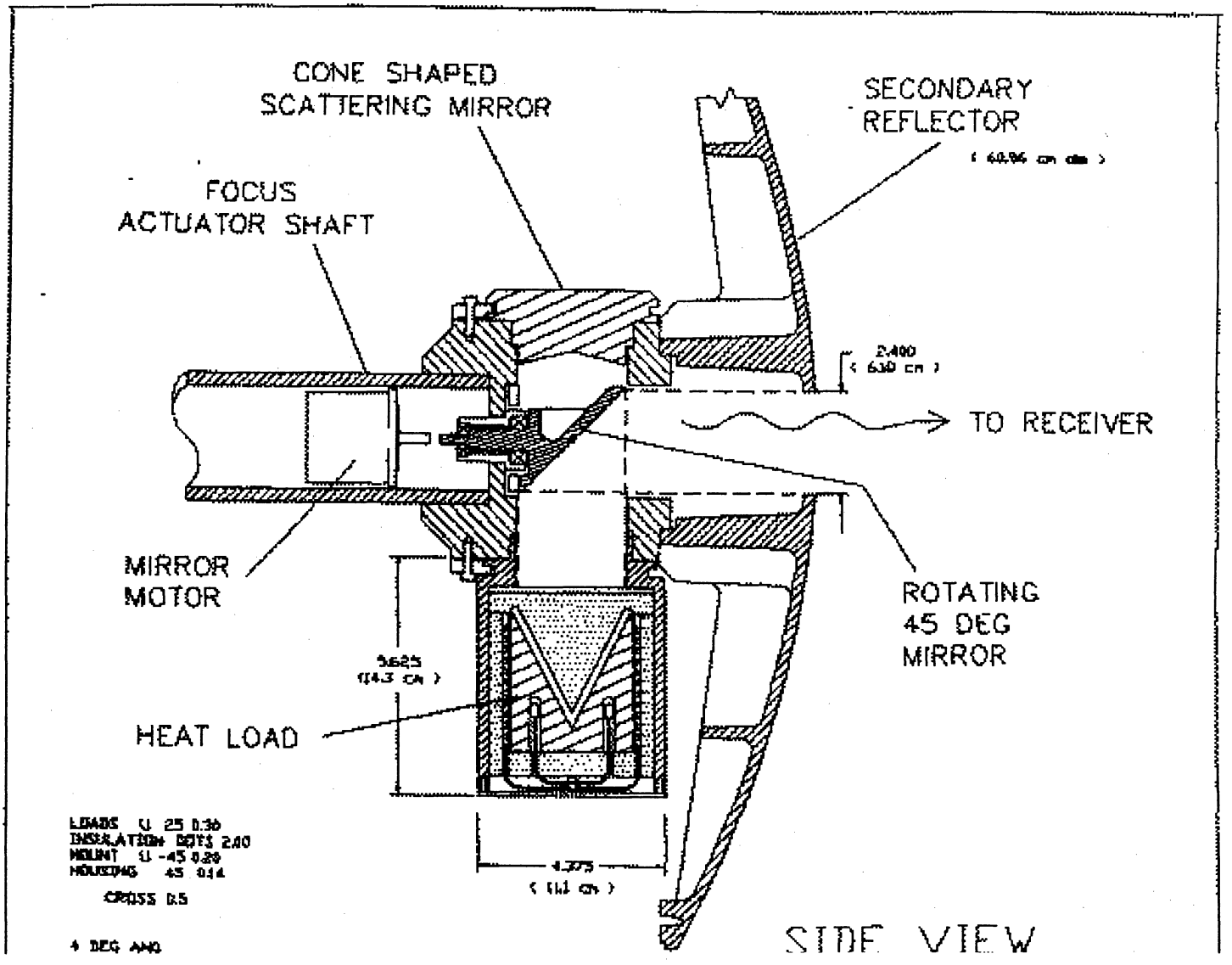
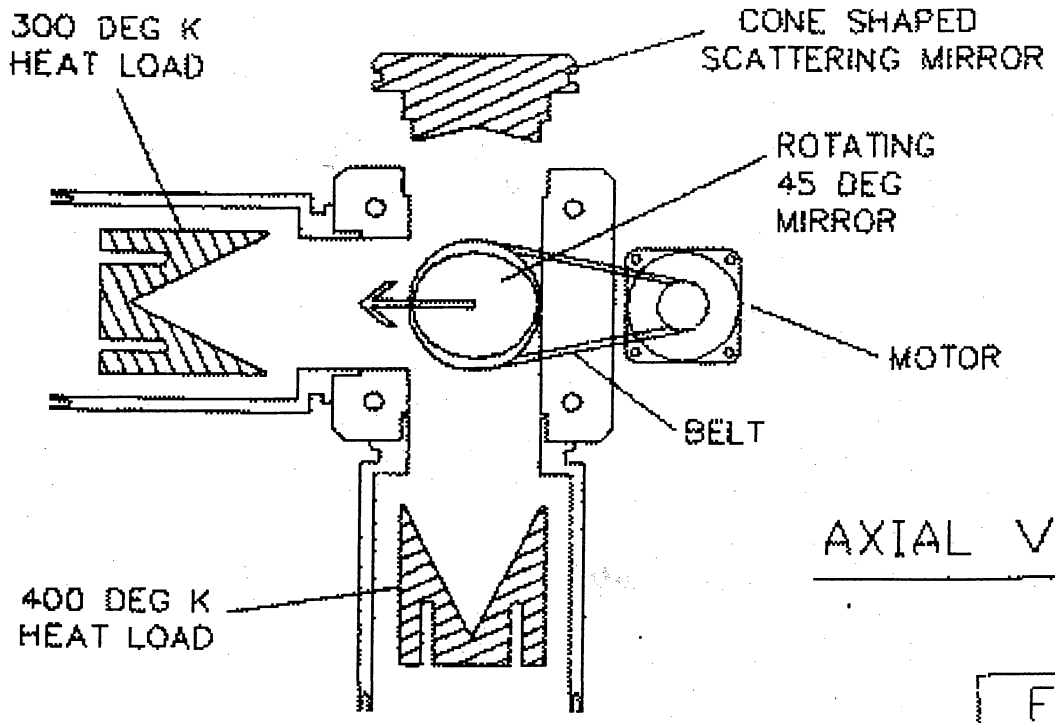


Figure 4 shows the arrangement of a rotating 45-degree mirror which will choose between one of two hot loads, whose temperatures differ by ~100 K, and the scattering cone from which rays which eventually reach cold sky.





AXIAL VIEW

FIGURE 4

ALMA Project Book, Chapter 4.

ANTENNAS

Torben Andersen and Peter Napier
Last changed 2001-Feb-06

Revision History:
1998-Aug-23: Added information from PDR; test plan, earthquake spec, receiver cabin power and mass requirement.
1998-Sep-01: Added Summary, corrected typo in fast switching time spec. Added requirement for metric compatibility, electrical supply voltage and frequency.Added chapter number to section numbers.
1998-Sep-18: Changed resonant frequency requirement in 4.2.8; added 3 phase voltage in 4.2.2; corrected wind speed typo in 4.2.1.3.
1999-Apr-13: Changes resulting from the decision to use 12 m diameter antennas, and updates on some of the specifications as a result of preparing the Request For Proposal. Significant specification changes include bottom elevation limit, survival stow position, operational cycles for fatigue design, optics layout and phase stability.
2000-May-31: Changes resulting from transition from MMA to full ALMA project. Also, most specifications firmed up as a result of placing the EIE and Vertex contracts.
2001-Feb-06: Corrections and updates after review by the ALMA Systems Group.

Summary

The ALMA radiotelescope is currently planned to consist of a goal of 64 antennas, each of 12 m diameter. In this chapter we outline the general requirements for the antennas and the detailed specifications can be found in the contract for the US prototype antenna (NRAO, 2000) and in the US Interface Control Documents (ICD) (ICD, 2000) which are part of the contract. The specifications and ICDs for the European prototype antenna can be found on the ALMA-Europe Antenna WWW Page. The principal requirements for the antennas are shown in Table 4.1.

Table 4.1 ALMA antenna principal performance requirements.

Configuration	Elevation-over-azimuth mount, Cassegrain focus
Frequency range	30 GHz to 950 GHz
Precision performance conditions	Nightime: wind 9 m/s Daytime: wind 6 m/s and sun from any angle
Reflector surface accuracy	20 microns rms, goal; 25 microns rms, spec
Pointing accuracy, rms	0.6 arcsec (offset, 2 deg in position and 15 min time), 2.0 arcsec (absolute)
Fast switching (settle to 3 arcsec pointing)	Move 1.5 deg in position in 1.5 seconds
Phase stability	15 microns rms
Close packing	1.25 dish diameters (15.0 m) between azimut'
Solar observing	Allowed
Transportability	Transportable on a rubber-tired vehicle

The antennas will be designed and built by one or more commercial companies. Prototype antennas are being built for the US and European ALMA partners by Vertex Antenna Systems LLC (Santa Clara, CA) and European Industrial Engineering (EIE) (Mestre, Italy) respectively.

4.1. Introduction

The "antenna" subsystem of ALMA is here defined to include the following equipment:

12 m diameter primary reflector including quadripod subreflector support legs.
Secondary reflector and its servo-controlled positioning platform, including nutation.
A receiver cabin and its HVAC system.
Alt/az mount, the drive systems on the mount and the servo-system controller for the drives.
Metrology instrumentation such as temperature probes, tiltmeters, laser metrology systems, etc.
Power distribution cabling on the antenna and the cable wraps for these cables.
Platforms for mounting auxiliary equipment such as cryogenic compressors.
Antenna foundation design but not fabrication.
Antenna transporter vehicle.

The detailed specifications for all of this equipment, except for the Antenna transporter, can be found in the contract for the prototype antenna (NRAO, 2000) and in the ICD's (ICD, 2000). A summary of the requirements, including the antenna transporter, is provided below.

4.2. Specifications and Requirements

4.2.1 Operating Environment

The following operating environment defines the environment on the ALMA site in Chile. The first two prototype antennas will be tested at the VLA site in New Mexico (altitude 2100 m). It is likely that all production antennas will be assembled and undergo preliminary testing at San Pedro de Atacama (altitude 2440 m). Survival conditions at these test sites are no worse than the survival conditions on the ALMA site, except for the temperature maximum which has been selected as adequate for the test sites.

4.2.1.1 Location: Northern Chile, latitude -23d01m S, longitude 67d45m W.

4.2.1.2 Altitude: 5000 m (16400 ft) The barometric pressure at this altitude is 55% of its sea-level value.

4.2.1.3 Maximum Wind Velocity: The antenna must survive 65 m/sec (145 mph) without damage when positioned in its stow position.

4.2.1.4 Temperature: The antenna must operate correctly in the temperature range -20 C to 20 C. It must survive without damage the range -30 C to 40 C. The annual average temperature on the site is -4 C.

4.2.1.5 Precipitation: Annual precipitation on the site is in the range 10 cm to 30 cm per year. Most of this falls as snow but thunderstorms do occur and so brief periods of heavy rain and hail are possible. The antenna must be designed to survive, without damage, the following conditions: maximum rate of rainfall 5 cm/hr, hailstones 2 cm diameter at 25 m/s, snow load 100 kg/sq.m on reflector surface, radial ice on all exposed surfaces 1 cm. Surface heating to prevent snow and ice buildup not required.

4.2.1.6 Humidity: The monthly average humidity in the summer (January) is 53% and in the winter (June) it is 31%. The annual average is 39%. The monthly average water vapor pressure in the summer (January) is

4.0 hPa (4 gm/sq.cm) and in the winter (July) it is 1.2 hPa. The annual average is 2.3 hPa.

4.2.1.7 Insolation: The site location on the southern tropic, the high altitude and low water vapor result in insolation rates amongst the highest in the world. The median midday solar flux in the wavelength range 0.3-60 micrometers for the months of December and June are 1290 w/sq.m and 840 w/sq.m respectively. Ultraviolet radiation will be approximately 70% higher than at sea-level.

4.2.1.8 Lightning. Thunderstorms occur on the site so the antenna must be equipped with a lightning protection system.

4.2.1.9 Dust and Grit. The site ground surface is volcanic soil and gravel with no vegetation of any kind to stabilize the surface. It is likely that wind-blown dust and grit will be a factor for machinery operating on the site but this problem has not yet been well quantified.

4.2.1.10 Earthquake. The ALMA site is in a seismically active zone, but the source of the earthquakes, the tectonic plate interface, is more than 100 km below the surface so that the strength of the earthquakes is lower than the strength experienced closer to the Chilean coast. Design for 0.3G horizontal or 0.3G vertical acceleration.

4.2.2 General Configuration

The antenna will be a symmetric paraboloidal reflector, of diameter 12 m, mounted on an elevation over azimuth mount. Subreflector support legs will be a quadripod configuration. A reflector surface consisting of machined aluminum panels, or a panel technology providing equivalent accuracy, will be used. The reflector surface will be mounted on a carbon fiber reinforced plastic (CFRP) reflector backup structure (BUS). The BUS could be built completely of CFRP or could consist of CFRP structures connected by metal connectors. The quadripod will be made of CFRP. The reflector surfaces of the antenna will not be painted.

All drawings will have metric dimensions. All fasteners will be metric. The use of standard metric cross-sections for construction materials is preferred but will not be required if it results in a cost increase.

The antenna will be designed for a lifetime of 30 years. For design purposes it will be assumed that the antenna will execute not less than 270,000 complete cycles of elevation motion, where a complete cycle of elevation motion is defined to be movement of the reflector from its lower elevation limit up to its upper elevation limit and back down to its lower elevation limit. During its lifetime the antenna will execute not less than 200,000,000 degrees of total motion about each axis.

The supply voltage for the antenna will be the European standards, 230 v (single phase), 400 v (3 phase). All electrical systems must operate correctly on both a 50 Hz or a 60 Hz supply.

The antenna will be designed so as to conform to all relevant Occupational Safety and Health Administration (OSHA) or European safety codes.

4.2.3 Reflector Geometry

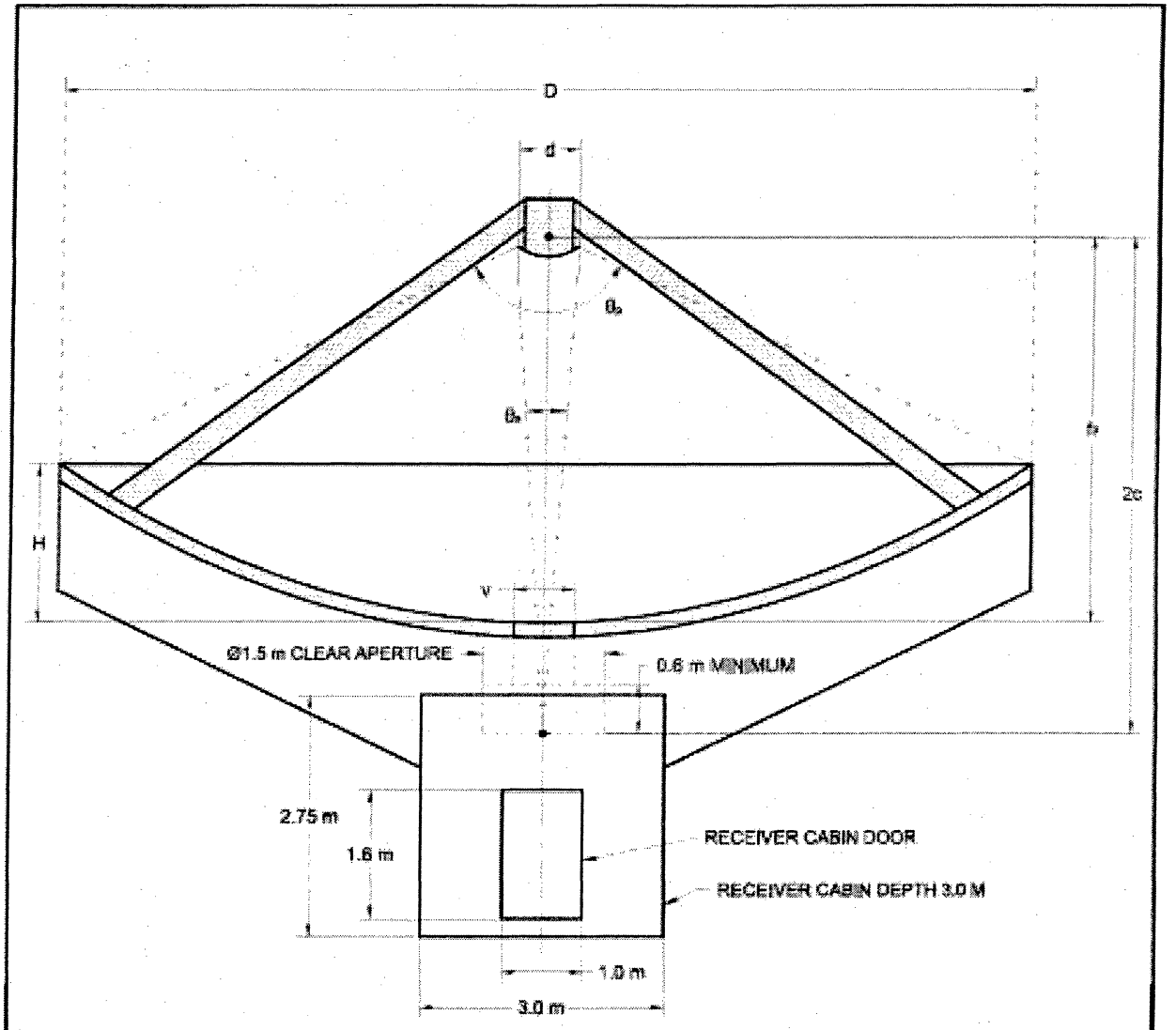
The receivers will be located at the secondary focus of the Cassegrain geometry shown in Figure 1 (taken from Lamb,1999).

Figure 1. Optics Layout for ALMA Antenna

APPENDIX A

ANTENNA OPTICAL CONFIGURATION

ALMA
OPTICAL CONFIGURATION



OPTICAL CONFIGURATION

D	Primary Aperture	12.0 m	472 in
f_p	Focal Length of Primary	4.8 m	189 in
	f_p / D of Primary	0.40	0.40
d	Secondary Aperture	0.75 m	29.5 in
	Final f / D	8.00	8.00
	Magnification Factor	20.0	20.0
θ_p	Primary Angle of Illumination	128.02°	128.02°
θ_s	Secondary Angle of Illumination	7.16°	7.16°
2c	Distance Between Primary and Secondary Foci	6.177 m	243.2 in
H	Depth of Primary	1.97 m	77.7 in
v	Primary Vertex Hole Clear Aperture	0.75 m	29.5 in

VERSION 2 (1999 MAY 21)

4.2.4 Range of Motion

Antenna foundations will be constructed so that the azimuth axis of an antenna is parallel to local gravity at the pad. For observations close to the zenith this will result in a difference in parallactic angle between antennas.

Minimum elevation angle for observations: 2 deg

Maximum elevation angle for observations: 125 deg. Cone of avoidance at the zenith: 0.2 deg in radius for normal sidereal tracking. Because of the high velocities and accelerations required for fast switching or on-the-fly mapping (see section 4.2.8 below) there will be a region around the zenith about 30 deg in radius, where azimuth switching times are degraded .

Stow position for wind survival: elevation 15 deg (this position was chosen so that, during a winter storm, the reflector can be oriented with its back into the wind to prevent build up of snow and ice in the dish. Stow position for maintenance: zenith (this position was chosen to prevent an antenna undergoing maintenance from mechanically interfering with an adjacent antenna in the most compact array).

Range of azimuth motion: 270 degrees either side of due north.

4.2.5 Reflector Surface Accuracy

The surface accuracy goal is 20 microns and the hard specification is 25 micrometers rms, including the subreflector contributions. The 20 micron and 25 micron surfaces will provide antenna surface efficiencies at 300GHz /900 Ghz of 94%/57% and 91%/41% respectively. At night this accuracy is to be achieved in a wind of 9 m/s which is approximately the 90^{th} percentile wind for nighttime (2000 hrs to 0800 hrs) observing. During the day this accuracy is to be achieved for any orientation of solar illumination in a wind of 6 m/s. During the day the focus can be calibrated astronomically every 30 min.

The final, precision measurement of the surface will be done by the ALMA Project using holography. For the prototype antenna the contractor will provide a surface setting accuracy of 100 microns.The panel adjusters will be calibrated so that an adjustment point can be moved with a resolution of 5 micrometers. A full surface adjustment must require no more than 16 person-hours of work.

4.2.6 Pointing Accuracy

A pointing accuracy in "offset" pointing mode (calibrator 2 deg away every 15 minutes of time) of 1/30th primary beamwidth rms at 300 GHz is required. The antenna specification is 0.6 arcsec RSS for offset pointing, 2.0 arcsec RSS for absolute pointing. At night this accuracy is to be achieved in a wind of 9.0 m/s which is the 90^{th} percentile wind for nighttime (2000 hrs to 0800 hrs) observing. During the day this accuracy is to be achieved for any orientation of solar illumination in a wind of 6 m/s.

4.2.7 Metrology

Provision will be made in the antenna design for inclusion of metrology equipment which will allow antenna pointing to be corrected for structural deformation caused by wind or thermal loading. Metrology systems for possible incorporation into the antenna include: a laser/quadrant-detector system to measure quadripod movement, tiltmeters, temperature probes and laser/retroreflector systems. The antenna contractor will provide any metrology that he considers essential for meeting performance specifications. An optical telescope with CCD camera for pointing on stars will be supplied by the ALMA Project.

4.2.8 Fast Motion Capability

Three observing modes require the ALMA antenna to have special fast motion capabilities: fast switching phase calibration, on-the-fly total power mapping and on-the-fly interferometric mosaicking.

Fast switching: The antenna will move 1.5 degrees on the sky and settle to within 3 arcsec peak pointing error, all in 1.5 seconds of time. The switching acceleration profile will be carefully designed so as to avoid exciting the lowest structural resonant frequency of the antenna. The maximum velocity and acceleration required for fast switching are 3 deg/sec and 12 deg/sec/sec on the sky respectively, with both axes able to move at this rate simultaneously. It is expected that this velocity and acceleration will be achievable in azimuth only for zenith angles greater than 30 deg (this implies maximum azimuth velocity and acceleration of 6 deg/sec and 24 deg/sec/sec respectively).

Analysis of the expected use of this fast switching mode (Holdaway, 1997) indicates that the antenna should be designed to survive 30-50 million cycles of fast switching during an assumed 30 year life.

On-the-fly mapping: In this mode the antenna will scan at a rate of up to 0.5 deg/sec across a large object, several or many beamwidths in size, and then turn around as rapidly as possible and scan back across the source in the opposite direction. A maximum acceleration of 12 deg/sec/sec is required for the turn around. While the antenna is scanning across the source the antenna position must be recorded at a rate sufficient to provide an angular sampling interval on the sky of wavelength/(2D) radians. For 0.5 deg/sec motion and 900 GHz observations this requires antenna position readout every 2 msec. The antenna positions should be accurate to 1 arcsec. As the antenna tracks across the source it is not necessary for the position at any time to be precisely a precommanded position; it is sufficient to simply know where the antenna is actually pointing and all antennas need not point precisely at the same position.

On-the-fly interferometric mosaicking requires interferometry data to be taken while the antenna is continuously scanning across the source. It is expected that the antenna velocity will be only one-tenth of its mapping-on-the fly value (see previous paragraph), but in this case all antennas must point to the same position at the same time to within 1 arcsec rms.

4.2.9 Subreflector Position Control

The subreflector will be supported on a platform which allows movement in all 3 linear directions. The precision of the mechanism will be adequate to allow the subreflector to be positioned, under computer command, with sufficient accuracy to prevent gain loss of more than 1% at 900 Ghz due to focus, comatic or astigmatic aberration. Position will be correctable on timescales of seconds. Total axial focus motion is 2.0 cm.

In addition to the above listed linear motions the prototype antennas will also be equipped with a subreflector nutator with the following specifications: 10 ms transition time, 10 Hz repetition rate, +- 1.5' throw, full dynamic balance. The decision as to whether all antennas will be equipped with nutators will be made after testing the prototype antennas.

4.2.10 Phase Stability

Phase errors caused by variations in the propagation pathlength through the antenna can be rapidly or slowly varying. Fast phase changes are primarily caused by the wind and the peak pathlength variation in a 9.0 m/s wind must be no more than 15 microns. Slow phase changes are primarily due to variations in the temperature of the antenna and the goal is to keep these phase errors small enough so that the residual errors after an astronomical phase calibration every 3 min are small enough to allow observations at 900 Ghz.

4.2.11 Close Packing

In the smallest array the antennas must be placed close together. It will be possible to place the antennas with their azimuth axis within 15 m (1.25 D) of each other without any possibility of the antennas hitting each other, no matter what the relative orientation of the two antennas.

4.2.12 Solar Observing

Direct observations of the sun will be allowed. All surface accuracy and pointing requirements must be met while observing the sun and a suitable surface treatment of the primary reflector surface must be provided to prevent solar heating damage of the subreflector or its support legs. When observing the sun the solar heating of the secondary focal plane must be less than 0.3 W/sq.cm.

4.2.13 Low Antenna Noise

Contributions to system noise from the antenna, due to such mechanisms as scattering of ground noise into the feed and resistive loss of reflector surfaces, will be minimized as much as possible without compromising the surface accuracy and pointing requirements. Design features to be considered to achieve this goal include supporting the subreflector support legs close to the edge of the reflector and shaping the underside of the support legs to reduce ground pickup. Geometrical blockage will not exceed 3.0% and resistive loss of any reflective surface will not exceed 1.0% at frequencies up to 950 GHz.

4.2.14 Transportability

To move the antennas from one array configuration to another the antennas will be picked up and carried on a transporter vehicle which runs on a gravel road on rubber tires. The transporter with an antenna on board will be able to negotiate a 15 % grade, turn a corner with a minimum turning radius of 10 m and travel at 10 km/hr on the flat and 5 km/hr up a 10% grade. An unloaded transporter must be able to travel at 20 km/hr on the flat. The transporter must be able to safely move an antenna in winds up to 16 m/s (this is approximately the 95th percentile for the winds on the site at 1600 hrs local time, the time at which the winds are maximum each day). A stationary transporter with an antenna on board will survive winds up to 65 m/s; if necessary, structure can be deployed to stabilize the transporter on the ground in this survival mode. To withstand the bumps and jolts of transportation and pickup/putdown the antenna will be designed to survive shock loads of 4 G vertical and 2 G horizontal acceleration.

The transporter will carry an auxiliary generator to keep all electrical systems on the antenna operational during a move. The transporter will pick up the antenna above its azimuth bearing so that the azimuth bearing and drive can be used to rotate the base of the antenna to simplify bolt hole alignment when an antenna is placed on a pad. It may be desirable to oxygenate the air in the transporter operator's cabin so the cabin must not have large air leaks.

When an antenna is picked up a time goal of 20 min is required from the time of arrival of the transporter to the time of departure with an antenna on board. When an antenna is placed down on a pad a time goal of 30 min is required from the time of arrival of the transporter until the transporter has departed and the antenna is ready to be pointed.

A current design concept for the transporter can be found in ICD NO. 5 ANTENNA/TRANSPORTER INTERFACE available at ICD, 2000
for the US antenna and ALMA-Europe Antenna WWW Page for the European antenna. Because of differences in the antenna design, the US antenna has 4 transporter attachment points and the European

antenna has 3. The current goal is to design the prototype transporter so that it can pick up either antenna.

4.2.15 Receiver Cabin

A receiver cabin with dimensions approximately as shown in Figure 1 will be provided at the Cassegrain focus. Temperature in the equipment air supply plenum will be maintained by an antenna mounted HVAC system at a setpoint in the range 16-22 C to an accuracy of +- 1C. The electrical power consumption of equipment in the cabin will not be greater than 10 kw. The mass of equipment in the cabin will not be greater than 1600 kg.

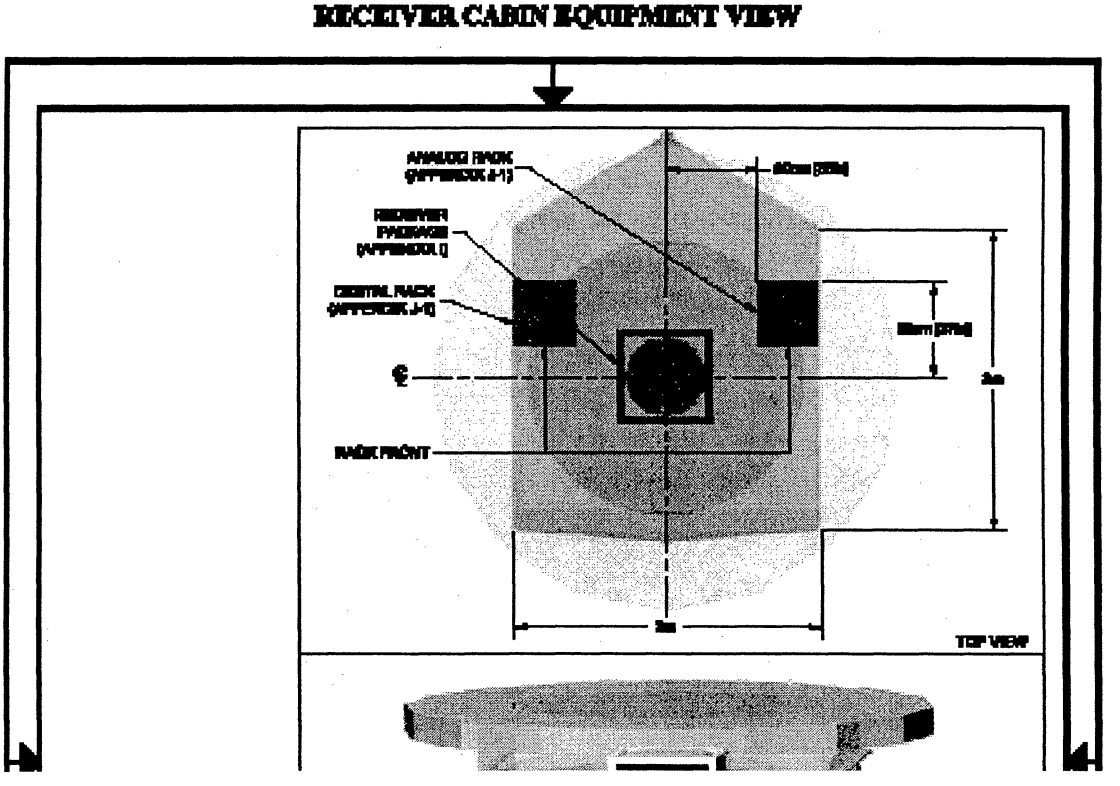
A built-in mechanism will be provided so that a receiver can be lifted from the ground, through the cabin door and into its observing location, all without significant man-handling of the receiver. Part of the installation of a receiver may involve the use of a separate special purpose vehicle, such as a high fork-lift, which lifts the receiver through the cabin door.

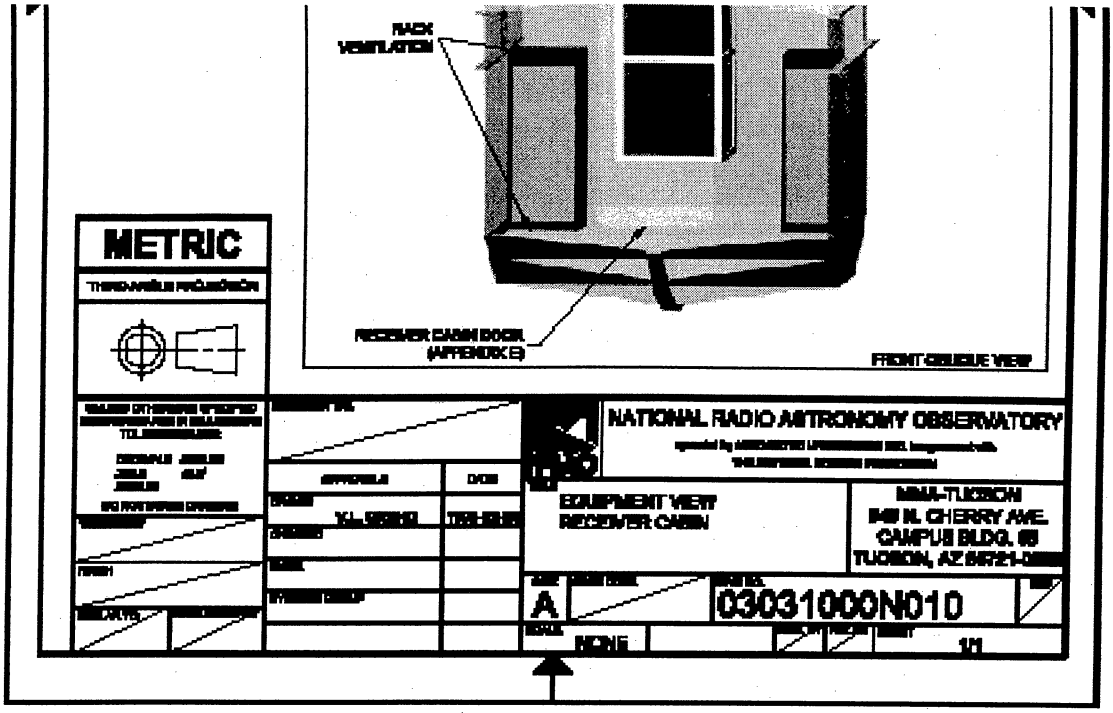
The cabin will be watertight and a thin RF-transparent membrane will cover the aperture through which the RF beam enters the cabin. A computer actuated shutter will be deployable to protect the membrane when necessary.

It may be desirable to oxygenate the cabin air when workers are inside so the cabin must not have large air leaks.

A typical layout for the receiver cabin is shown in Figure 2.

Figure 2. Typical layout for ALMA receiver cabin.





J-2

4.2.16 Monitor and Control

Lists of monitor and control points for the antenna can be found in
ICD NO. 9 ANTENNA/MONITOR AND CONTROL INTERFACE, available at (ICD, 2000) for the US antenna and at ALMA-Europe Antenna WWW Page for the European antenna.

4.2.17 Interfaces to Other Subsystems

Interfaces to the various ALMA subsystems are defined in the ICDs available at (ICD, 2000) for the US antenna and at ALMA-Europe Antenna WWW Page for the European antenna.

ICD NO. 1 ANTENNA/RECEIVER INTERFACE
ICD NO. 2 ANTENNA/APEX INTERFACE
ICD NO. 3 ANTENNA/SITE ELECTRIC POWER INTERFACE
ICD NO. 4 ANTENNA SITE FOUNDATION INTERFACE
ICD NO. 5 ANTENNA/TRANSPORTER INTERFACE
ICD NO. 6 ANTENNA/CABLE WRAP INTERFACE
ICD NO. 7 ANTENNA/HELIUM COMPRESSOR INTERFACE
ICD NO. 8 ANTENNA/OPTICAL POINTING TELESCOPE INTERFACE
ICD NO. 9 ANTENNA/MONITOR AND CONTROL INTERFACE
ICD NO. 10 RECEIVER CABIN EQUIPMENT RACK INTERFACE
ICD NO. 11 BASIC ANTENNA DEFINITIONS

4.2.18 Maintenance and Reliability

Because of the remote site and large number of antennas the reliability and maintainability of the antennas are important. The antennas will be designed so that, with normal preventive maintenance, they should operate for 30 years without requiring elevation or azimuth bearing or reflector surface replacement. Although they should not be required, straightforward elevation and azimuth bearing replacement procedures must be included in the antenna design. All normal repair and maintenance actions should be able to be completed by a two- person crew in 4 hours. To the maximum extent possible all equipment on the antenna should be "modularized" so that a failure can be cured by simply swapping out the failed component without the need for any repair in place. Examples of equipment which should be designed for easy replacement includes gear boxes, drive motors, HVAC equipment, servo-system electronic components and the subreflector position control mechanism.

4.2.19 Manufacture and Assembly

The antenna will be designed for economic production costs.

The first two antennas will be tested initially at the VLA site in the US. At least one of the prototype antennas will later be shipped to the ALMA site so the ability to disassemble the antenna into pieces for overseas shipping is required.

The high altitude and remoteness of the ALMA site make it desirable to minimize the amount of work required on the high site. It is expected that the antennas will be assembled, outfitted and tested at an Operations Support Facility in San Pedro de Atacama 50 km from the ALMA site at an altitude of 2400 m. They will be carried to the ALMA site on the transporter vehicle or, in the event that this proves not to be feasible, they will be disassembled into just two pieces, the mount and the reflector, for transportation to the site on trucks. Thus the antenna will be designed for easy disassembly at the elevation axis and both the reflector and mount must have pickup points for handling as single units.

4.3. Design Concepts

Four antenna concepts were developed within the ALMA Project.

A 10 m concept developed principally at NRAO and BIMA(Lugten et.al., 1999)
A 10 m concept developed principally at OVRO(Woody and Lamb, 1999)
A 12 m concept developed principally at ESO (Andersen, 1999)
A 12 m concept developed principally at IRAM (Plathner, 1999)
These concepts were provided to the bidders during the proposal phases for the NRAO and ESO contracts to assist the bidders in making their proposals..

Concept pictures of the antennas being designed by the two contractors are shown in Figures 3 and 4.

Figure 3. European Industrial Engineering antenna concept.

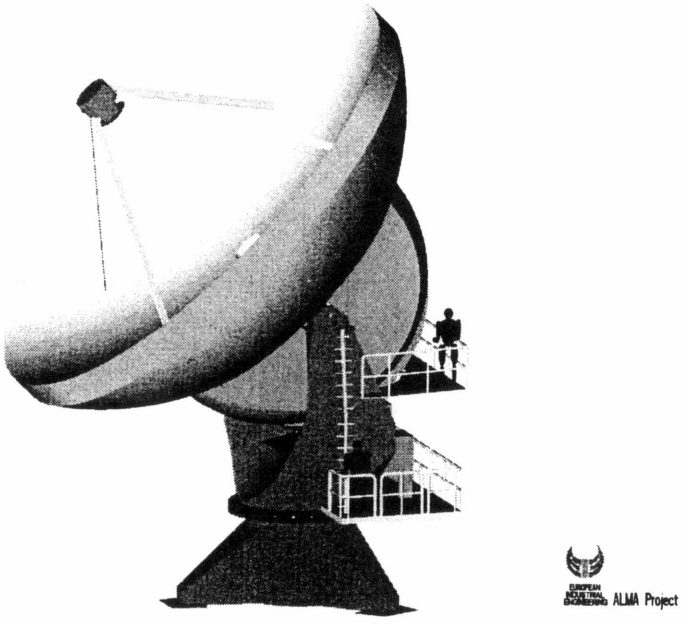
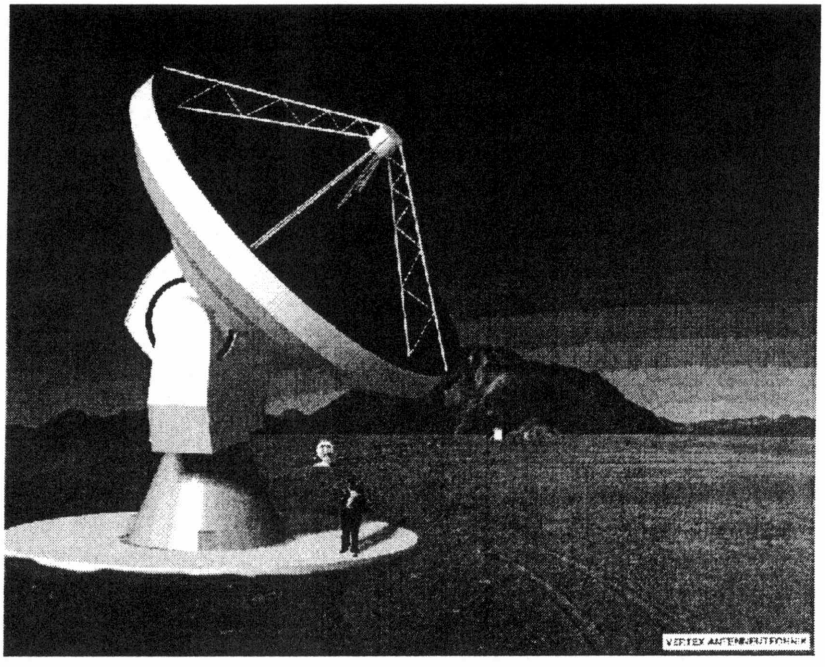


Figure 4. Vertex Antenna Systems antenna concept



4.4. Procurement and Construction Plans

The following is a list of the principal tasks, with date goals for the major milestones, required in the ALMA Project to progress from prototype antenna delivery to a contract for production antennas:

Installation of prototype antennas at VLA site - 4Q 2001
Contractor Acceptance Tests - 4Q2001
Tests
Evaluation
Selection of production design
Specification for Production Antennas (build to print) - 2Q 2003
Call for Tenders

Bidding
Selection
Approval
Contracting - 4Q 2003

4.5. References

T. Andersen, "Feasibility Study for a 12 m Submillimeter Antenna", ALMA Memo 253, Feb 1999.

M. Holdaway, "How many switching cycles will the MMA make in its lifetime", ALMA Memo 174, 1997.

Antenna Interface Control Documents, ALMA, NRAO, (ICD, 2000), Feb 2000.

J. Lamb, 1999, Optimized Optical Layout for MMA 12-m Antennas, ALMA Memo 246, Jan 1999.

J. Lutgen, J. Kingsley, J. Cheng, V. Gasho, M. Fleming, "A 10-m Antenna Design for the Millimeter Array", ALMA Memo 240, Feb 1999.

NRAO, ALMA US Prototype Antenna Purchase Order, February, 2000.

D. Plathner, "A 12m Telescope for the MMA-LSA Project", ALMA Memo 259, April 1999.

D.P. Woody, J.W. Lamb, "A Design for a Precision 10-m Sub-Millimeter Antenna", ALMA Memo 241, Mar 1999.

ALMA Project Book, Chapter 5.

ALMA FRONT ENDS

Wolfgang Wild & John Payne
Last revised 2001-Feb-07

Revision History

2000-12-12: First ALMA version
2001-02-07: Figure 5.1 inserted

5 Alma Project Book. Front End

5.1 Introduction

This chapter of the Project Book describes the front ends that will be built for ALMA. As with all of the Project Book this chapter should be regarded as a "living document" subject to change at any time. Many of the details of the front end are undecided at the present time and will be included as the final design evolves. The present form of the front end is the result of efforts by several groups, each of which has contributed to the Project Book. From its conception it was recognized that the front end for ALMA would be quite different from any front end previously built for radio astronomy. Listed below are the major considerations that have driven the concepts behind the present front end design.

- For reasons of access, weight and all the usual reasons the decision was made at the start of the project to install the front end at the Cassegrain focus.
- Good performance over the complete range of frequencies to be covered by ALMA. This resulted in the frequency range of 31 GHz to 950 GHz being divided into ten separate bands. This division permits the optimization of both noise performance and optical coupling at all frequencies over the array's operating frequency band.
- Each frequency band will have two channels tuned to identical frequencies. The decision has been made to have these two channels receive two linearly polarized orthogonal signals.
- High reliability. It is recognized that building a front end that must be replicated at least 64 times is quite different from building a single front end for one telescope. High reliability is obviously a major consideration. It has been recognized that high reliability and the very best performance may not always be achievable together.
- The front end should be modular so that one easy to install self-contained receiver should cover that one particular frequency band. This was felt to be necessary to accommodate the desire to have groups in different locations produce receivers for the different bands. These self-contained receivers have become to be known as "cartridges".
- The front end itself, containing the ten cartridges should be one self-contained unit, easily removable from the antenna for servicing.
- The front end and all its components should be able to be produced, assembled and tested in a manner appropriate to the manufacture of front ends on this scale. It is recognized that the resulting design may be quite different from that produced for optimum performance on a single telescope.
- Servicing and all matters to do with installation on the antenna should be as easy as possible and appropriate to conditions at the high site.
- The front end should operate for long periods - around one year - with no maintenance. Experience at various telescopes in operation for many years suggest that this is a realistic goal although it was felt that this requirement ruled out the use of a "hybrid" cryostat involving the use of liquid helium.

5.2 Specifications

The document *Specifications for the ALMA Front End Assembly* (latest version) contains the detailed specifications. These have been approved by the AEC with a pending change request regarding the extension of band 3 down to 84 GHz. The main specifications are:

- Frequency coverage: from 31.3 to 950 GHz in 10 bands (see Table 5.1)

- simultaneous reception of two orthogonal polarizations
- receiver noise between 6 and 10 times $h\nu/k$ over 80% of the band, with a goal of achieving 3 to 8 times $h\nu/k$, depending on the band
- IF bandwidth 8 GHz total per polarization
- observations at one frequency at a time (no dual frequency obervations)
- inclusion of a water vapour radiometer using the 183 GHz line for phase correction.

For details, see the full *Specifications for the ALMA Front End Assembly* (latest version).

Table 5.1 – Frequency bands for ALMA

Band	from (GHz)	to (GHz)
1	31.3	45
2	67	90
3	89*	116
4	125	163
5	163	211
6	211	275
7	275	370
8	385	500
9	602	720
10	787	950

* change request to 84 GHz underway

5.3 Overall System Description

The following is a brief description of the overall front end system. Details may be found in the relevant sections of the chapter. The receiver consists of a circular dewar 1.0 m in diameter and 0.7 m in height. The individual receivers (cartridges) are inserted into the bottom of the circular dewar. LO, IF and circuit connections are made to cartridges from the bottom of the cartridge: the millimeter/sub-millimeter signal enters the top of the cartridge via a vacuum window and infra-red filter. The entrance to the various frequency band cartridges is in the focal plane of the antenna so frequency selection is achieved by adjusting the telescope pointing. This very simple configuration is shown in Figure 5.1

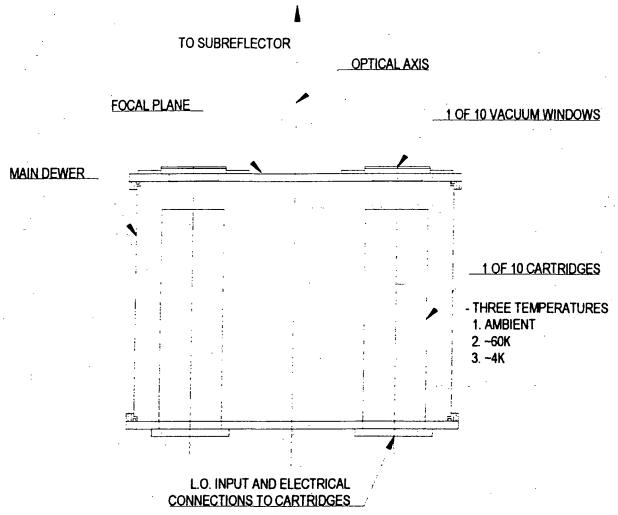


Figure 5.1

The necessary IF processing, various control circuits and computer interfaces will be packaged external to the dewar with the configuration yet to be decided. The overall configuration of the front end is illustrated in block diagram form in Figure 5.2 (the front end configuration) and Figure 5.3 (the cartridge configuration).

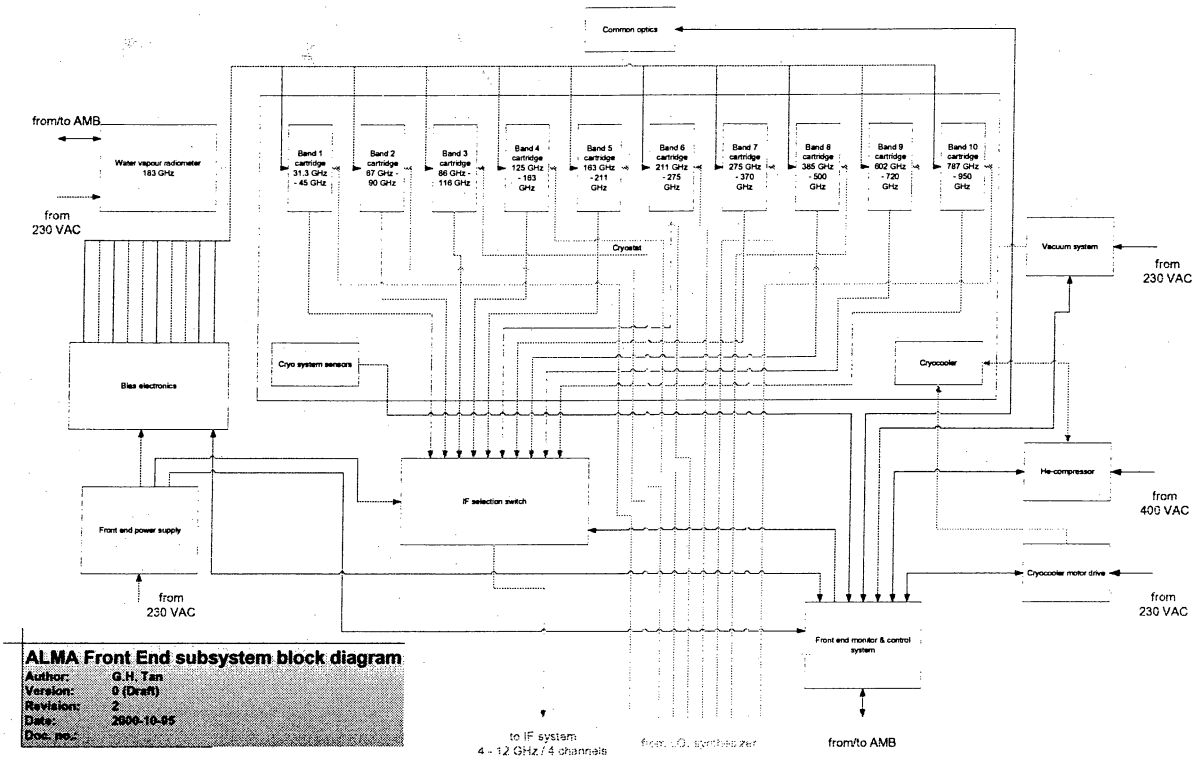
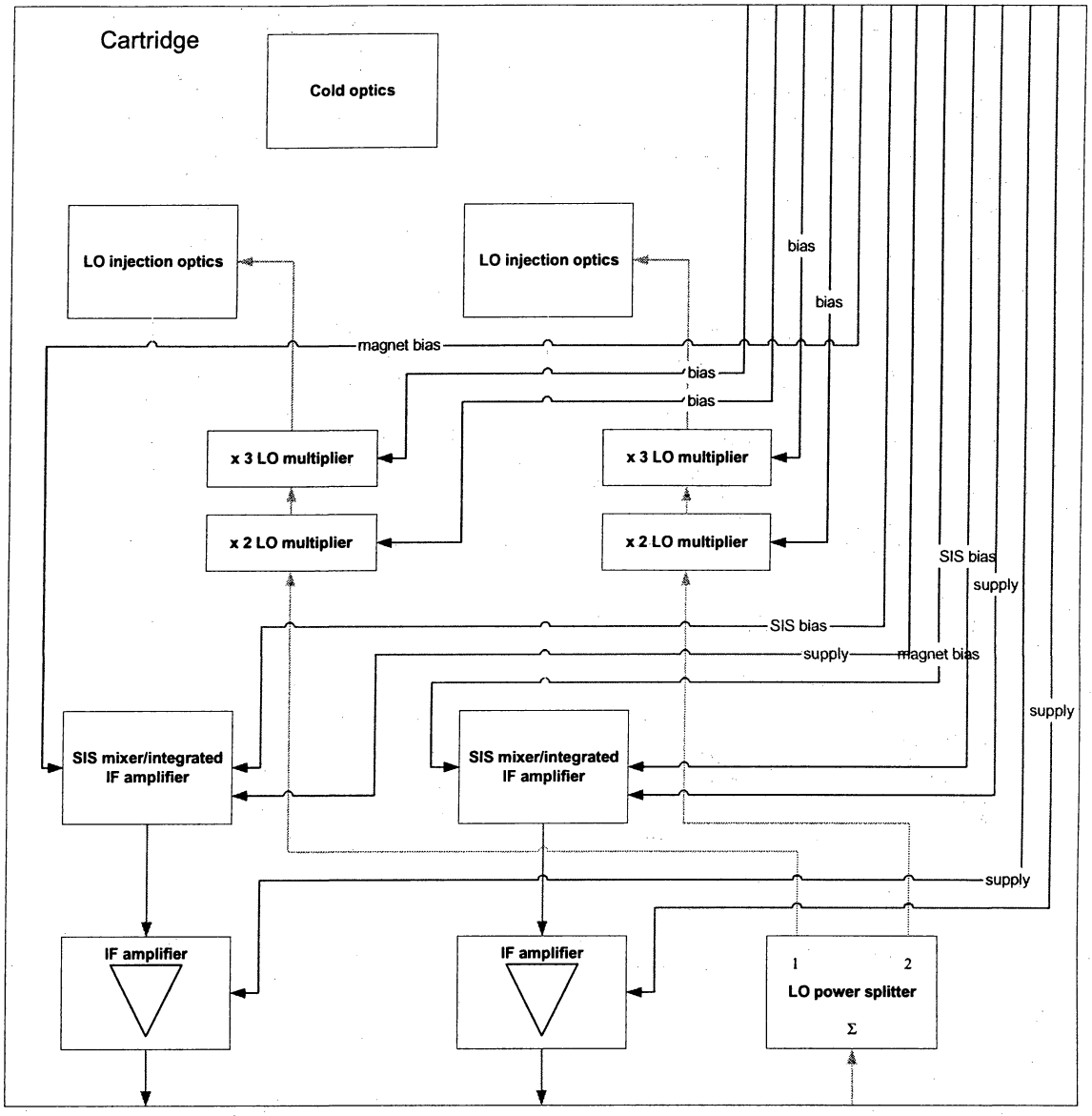

Common optics
from/to AMB
Water vapour radiometer 183 GHz
from 230 VAC
Band 1 cartridge 31.3 GHz - 45 GHz
Band 2 cartridge 67 GHz - 90 GHz
Band 3 cartridge 86 GHz - 116 GHz
Band 4 cartridge 125 GHz - 183 GHz
Band 5 cartridge 163 GHz - 211 GHz
Band 6 cartridge 211 GHz - 275 GHz
Band 7 cartridge 275 GHz - 370 GHz
Band 8 cartridge 385 GHz - 500 GHz
Band 9 cartridge 602 GHz - 720 GHz
Band 10 cartridge 787 GHz - 950 GHz
Cryostat
Vacuum system
from 230 VAC
Cryo system sensors
Cryocooler
Bias electronics
He-compressor
from 400 VAC
IF selection switch
Front end power supply
Cryocooler motor drive
from 230 VAC
from 230 VAC
Front end monitor & control system
ALMA Front End subsystem block diagram
Author: G.H. Tan
Version: 0 (Draft)
Revision: 2
Date: 2000-10-05
Doc. no.:
to IF system
4 - 12 GHz / 4 channels
from/to AMB

Figure 5.2

bias/supply lines from bias electronics



IF outputs to IF selection switch

from LO_1 synthesizer

ALMA band 9 cartridge block diagram

Author:	**G.H. Tan**
Version:	**0 (Draft)**
Revision:	**3**
Date:	**2000-10-05**
Doc. no.:	

Figure 5.3

5.4 The optical arrangement

(See separate sub-chapter on Optics in Project Book Table of Contents)

5.5 The dewar and Cryogenic Cooler

(See separate chapter on Cryogens in Project Book Table of Contents)

5.6 Receiver Band Cartridges

5.6.1 Introduction

The concept of receiver cartridges is being developed for various reasons. The idea of a millimeter front end consisting of various well defined inserts is not new and was developed at NRAO many years ago. For the ALMA receivers the idea of each receiver band being a single unit, testable separately of the main receiver Dewar was particularly appealing given the participation of various groups geographically separated and the desire of these groups to be responsible for different receiver bands. The cartridge approach also minimizes the number of interfaces (optical, mechanical, electrical and thermal) and allows that a cartridge may be built and tested in one location and installed in the main front end dewar later. The constraints on cartridge size are outlined in *** along with drawings of the basic cartridge. The *ALMA Scientific Advisory Committee (ASAC)* has identified four receiver bands out of ten as first priority for development and installation. These are bands 3 (89 – 116 GHz), 6 (211 – 275 GHz), 7 (275 – 370 GHz) and 9 (602 – 720 GHz). The design approaches for these four initial bands are described in the following sections.

5.6.2 Band 3 Cartridge development at NRAO

Last revised on November, 05 – 2000 by A.R. Kerr, S.-K. Pan and John Webber

Revision History: 2000-10-05: New

5.6.2.1 Introduction

This band is presently defined as covering 89-116 Ghz. However there is a change order in process to change the lower end of this band to 84 Ghz. In recent years HFET amplifiers have been developed which meet the ALMA specifications with one exception and would be most attractive to use. Since the ALMA receivers are intended for both interferometric and single-dish total power observations the radiometric stability of the receivers is important. Based on the work of Wollack and Pospiezalsky [?] the so called 1/f noise produced by a wideband HFET amplifier would exceed the ALMA specifications. This problem has been well summarized by Webber (see section 5.6.2.2.3). Due to this potential problem work had progressed on the development of a fixed tuned SIS mixer for band 3 as described below.

5.6.2.2 SIS Mixer Development for band 3

5.6.2.2.1 Summary

This section describes the SIS mixer development plan for the ALMA front-ends for Band 3, nominally 90-116 GHz, for which the science group has requested assessment of the feasibility of extension to 86-116 GHz. The primary driver for this development is the 1/f gain noise of HFET receivers (discussed below). The goals for the design and development phase are:

1. carry out a thorough study on the feasibility of developing balanced sideband-separating mixers with integrated IF amplifiers meeting the ALMA specifications,
2. develop and evaluate a fully-integrated (MMIC) fixed-tuned waveguide mixer and use it as a building block in the balanced and sideband separating mixer, and
3. provide technical and budgetary information gathered in this study to ALMA management and scientific advisory committee as one of the basis of choosing SIS or HFET receivers for this band. If it is decided to use SIS receivers in this band, the goals for the construction phase are to mass-produce SIS mixers with repeatable performance at minimum total cost.

Table 5.2 - SIS Receiver Specifications

Item	Specification
Receiver noise temperature	Noise sufficiently low to produce single sideband receiver noise referred to the vacuum window of 60 K over 80% of band and 80 K at any frequency
Frequency band covered	Band 3, 90-116 GHz, extended to 86-116 GHz if possible
IF bandwidth	Minimum of 4 GHz total, falling in band 4-12 GHz; want 8 GHz for each sideband if possible
Linearity	TBD
Configuration	Balanced operation, sideband separation > 10 dB, no mechanical tuners

5.6.2.2.2 Development

5.6.2.2.2.1 Design Requirements

In order to meet the receiver specifications listed in Table 5.2, the following properties are required in Band 3 SIS mixers:

- Low mixer noise temperature.
- Low mixer conversion loss (~0 dB DSB). While gain is possible in SIS mixers, substantial conversion gain is undesirable because of the reduced dynamic range and greater possibility of out-of-band instability.
- High saturation power. Receivers should be capable of performing solar observations.
- Wide RF bandwidth (minimum of 26 GHz total, from 90 to 116 GHz, but extended to 30 GHz total, from 86 to 116 GHz, if possible).
- Wide IF bandwidth (minimum of 8 GHz total, from 4 to 12 GHz).
- A moderately well matched input.
- Operation into a 50-O IF amplifier with no matching impedance transformer is desirable. SIS mixers with matched output tend to have poor input match and, in certain cases, may have input reflection gain, which may increase the baseline ripples and the receiver's instability.

5.6.2.2.2.2 Single-Junction vs. Array

Theoretically, the performance of an N-junction array is the same as that of a single junction, which has the same overall impedance, provided that

- current is in phase all along the array and
- that all of the junctions of the array are identical.

The advantages of using N-junction arrays are:

- a greatly increased dynamic range (proportional to N^2),
- easy suppression of Josephson-effect noise
- less susceptibility to electric transients and
- easier fabrication (better yield).

The disadvantages of using arrays are:

- some experiments have shown that, contrary to the theoretical predictions, array mixers may have higher noise temperature than single-junction mixers and
- it requires N^2 more LO power to operate.

Since the NRAO is experienced in developing and operating quantum-limited low-noise array mixers in this frequency band and sufficient LO power is not an issue in Band 3, we have decided to use arrays in this band.

5.6.2.2.2.3 MMIC Design vs. Waveguide Hybrids

There are many ways to construct balanced sideband-separating mixers in the millimeter- and submillimeter-wave bands.

Two designs, a single-chip (MMIC) design developed at NRAO [1-3] and a design based on waveguide hybrids reported in ALMA Memo 316 [4], are in particular suitable for ALMA balanced sideband separating mixer development work. However, for Band 3, because the large size of single-chip balanced sideband-separating mixers will permit very few mixers per wafer, the approach using waveguide hybrids may be preferable to the MMIC approach.

5.6.2.2.2.4 Junction Parameters

Kerr and Pan [5] and Ke and Feldman [6] have developed SIS mixer design procedures. At 100 GHz, both procedures give similar optimum source and load conductance and junction ? R_NC product. Table 5.3 lists the junction parameters calculated using the design rules outlined in [5] with a source impedance of 35 O and load impedance of 50 O, a specific capacitance C_S = 65 fF/μm^2, R_NI_C = 1.8 mV and ? R_NC = 3.5 at 115 GHz.

Table 5.3 - Band 3 Device Parameters for UVA=s Niobium Trilayer Circuit Process

Jc	2,500 A/cm^2
Junction size (diameter)	2.3 μm
Normal Resistance of the Array	70 O
Cs	65 fF/μm^2
SiO dielectric constant	5.7
I2 (SiO)	2,000 Å

M3	4,000 Å
Pd/AU	300 Å

5.6.2.2.2.5 Mixer Design

A fully integrated (MMIC) fixed-tuned 86-116 GHz SIS waveguide mixer, similar to the NRAO 373 mixer [7], is currently being developed at the CDL for ALMA Band 3. Special design efforts have been made to meet ALMA's specifications. The circuit parasitics (capacitance and inductance) seen at the mixer's IF port have been minimized using a circuit layout similar to that described in [7] to meet ALMA's IF bandwidth specification. An additional RF matching circuit has been implemented to increase the RF bandwidth.

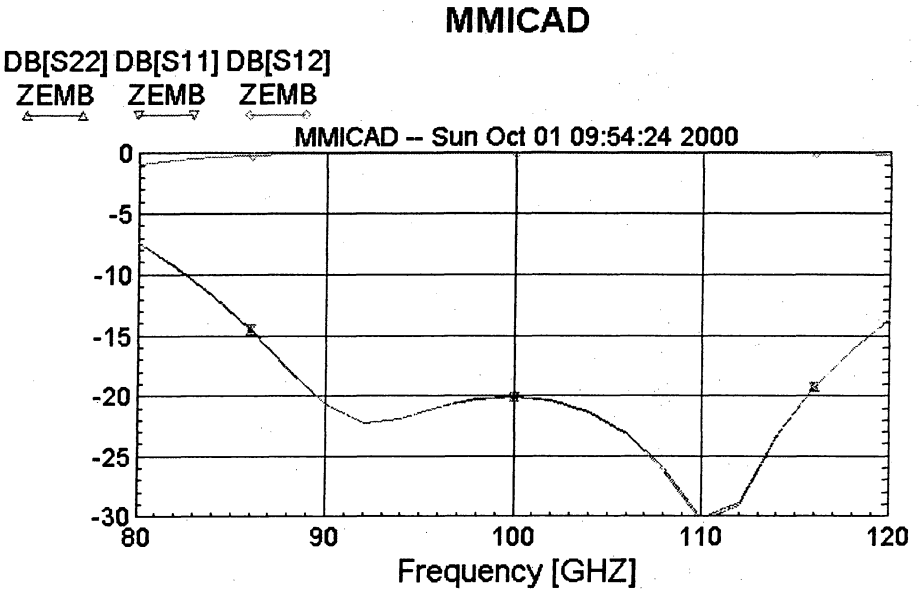


Figure 5.4 - Return loss of the coupling network to the SIS array. The return loss is the match seen at the 50 O waveguide probe to a 35 O optimum array source impedance.

Initial circuit analysis using MMICAD [8], shown in Figure 5.4, shows that it is possible to design a coupling network to provide good matching between waveguide probe and the array's optimum source admittance over the entire ALMA Band 3 frequency range. The RF embedding admittance seen by the array is shown in Figure 5.5.

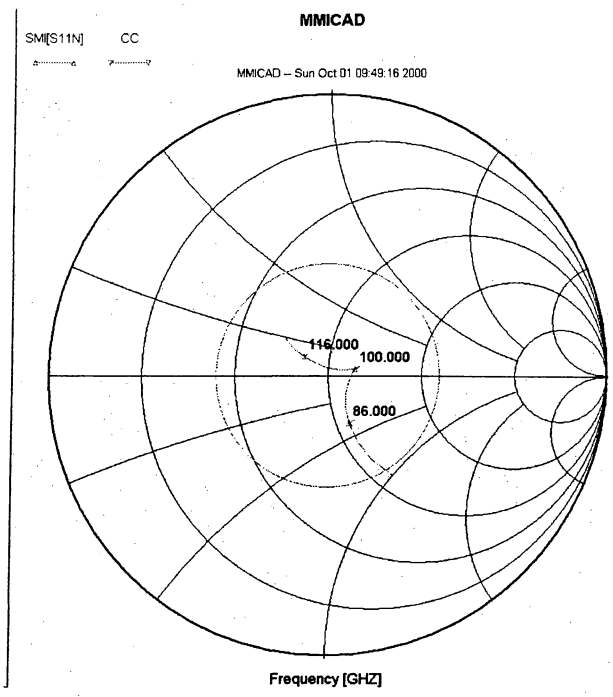


Figure 5.5 - RF embedding admittance seen by the array. The junction capacitance is taken as part of the embedding circuit. The circle is at |?| = 0.4. The Smith chart is normalized to the optimum source conductance.

The Smith chart is normalized to the optimum source admittance for the array, 0.0286 mhos in the present case. The junction capacitance is taken as part of the embedding circuit in this calculation. The circle at |?| = 0.4 indicates the range of embedding admittance within which acceptable SIS mixer performance will be attained.

5.6.2.2.2.6 Band 3 Milestones

Table 5.4 - Band 3 building block mixer milestones

shows the proposed development schedule for Band 3 SIS building block mixer:

Table 5.4 - Band 3 building block mixer milestones

Finish mixer circuit analysis	2000-11-30
Mask layout	2000-12-21
Mask fabrication	2001-01-15
Junction fabrication by UVA	2001-02-15
Mixer block fabrication	2001-02 -15
Mixer evaluation	2001-03-15

5.6.2.2.3 1/f Gain Fluctuations

Since the ALMA receivers are intended to perform duty both for interferometric and for single-dish total power observations, the radiometric stability of the receivers is important. M. Pospieszalski of NRAO has already developed a wideband 68-116 HFET amplifier with noise performance which nearly meets the ALMA specification. However, based on work by Wollack [9] and Wollack and Pospieszalski [10], it may be calculated that the 1/f gain fluctuation of a single-channel radiometer (no switching) would produce total power fluctuation of about $3 \cdot 10^{-4}$ in one second, exceeding the ALMA receiver specification of $1 \cdot 10^{-4}$ in one second. Preliminary results on a 230 GHz laboratory SIS receiver indicate that it meets the ALMA specification; a detailed investigation is in progress.

5.6.2.2.4 Section References

[1] A. R. Kerr and S.-K. Pan, ADesign of planar image-separating and balanced SIS mixers,@ *Proceedings of the Seventh International Symposium on Space Terahertz Technology,* pp. 207-219, 12-14 March 1996. Available as ALMA Memo151 at http://www.alma.nrao.edu/memos/html-memos/alma151/memo151.pdf

[2] A. R. Kerr, S.-K. Pan, A. W. Lichtenberger, N. Horner, J. E. Effland and K. Crady, "A single-chip balanced SIS mixer for 200-300 GHz," *Proceedings of the 11th International Symposium on Space Terahertz Technology,* May 1-3, 2000. Available as ALMA Memo 308 at http://www.alma.nrao.edu/memos/html-memos/alma308/memo308.pdf

[3] A. R. Kerr, S.-K. Pan and H. G. LeDuc, AAn integrated sideband-separating SIS mixer for 200-280 GHz,@ *Proceedings of the Ninth International Symposium on Space Terahertz Technology,* pp. 215-221, 17-19 March 1998. Available as ALMA Memo 206 at http://www.alma.nrao.edu/memos/html-memos/alma206/memo206.pdf

[4] S. M. X. Claude, C. T. Cunningham, A. R. Kerr and S.-K. Pan, "Design of a sideband-separating balanced SIS mixer based on waveguide hybrids," ALMA Memo 316, available at http://www.alma.nrao.edu/memos/html-memos/alma316/memo316.pdf

[5] A. R. Kerr and S.-K. Pan, "Some recent developments in the design of SIS mixers," *Int. J. Infrared Millimeter Waves,* vol. 11, no. 10, pp. 1169-1187, Oct. 1990.

[6] Q. Ke and M. J. Feldman, "Optimum source conductance for high frequency superconducting quasi-particle receivers," *IEEE Trans. Microwave Theory Tech.,* vol. MTT-41, no. 4, pp. 600-604, April 1993.

[7] A. R. Kerr, S.-K. Pan, A. W. Lichtenberger and H. H. Huang, "A tunerless SIS mixer for 200-280 GHz with low output capacitance and inductance," *Proceedings of the Ninth* International Symposium on Space Terahertz Technology, pp. 195-203, 17-19 March 1998. Available as ALMA Memo 205 at http://www.alma.nrao.edu/memos/html-memos/alma205/memo205.pdf

[8] MMICAD is a microwave integrated circuit analysis and optimization program, and is a product of Optotek, Ltd., Ontario, Canada K2K-2A9.

[9] E. J. Wollack, "High-Electron-Mobility Transistor Gain Stability and it Design Implications for Wide Band Millimeter Wave Receivers", 1995, *Rev. Sci. Instrum.*, vol. 66, no. 8, pp. 4305-4312.

[10] E. J. Wollack and M. W. Pospieszalski, "Characteristics of Broadband InP Millimeter-Wave Amplifiers for Radiometry", 1998, *IEEE MTT-S Digest*, pp. 669-672.

5.6.2.3 Orthomode Transducer for band 3.

As mentioned previously each ALMA band is divided into two channels, each channel responding to a linear polarization with the two polarizations being orthogonal. For the lower frequency bands we are developing waveguide orthomode junctions based on the Biofort junction. Ed Wollack, now at NASA Goddard has pioneered this work (We need references and possibly results here). We now have good results from such an orthomode transducer for band 3 and are working now on a similar design for band 6. An outline drawing of the OMT is given below.

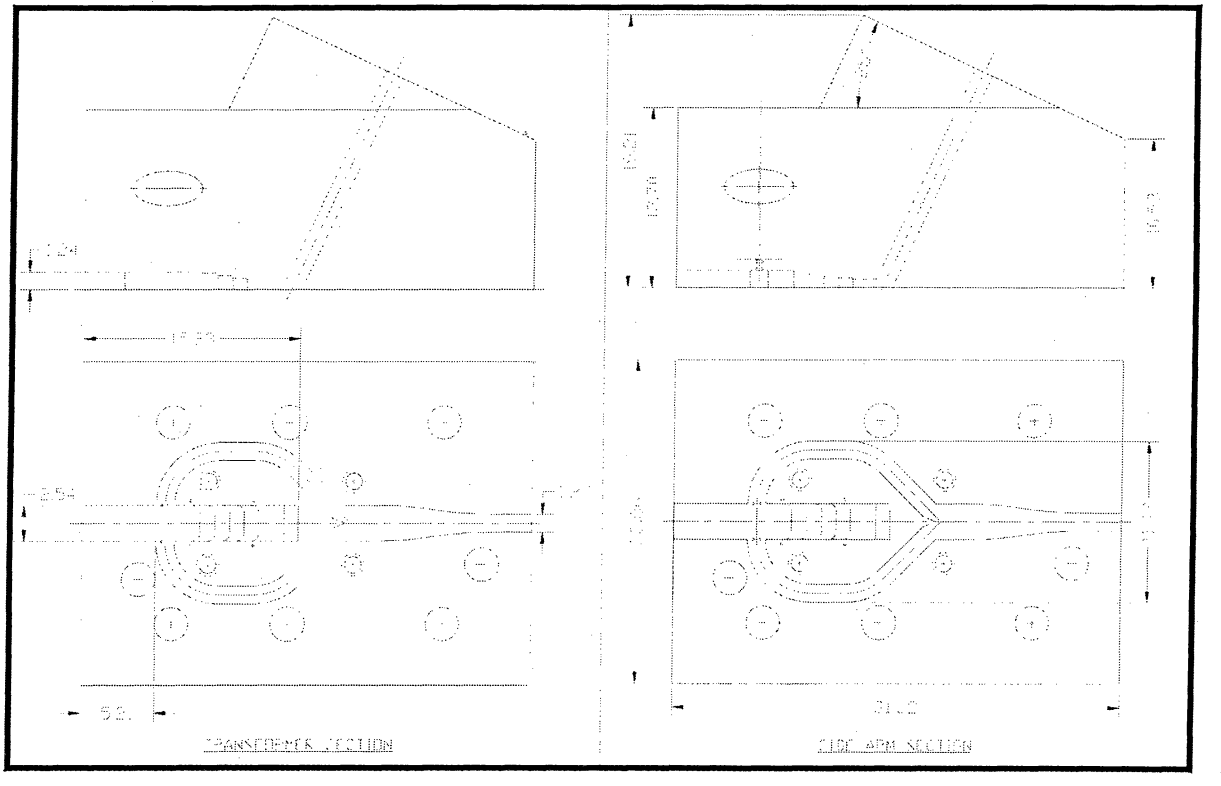


Figure 5.6

5.6.2.4 Band 3 Cartridge outlines.

A preliminary outline of a cartridge design that will satisfy the mechanical dimensions of the present cartridge design is given below.

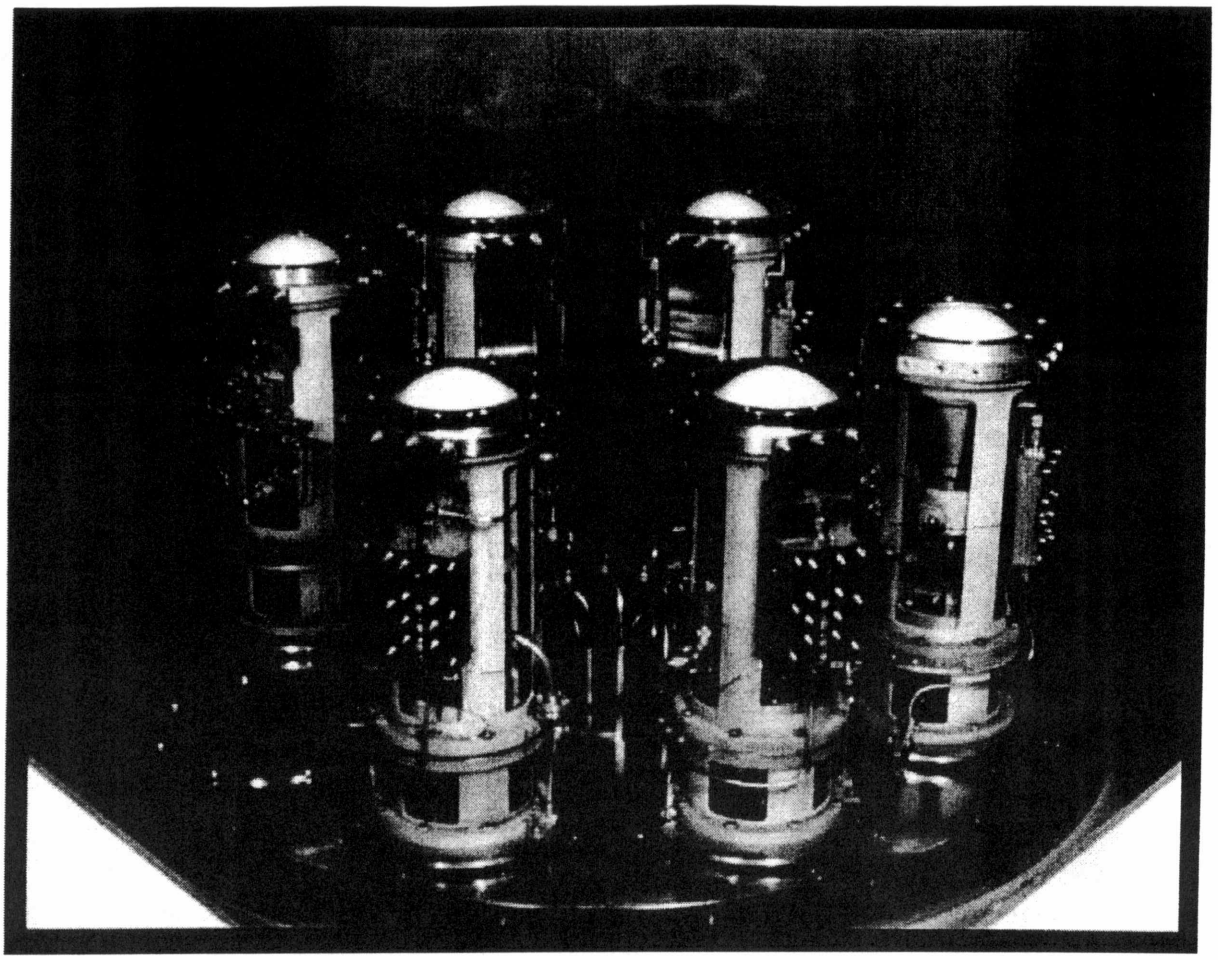


Figure 5.7

5.6.3 Band 6 SIS mixer development at NRAO

Last revised on November, 05 – 2000 by A.R. Kerr, S.-K. Pan and John Webber

Revision History: 2000-10-05: Revised from 1999 MMA version for ALMA.

5.6.3.1 Summary

This section describes the SIS mixers to be used in ALMA front ends for Band 6, 211-275 GHz. The goals for the design and development phase are to produce working prototypes of balanced, sideband-separating mixers with internal IF amplifiers (see section 5.6.4) meeting the general specifications. The goals for the construction phase are to produce large numbers of mixers with repeatable performance at minimum total expense.

Table 5.5 - SIS mixer specifications

Item	Specification
Receiver noise temperature	Noise sufficiently low to produce single sideband receiver noise referred to the vacuum window of 77K over 80% of band, 126K

	at any frequency
Frequency band covered	Band 6, 211-275 GHz
IF bandwidth	Minimum of 8 GHz total, falling in band 4-12 GHz; want 8 GHz for each sideband if possible—otherwise, 4 GHz per sideband is acceptable
Linearity	TBD
Configuration	Balanced operation, sideband separation >10 dB, no mechanical tuners

Table 5.6 - SIS mixer Band 6 milestones

First sideband-separating (SBS), balanced mixer tests	2000-10-30
Integration of SBS, balanced mixer with 4-12 GHz IF amplifiers	2001-03-01
Critical Design Review	2001-04-01
Beginning of production	2001-06-01

5.6.3.2 Performance

Figure 5.8 shows the DSB noise temperatures of SIS receivers reported in the last few years. The best fixed-tuned receivers have DSB noise temperatures in the range 2-4 hν/k up to ~700 GHz. Above ~700 GHz, receiver noise temperatures rise rapidly because of RF loss in the Nb conductors. Work on new materials is likely to improve high frequency results in the next few years (e.g., NbTiN for 700-1200 GHz).

Note that in calculating SSB system noise temperatures from DSB receiver noise temperatures, care must be taken to include the appropriate image input noise. The appropriate value of SSB receiver noise temperature is given (in the absence of window, lens, mirror, and IR filter losses) by:

$$T_{RSSB} = 2T_{RDSB} + T_{image}$$

This formula applies to a SSB receiver composed of a DSB receiver with a sideband separating network at its input. Since T_{RDSB} presumably includes window plus IR filter plus horn loss, that will be included in both signal and image channels, so the value of T_{RSSB} above is pessimistic.

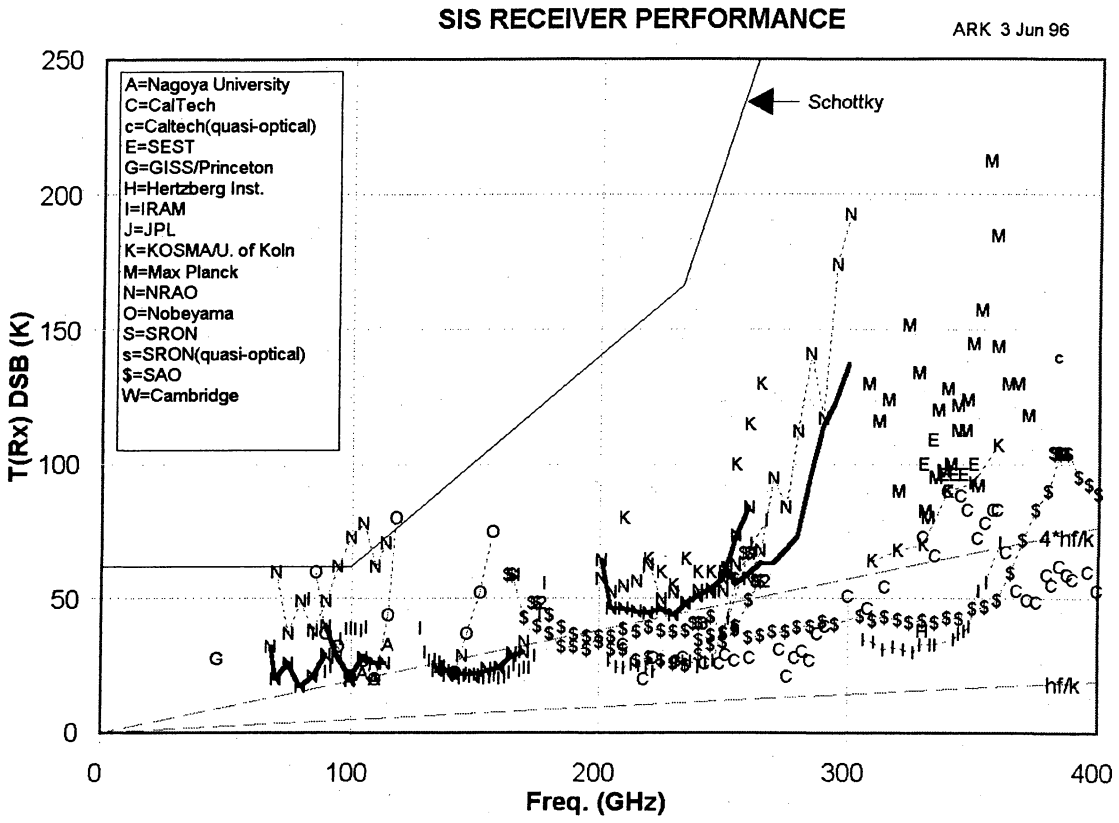


Figure 5.8 - Reported SIS mixer DSB receiver temperatures

Most of these receivers use a ~1.5 GHz IF, an exception being the SAO receivers which use 4-6 GHz. The IF for ALMA is chosen as 4-12 GHz to give the desired 8 GHz IF bandwidth. The best individual tunerless SIS receivers reported to date in the 150-400 GHz range have frequency ranges 1.37:1, 1.42:1, and 1.54:1. Their noise temperatures degrade quite precipitously beyond the band edges. In making the 64 receivers required for each band on ALMA, we cannot expect to achieve identical Tr vs. freq. characteristics, and the maximum bandwidth common to all 80 receivers will be somewhat less than that of the individual receivers. (Nb process control is something we are starting to work on with our SIS fabricators, but hitherto there has been little consideration given to such matters in SIS mixer production). It is hoped that by the time we start building the ALMA receivers we will be able to achieve a 1.5:1 common bandwidth, but until this is actually demonstrated we should be conservative to ensure we do not end up with unexpected gaps in the frequency coverage. This has governed the choice of frequency bands for the SIS receivers.

5.6.3.3 Development

5.6.3.3.1 Capacitively loaded coplanar waveguide

To achieve wide RF bands (an upper to lower frequency ratio of 1.3 or greater) without mechanical tuning, a fully integrated (MMIC) mixer design is desirable. The resulting "drop in" mixer chips are relatively easy to mount in blocks in which they are coupled to RF and LO waveguides. Conventional microstrip MMIC technology is difficult to use above ~100 GHz because of the very thin substrates necessary to prevent coupling to unwanted substrate modes. The use of coplanar waveguide (CPW) circuits allows a thick substrate, but is prone to odd-mode resonances excited by bends or near-by obstacles, and has poor isolation between adjacent lines. CPW also requires inconveniently narrow gaps when a substrate of low dielectric constant is used. To overcome these difficulties, we have developed capacitively loaded coplanar waveguide (CLCPW), a CPW with periodic capacitive bridges. The bridges are grounded at the ends, thus suppressing the odd mode, but they also add a substantial capacitance per unit length to the CPW, which allows desirable characteristic impedance levels to be obtained with convenient dimensions. Figure 5.9 shows a 200-300 GHz quadrature hybrid composed of CLCPW with periodic capacitive bridges.

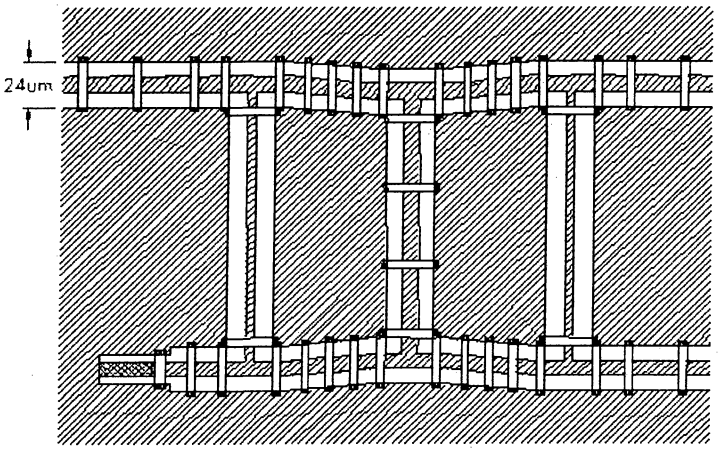


Figure 5.9 - A 200-300 GHz quadrature hybrid using capacitively loaded coplanar waveguide (CLCPW).

The bridges are 4 microns wide, and are connected to the ground plane at their ends. The fourth port (lower left) has a built-in matched termination. The substrate is 0.0035" fused quartz.

5.6.3.3.2 Sideband separating mixer

Even at the proposed site in Chile with its low atmospheric water vapor, atmospheric noise in the image band of an SIS receiver will add substantially to the system noise. The advantages of sideband separating mixers with their image terminated in a 4 K cold load have been discussed (see ALMA Memos 168 and 170), and we expect to use sideband separating mixers in at least the lower frequency SIS receivers. A developmental MMIC 211-275 GHz sideband separating mixer is shown in (Figure 5.10, Figure 5.11, Figure 5.12, Figure 5.13 and Figure 5.14). The IF outputs from the mixer are combined in an external quadrature hybrid which phases the down-converted signals from the upper and lower sidebands so they appear separately at the output ports of the hybrid. A useful property of the sideband separating SIS mixer is that the sidebands can be swapped between the two outputs simply by reversing the polarity of the bias on one of the component mixers.

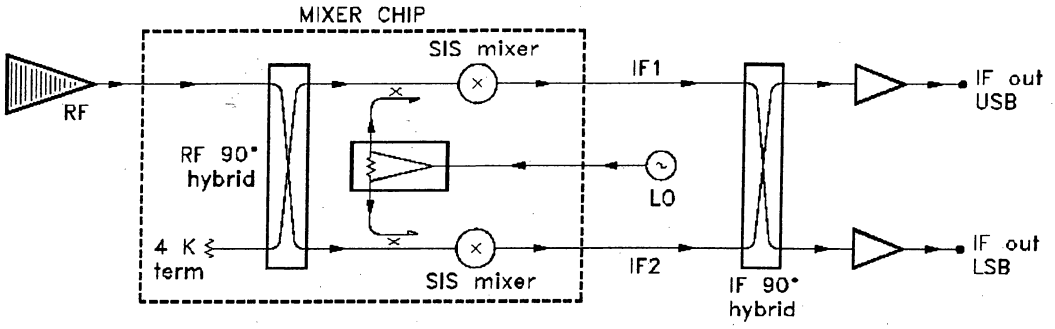


Figure 5.10 - Block diagram of an SIS sideband separating mixer

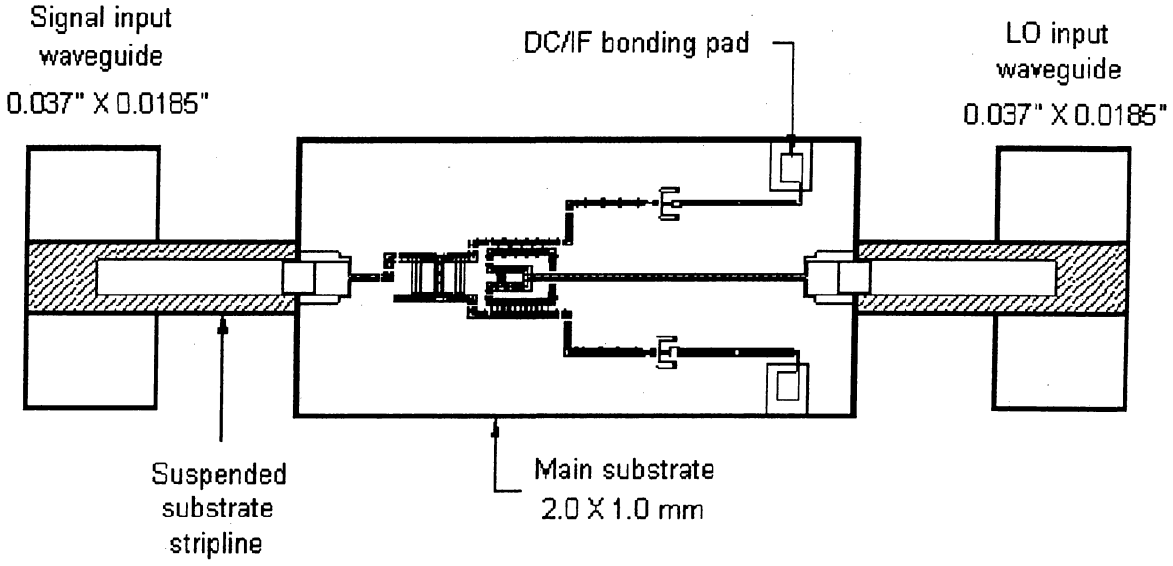


Figure 5.11 - 211-275 GHz sideband separating mixer, showing the signal and LO waveguides, suspended stripline coupling probes, and the main substrate.

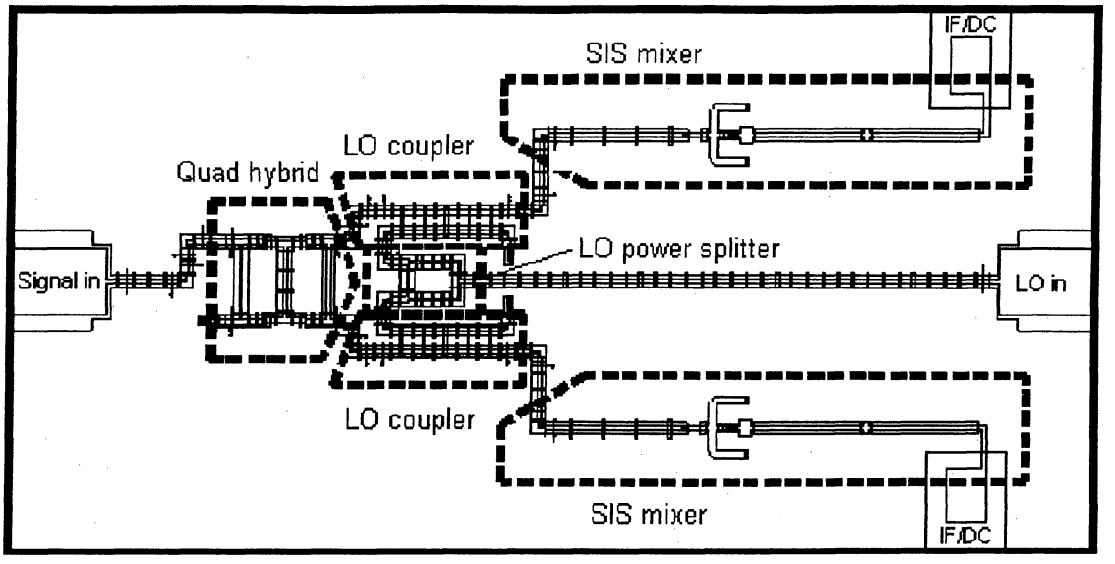


Figure 5.12 - Substrate of the 211-275 GHz sideband separating mixer, showing the main components.

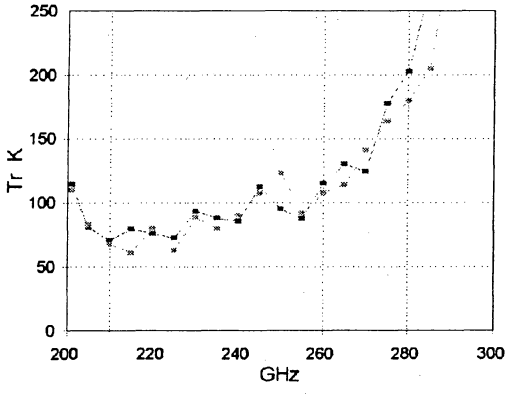


Figure 5.13 - Receiver temperature for the experimental mixer.

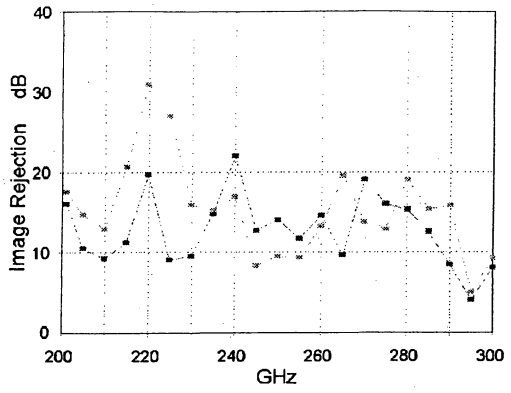


Figure 5.14 - Receiver sideband separation for the experimental mixer.

5.6.3.3.3 Balanced mixer

The use of balanced SIS mixers has two potential advantages for ALMA. Compared with the usual ~20 dB LO coupler or beam splitter in front of the mixer, a balanced mixer requires ~17 dB less LO power. This greatly eases the task of developing wideband tunerless LOs. The other benefit of a balanced mixer is its inherent rejection of AM sideband noise accompanying the LO. A MMIC balanced mixer design is shown in (Figure 5.15, Figure 5.16 and Figure 5.17).

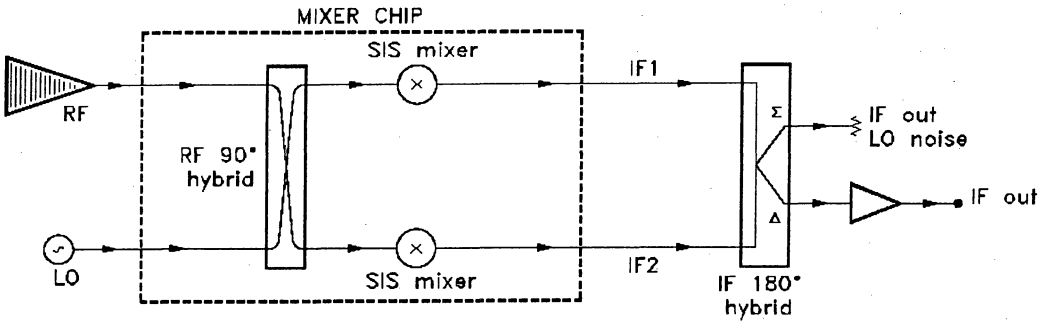


Figure 5.15 - Block diagram of a balanced SIS mixer.

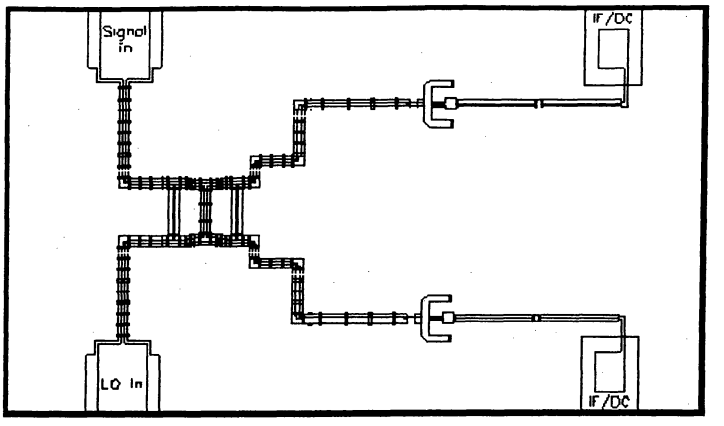


Figure 5.16 - Substrate of a 211-275 GHz balanced mixer, showing the quadrature hybrid and two SIS mixers.

ALMA Memo 308 describes the 211-275 GHz balanced mixer depicted in Figure 5.16. The measured noise temperature is shown *vs.* frequency in Figure 5.17. The first such chip tested was tuned slightly high due to normal variation of wafer parameters, but it exhibits good noise performance and LO noise rejection. The LO noise rejection was >10dB over the tuning range.

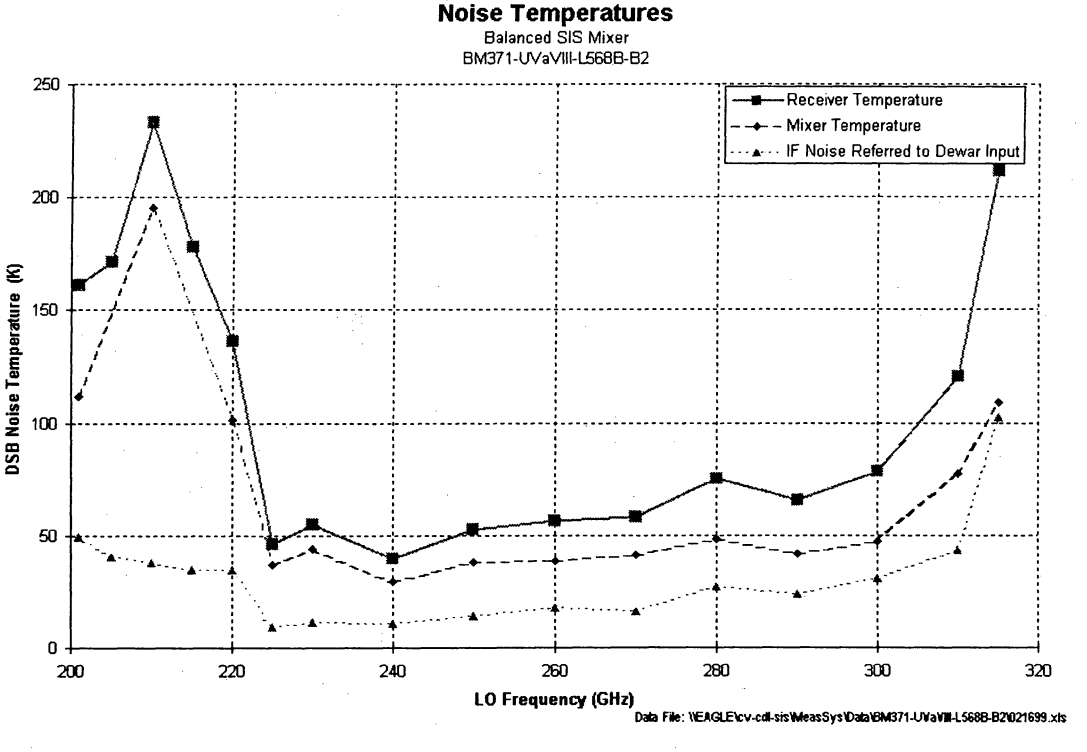


Figure 5.17 - Noise of a balanced SIS mixer.

5.6.3.3.4 Sideband-separating balanced mixers

Now that the designs of the sideband-separating and balanced mixers have been verified, we have designed and expect soon to test a mixer which incorporates both these features: a balanced, sideband-separating mixer. This will incorporate the circuit elements whose design has already been proven individually. This will produce for the MMA a mixer that requires a minimum of LO power, provides good immunity to LO noise, and substantially reduces the contribution to system noise of atmospheric noise in the unwanted sideband.
A photograph of a single MMIC chip is shown in Figure 5.18.

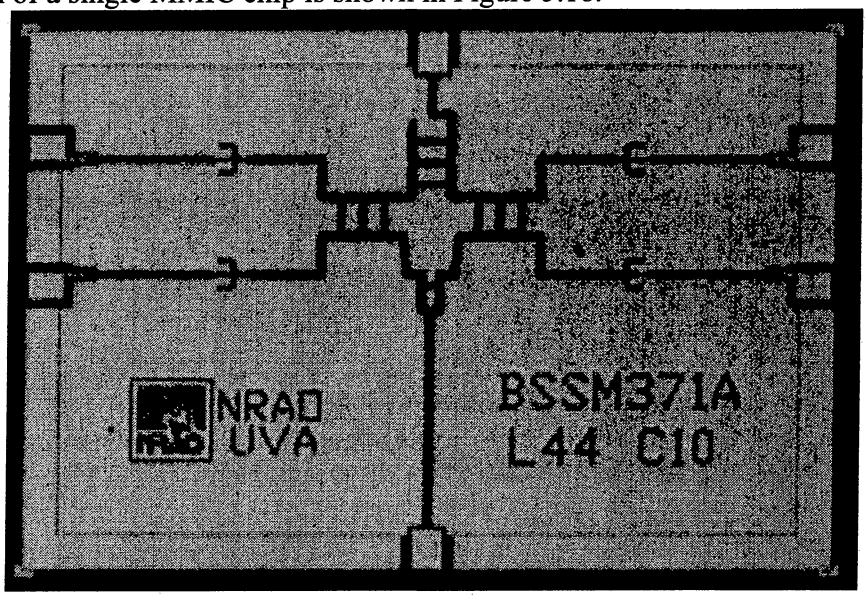


Figure 5.18 - Photograph of a balanced, sideband-separating SIS mixer chip.

5.6.3.3.5 Balanced sideband-separating balanced mixers in waveguide hybrids

An alternate means of achieving balanced, sideband-separating, and balanced sideband-separating operation with SIS mixers is by the use of waveguide hybrids and power splitters with two or four simple DSB mixer chips. The waveguide components can all be machined into a single split-block which also serves as the SIS mixer block. An example of this approach appears in ALMA Memo 316. We have designed and tested such waveguide hybrids in WR-10 waveguide (the highest band for which band we have a vector network analyzer). Figure 5.19 and Figure 5.20 show the computed and measured results for an experimental WR-10 quadrature hybrid. The performance of these experimental structures is satisfactory for use in ALMA receivers, and the required tolerances appear achievable with modern CNC machining techniques for all bands except, possibly, band 10 (787-950 GHz).
This configuration may be preferable to the single-chip balanced sideband-separating mixers in the following circumstances:

1. at the lowest SIS mixer band, for which a completely integrated chip would be relatively large, so a production wafer would contain only a few mixers;
2. at the highest bands, for which ohmic losses in the niobium transmission lines of a single-chip mixer may be too high;
3. if the yield of junctions of acceptable quality were low, so the chance of obtaining four good component mixers on a single chip was reduced to an unacceptable level.

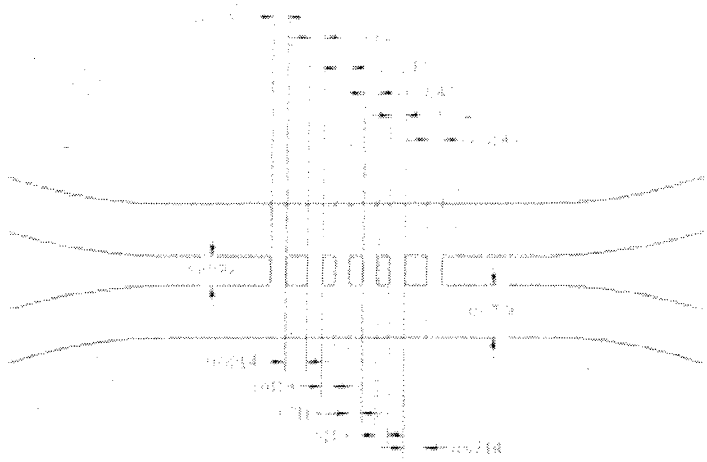


Figure 5.19 - An experimental WR-10 quadrature hybrid.

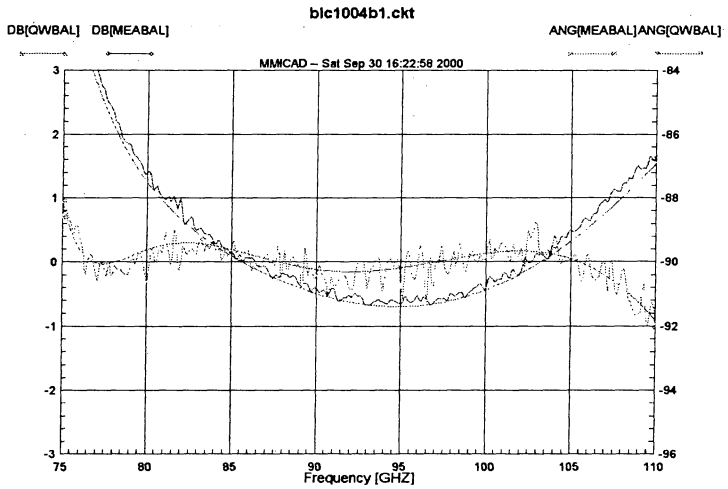


Figure 5.20 - Comparison of a simulation using QuickWave with measured results. The smooth curves are the predictions, and the noisy curves are measured data.

In order to achieve the 8 Ghz bandwidth needed to satisfy the ALMA specifications new techniques are required. One option is to integrate the IF amplifier into the mixer. Work carried out at the CDL in this regard is described below.

5.6.4 Integrated IF Amplifier

Last revised on November, 05 – 2000 by Eugene Lauria, A.R. Kerr, S.-K. Pan, J.C. Webber.

Revision History: 2000-10-05: Revised from 1999 MMA version for ALMA.

5.6.4.1 Introduction

Two options were considered for the 8-GHz-wide IF in the 211-275 GHz SIS receivers for ALMA. The conventional approach uses an IF isolator between the mixer and IF amplifier, while a new scheme, based on earlier work done at OVRO in collaboration with the NRAO, uses an IF amplifier stage inside the SIS mixer block and no isolator. The latter scheme allows an IF covering more than an octave, chosen as 4-12 GHz. The need for an isolator in the conventional scheme would force the IF center frequency to at least 12 GHz (IF = 8-16 GHz) to achieve an 8 GHz bandwidth, probably with a significant noise penalty. The penalty is not simply a result of the increase in amplifier noise temperature at the higher frequency, but includes the noise from the cold termination of the isolator which is reflected from the mixer output.
The use of a high intermediate frequency, as required by both the above schemes, imposes a constraint on the output capacitance and inductance of the SIS mixer. In most SIS mixers, the RF tuning circuit adds substantial IF capacitance in parallel with the SIS junction. We have developed an SIS mixer with low IF capacitance, and this design was used as a building block in the sideband separating and balanced mixers described in section 5.6.3.3.4.

5.6.4.2 Development

In collaboration with M. Pospieszalski of the NRAO Central Development Laboratory, we have developed and interfaced to a 211-275 GHz SIS mixer a 3-stage IF amplifier covering 4-12 GHz (Figure 5.21). This amplifier uses discrete InP HFET transistors to achieve minimum noise and power dissipation, a critical factor in maintaining the SIS junctions at the lowest possible temperature. Due to the high f_T of InP devices, the frequency dependence of their noise parameters is much lower than that of GaAs devices. This is important in order to obtain low noise over broad bandwidths.
Initially, the IF amplifier was optimized for minimum noise with a 50 ohm input load impedance. The SIS mixer is connected to the IF amplifier with a single bond wire and requires no additional matching circuitry. In this particular case, further optimization of the input circuit does not yield any substantial improvement in noise performance over the existing network used for the amplifier by itself. Although this matching network happens to work in this case, it may not work for other mixers. Having the input matching circuit optimized for an input load impedance of 50 ohms makes it handy for testing the amplifier because the mixer block and a type-K connector can be interchanged. To minimize parasitic reactance between the mixer and amplifier, the bias circuit for the mixer is incorporated in the existing amplifier block. This has the added advantage that the amplifier and the mixer bias circuit are tested together which reveals any undesirable interaction between them.

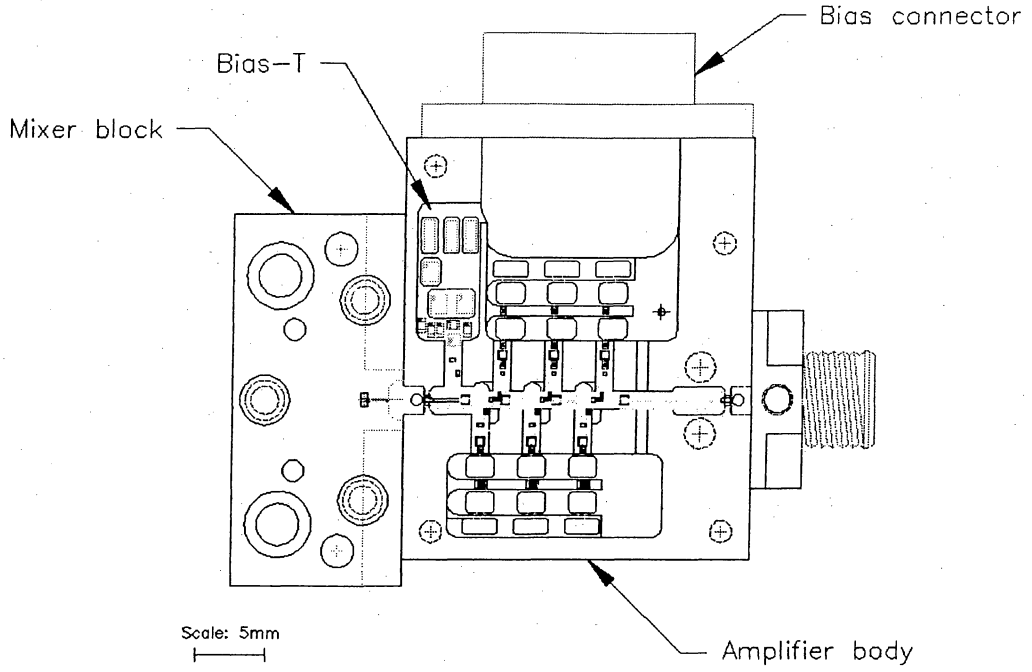


Figure 5.21 - Physical layout of the experimental integrated mixer/amplifier.

Initial experiments have been carried out with a single-ended building-block mixer. The results are shown in Figure 5.22 and Figure 5.23. The performance with the 4-12 GHz IF as a function of RF frequency is essentially the same as for the 1.5-GHz IF chain.

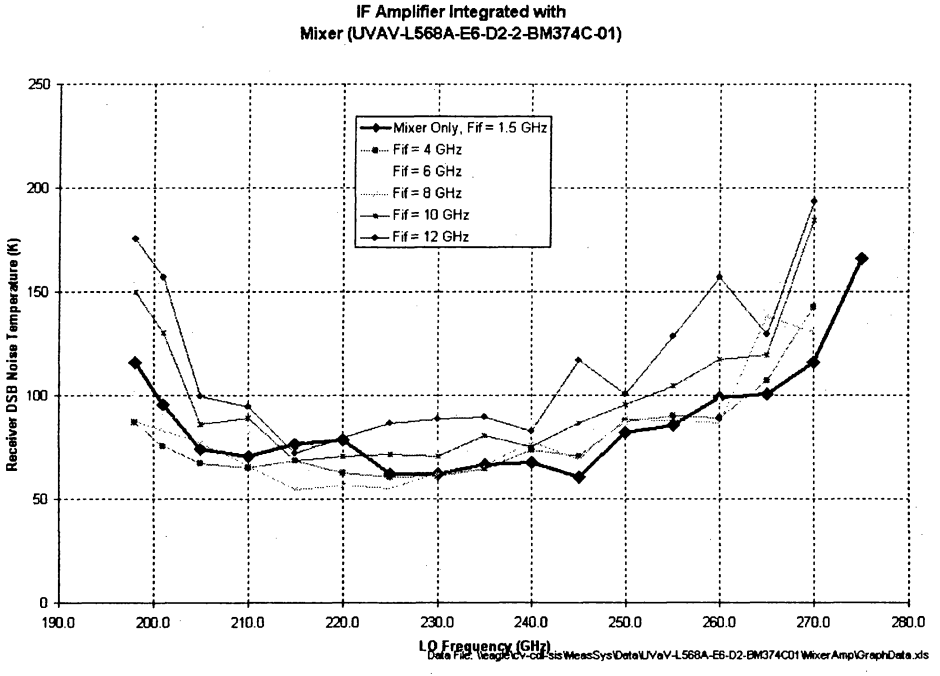


Figure 5.22 - Initial results for a SIS mixer with the experimental IF amplifier, as a function of LO frequency.

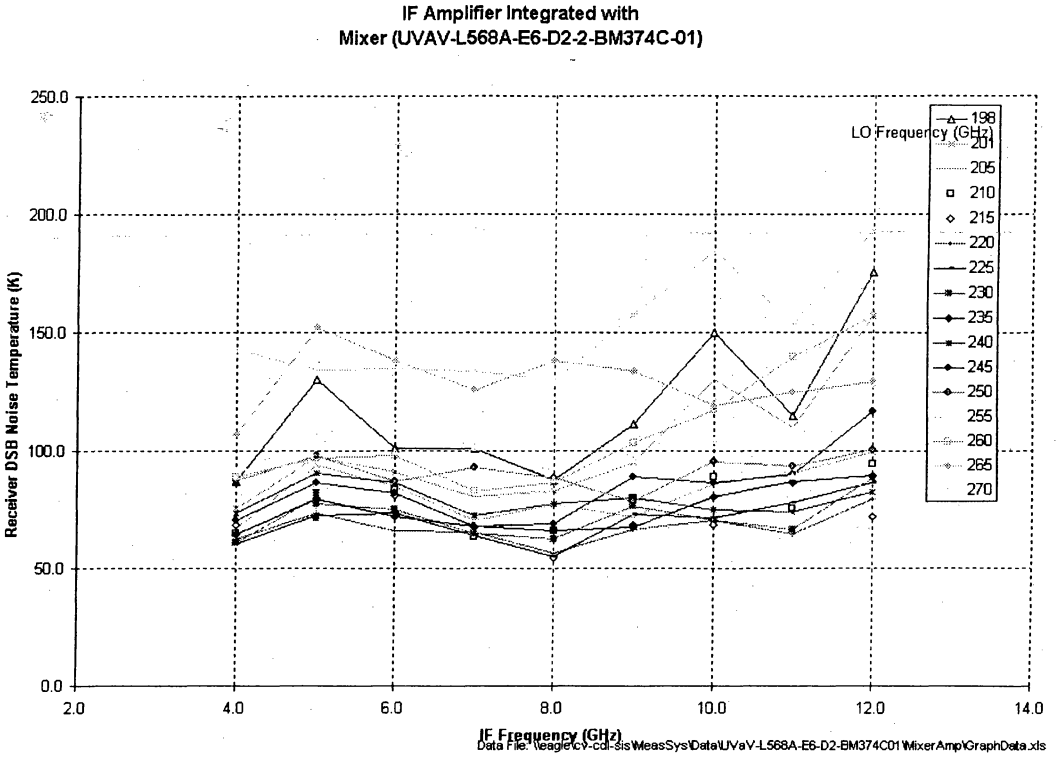


Figure 5.23 - Initial results for the experimental integrated SIS mixer/wideband IF amplifier, as a function of intermediate frequency.

5.6.4.3 Further plans

The next step will be to try different mixers to see how they interact with the amplifier. There may be some cases in which the amplifier will see a negative input load impedance from the mixer. It is not certain how the amplifier will perform if it sees such an impedance. Also, integration of a balanced image-separating sideband mixer will be undertaken. In this configuration, the amplifier input circuit has to allow for two bias-T's for the biasing the two building block mixer junctions of each balanced mixer. Since there are two balanced mixers (four junctions), two amplifiers will be required. The output of these two amplifiers will be combined by a quadrature hybrid which separates the upper and lower sidebands across the IF band.

5.6.5 Band 7 Mixer Development at Onsala Space Observatory, Chalmers University

Last revised on November 23, 2000 by V. Belitsky

Revision history: 2000-10-30: New

5.6.5.1 Introduction

A baseline for *ALMA Band 7 SIS mixer design,* proposed by Onsala Space Observatory, is a sideband separation mixer using quadrature scheme with two identical DSB SIS mixers pumped by a local oscillator (LO) with 90° phase difference. This technology at short mm-waves was pioneered by NRAO and demonstrated at 200-280 GHz band [1]. The main advantage of the sideband separation scheme is that no further tuning is required to provide single side band (SSB) operation compare to other schemes even though fixed-tuned DSB mixers are used. The upper (USB) and the lower sidebands (LSB) are available simultaneously at the two mixer outputs and this relaxes the ALMA requirement of having 8 GHz IF frequency band by a polarization channel, allowing to use a sum of USB and LSB with 4 to 8 GHz IF band for each SIS mixer. The description below outlines the suggested design of the mixer for one polarization channel with assumption of having identical mixer for the second polarization channel.

The block-diagram of a quadrature sideband separation mixer is presented in Figure 5.24: the input RF signal is divided and distributed between the two identical DSB mixers, the LO power is also divided and coupled to the mixers with 90° phase difference. The IF outputs of the two mixers are connected to an IF quadrature hybrid, thus the down-converted USB and LSB signals appear separately at the two output ports of the hybrid.

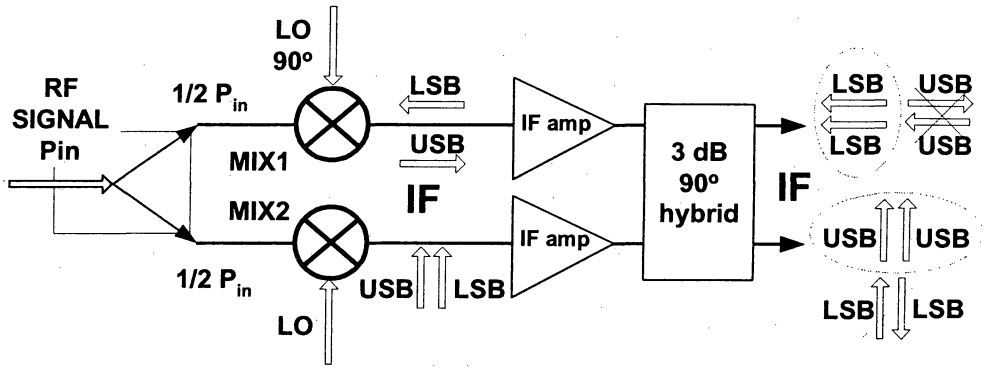


Figure 5.24 - Layout of the sideband separation mixer. The crossed out items at the hybrid outputs are the rejected sidebands (180° phase difference). LSB and USB stand for lower side band respectively.

5.6.5.2 Mixer Block Layout

In the present design we take advantage of a new device, a *double-probe coupler* structure that splits the input RF signal between the two ports, apparently, with minimum losses over a wide frequency band and provides transition from a waveguide to a microstrip line for easy integration of the SIS mixers [2]. In that design the SIS mixers are integrated on the same substrate as the double-probe structure. The layout of the mixer, corresponding to the block-diagram in the Figure 1 and employing the double-probe coupler is presented in the Figure 5.25. It is possible to use the spilt-block technique and CNC machine for the mixer block fabrication, which would ease mass fabrication; both mixers are located on the same substrate providing a high degree of *similarity* in the SIS junction performance and the geometry of all the mixer elements including the transmition lines. Balance between the two mixers is extremely important to keep symmetric

phase and amplitude for the signal and LO and achieve required image band rejection (>10 dB) [3].

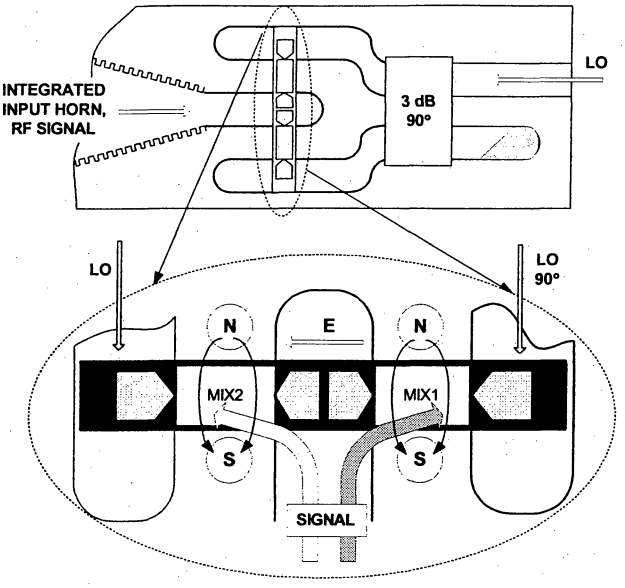


Figure 5.25 - *On the top:* The layout of the sideband separation mixer employing the double-probe coupler; *on the bottom:* the substrate with the two-probe coupler, the two SIS mixers and the single-probe LO injecting feeds.

5.6.5.3 Mixer Chip Layout

The substrate penetrates the three waveguides; the middle waveguide is coupled to the integrated corrugated horn. The two outer waveguides bring the LO signal from the outputs of the 3-dB, 90° branch-line waveguide coupler of a similar type as in [4]. On the chip we place the two mixers with their respective tuning circuits and LO injection coupler with local oscillator guiding circuitry. Figure 5.26 shows schematically layout of the mixer chip.

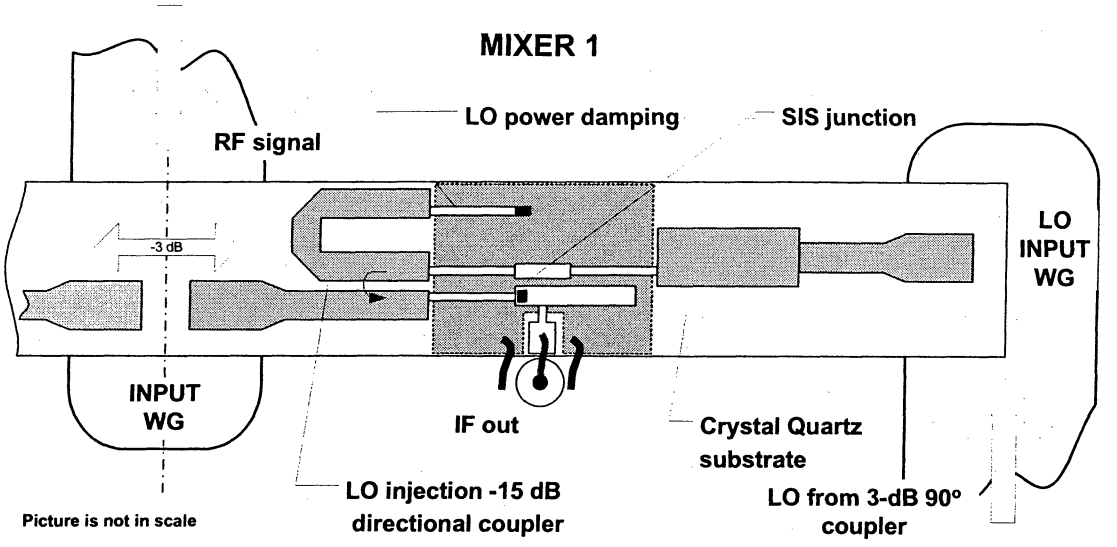


Figure 5.26 - Mixer chip layout: the figure covers area around *Mixer 1* (as in Figure 5.25). SIS mixer tuning circuit consists of an inductive section followed by an open quarter-wave stub. A quarter-wave transformer is coupled from another side for matching of the tuned mixer to the double probe structure connected through the LO injection quarter-wave coupler. All shown lines are microstrip type transmition lines.

5.6.5.4 Mixer Interfaces

5.6.5.4.1 Optics

At the moment of writing these notes the ALMA optical design is still in the discussion stage and no optical interfaces for cartridges have been defined. The cartridge design is pending readiness of the main receiver optics design.

The mixer described above will be fabricated using split-block technique and employing CNC milling machine. Depending on complexity requirements by the optical interface we plan to integrate the scalar horn into the mixer block (fast beam) as it is depicted in Figure 5.25. Alternatively, if the horn should be long (slow beam required) we will use a standard waveguide flange connection between the mixer and the horn.

5.6.5.4.2 LO Feed and LO Power

The LO power required for one polarization channel was calculated as follows: we included in the model 2 SIS junctions (R_n=5 Ω, A=5 μm^2), –15 dB for the LO injection via the coupler, frequency dependent loss in the transmition lines on the substrate and waveguides, 2 dB ripple in both couplers (on the substrate and WG 3-dB coupler) and margin 3 dB. Calculations of the LO power are made following model suggested in ALMA MEMO 264; additionally we added a provisional dependence of the SIS power coupling with RF integrated tuning.

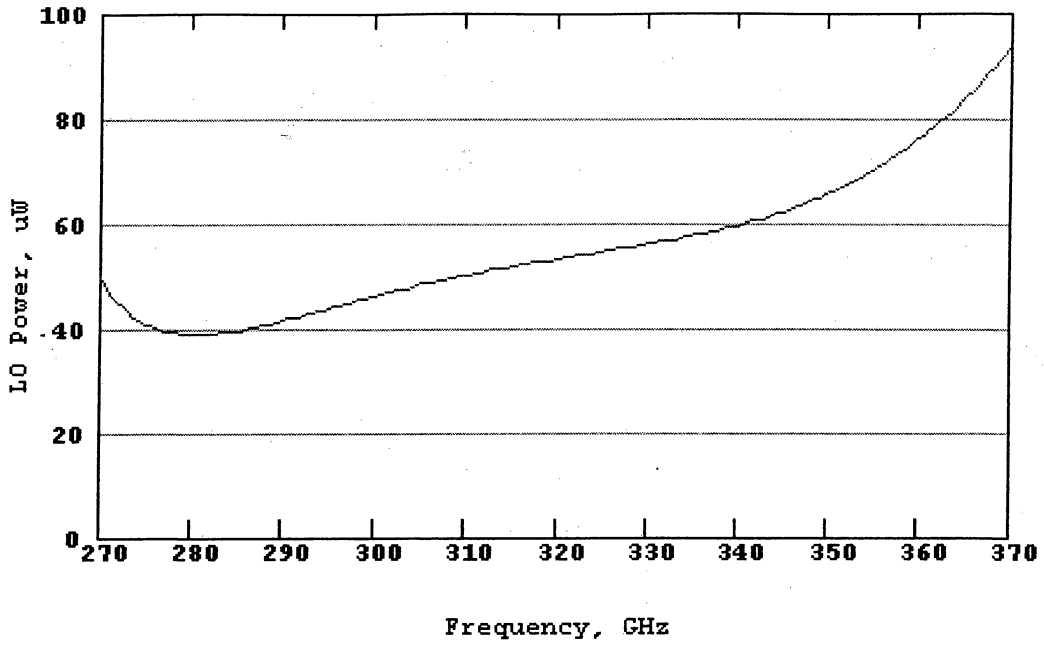


Figure 5.27 - LO power required by the sideband-separating mixer for one polarization channel.

For the mixer described here the LO interface would be just a waveguide feed connected to the input of the integrated 3-dB 90° coupler as in Figure 5.25. The two mixers fabricated on the same substrate are matched with respect to the required LO and we do not consider at the moment any individual LO level adjustment inside the mixer. We expect though that the LO distribution circuitry for different polarizations will have a ***balance attenuator*** to allow different level of LO power between the polarization channels (LO injection scheme pending finalizing of the optical interfaces and the cartridge design). As a result, the total power for the two-polarization system would be somewhat more than 2 times higher (insertion loss) than the one depicted in Figure 5.27 .

5.6.5.4.3 Intermediate Frequency

The mixer will use two cryogenic IF amplifiers connected to the mixers with isolator and coupled to 3-dB 90-degree coupler at the output. IF block diagram is shown in Figure 5.24. ALMA IF band is 4 – 8 GHz and two options for IF amplifiers are considered: ***i.*** integrated amplifier with direct connection to the SIS mixers (no isolators, bias tee integrated into the first stage of the amplifier); ***ii.*** IF amplifier based on discrete components with an isolator at the input. The first option is under development at NRAO, Charlottesville. Onsala group, in collaboration with Microwave Technology Dept., Chalmers University, works on the developing of a low-noise cryogenic amplifier based on GaAs HEMT transistors for 4 – 8 GHz band (later plan to go for InP HEMT). The amplifiers should be matched in phase and gain to achieve required sideband rejection. Onsala design considers built-in adjustment of the amplifier gain to equalize overall gain (including mixers) in both channels of the mixer.

5.6.5.4.4 DC Bias, Magnetic Field, Heater and Temperature

DC bias uses circuit with a shunt resistor. The circuit similar to the presented in Figure 5.28 is in use at Onsala SIS mixers (and many other places) and has a number of advantages, including protection against static discharges (SIS junction is always connected to the ground via the shunt resistor (2 resistors of 10 **Ω** in series in our case).

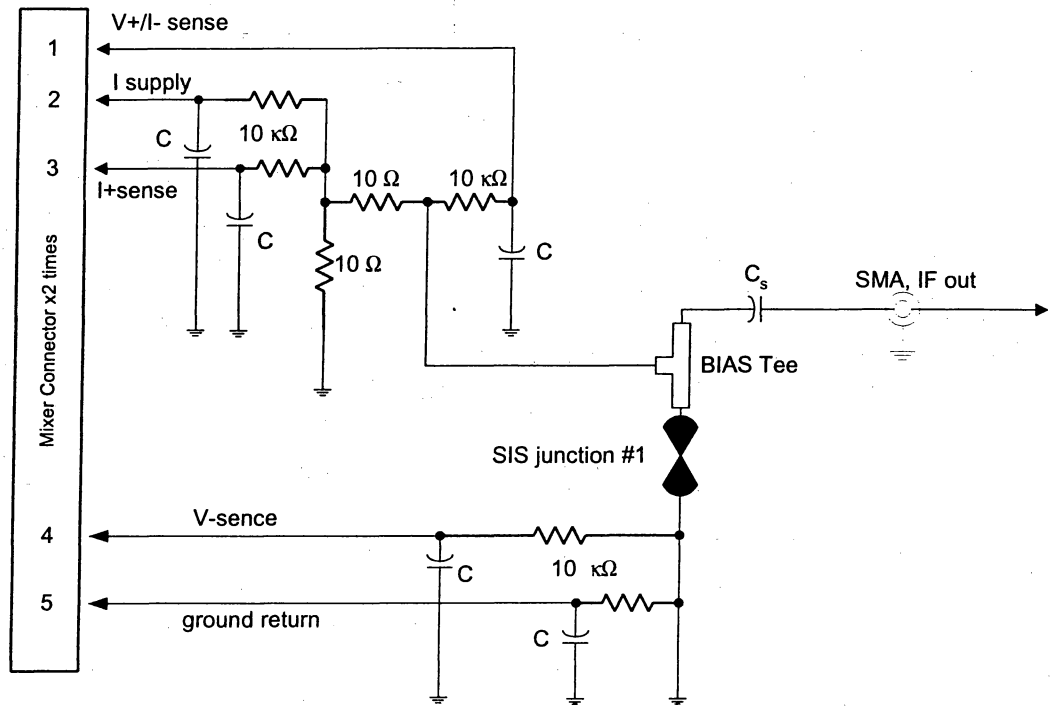


Figure 5.28 - DC bias circuit diagram with 10 Ω shunt resistor. Normally, all the shown components, chip resistors and capacitors, are integrated into the mixer block. Depending on the type of IF amplifier used, the bias tee would be integrated into the amplifier instead.

Both passive bias voltage stabilization (stable voltage between the contacts #2 and #5 as in Figure 5.28) and active voltage/current source with feedback are possible to use with the circuitry above. To avoid problem with ground loops we suggest using floating DC bias with the only grounding point at SIS junction mounting in the mixer block (separate ground return). Each of the two SIS junctions will require 5 wires for DC bias, total 10 wires.

Magnetic field to suppress Josephson current will be generated by the two separate coils dedicated for each of the two SIS junctions (Figure 5.25 shows position of magnetic poles around the two SIS mixers). We plan to use magnetic field guiding (high μ metal) to minimize required currents. The coils should be fed by a separate driving electronics (current stabilization, two-wire circuitry) to provide individual tuning for the two SIS mixers, preferably with floating power supply to avoid ground loop problem. Magnetic field will require in total 4 wires.
In order to better control Josephson current via applying of the magnetic field we need also control over magnetic fields while the system is cooling down and possibly "frozen-in" or trapped fluxes of the magnetic field could be a potential problem. The fluxes may also appear as a result of abrupt change of the DC bias current or using electro powered instruments nearby the

receiver. Temporary warming up of the mixer above the temperature of the superconducting transition, $T_c \sim 9.2$ K for Nb film would allow us to remove the frozen fluxes. We suggest including a heater for each mixer to control frozen fluxes and simplify suppression of the Josephson current. The two-wire circuitry with current stabilization, floating power supply, total 2 wires per mixer block.

Mixer ambient temperature information is very essential for understanding and solving possible problems during SIS mixer operation. We suggest installation of a temperature sensor on every mixer block. If the IF amplifiers employing circulators would be used, we suggest to monitor the temperature of the termination load installed on the circulators. Standard Lake-Shore temperature sensors or simular, 2 wires per sensor.

5.6.5.5 References

[1] A. R. Kerr, S.-K. Pan and H. G. LeDuc, "*An integrated sideband separating SIS mixer for 200-280 GHz*", Proc. of the Ninth Space Terahertz Technology Symposium, Pasadena, USA, March, 1998.

[2] V. Vassilev, V. Belitsky and R. Booth, "*A New Sideband Separation SIS Mixer for ALMA*" Proc. of SPIE, volume 4015, March 2000. Can be obtained via http://gard04.mc2.chalmers.se/papers/SPIE_2000.pdf

[3] A. R. Kerr, S.-K. Pan, A. W. Lichtenberger, N. Horner, J. E. Effland, and K. Crady, "A Single-Chip Balanced SIS Mixer For 200-300 GHz ", ALMA Memo Series, *Memo 308*, http://www.mma.nrao.edu/memos/html-memos/abstracts/abs308.html

[4] S. M. X. Claude, C. T. Cunningham, A. R. Kerr, and S.-K. Pan "*Design of a Sideband-Separating Balanced SIS Mixer Based on Waveguide Hybrids*", ALMA Memo Series, *Memo 316,* http://www.alma.nrao.edu/memos/html-memos/abstracts/abs316.html

5.6.6 Band 7 Mixer and Cartridge Development at IRAM

Last revised on 27 Nov 2000 by S. Claude, IRAM

Revision history: 2000-11-23 first version

5.6.6.1 Summary

We describe here IRAM's development for the realisation of Band 7 cartridge. Our baseline design involves two DSB mixers for each orthogonal polarization with waveguide couplers LO injection. The polarizations are separated by one grid. The 17 dB crossguide coupler allows LO injection in a compact configuration. The mixers to be used in the cartridge are DSB, fixed tuned across the RF frequency range (275-370 GHz) and low noise for an IF covering 4 to 8 GHz. Future development involves integrated sideband separation mixers in waveguide.

5.6.6.2 *Cartridge layout and optics*

The cartridge contains three cold plates, one at 4K for the mixers, mirrors and grid, one at 15 K for the IF LNA and one at 70 K for the input LO multiplier. The baseline system involves two orthogonal polarizations with two DSB (double side band) single-ended mixers. Fig. 1 shows the layout of the components in the cartridge with the input optics.

A more precise description of the optics layout will be given in the chapter on the receiver optics, and we give here only a brief description (Fig. 2). The telescope beam waist is at the cryostat top plate, which is situated 130mm off the cryostat axis. The beam enters into the cryostat and is reflected at an angle of 116° on to an elliptical mirror. A reflection of 26° gives an intermediary waist of 2.4mm, where a compact polarization grid is placed. The two orthogonal polarizations are then directed onto two elliptical mirrors of reflection angle of 26° and then reflected onto flat mirrors into the corrugated feed horns.

The reasons for the added complexity of the optics were to ease the layout for the mixers and the local oscillator injection. Moreover, the low angles of reflections on the input elliptical mirror will reduce any crosspolar. With this scheme we have two signals of orthogonal polarizations coming into parallel waveguide paths of the same orientation, which allows a series of upgrades for the mixers without modification of the optics.

The LO is injected in both mixers from a multiplier on the 70 K stage via two crossguide couplers in series. A magnetic coil is attached to each mixer to suppress the Josephson current. Finally the output IF signals are amplified by cooled HEMT amplifiers at 15 K.

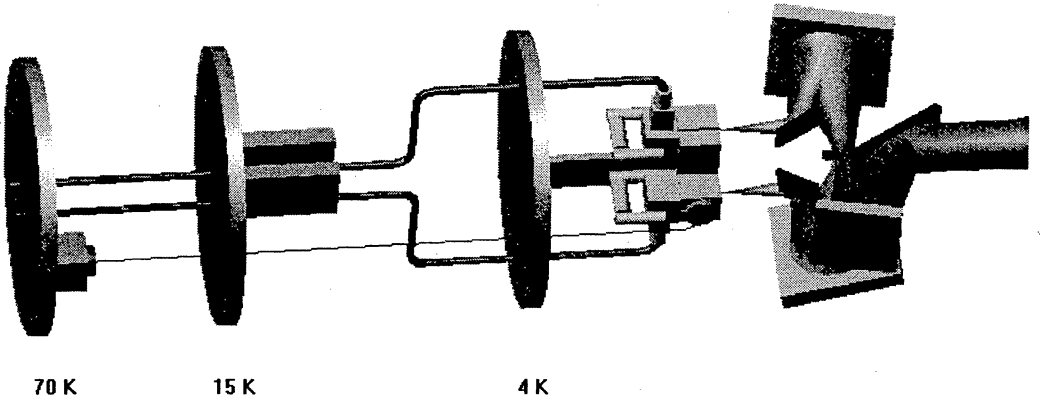


Figure 5.29 - Layout of the main components in the cartridge

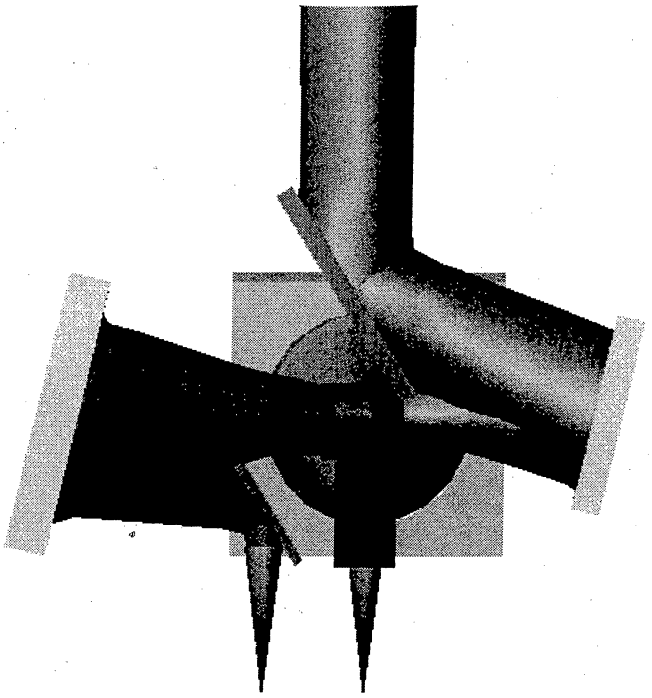


Figure 5.30 - Close up view of the optics

5.6.6.3 Component development

5.6.6.3.1 LO injection: a compact crossguide coupler

A compact crossed guide coupler for the frequency range of 275 to 370 GHz has been developed for the injection of the LO signal. The design is shown in Fig. 3.

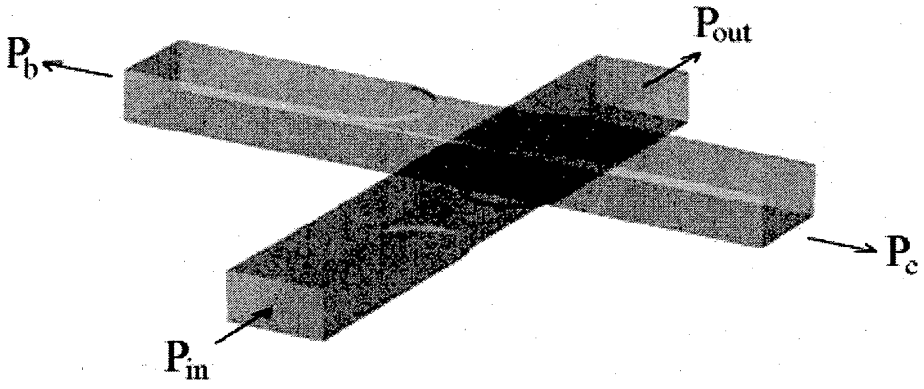


Figure 5.31 - Schematic view of the crossed guide coupler with Pc and Pb corresponding to forward coupled power and backward coupled power, respectively.

Coupling is achieved via two round holes. Simulations carried out with an electromagnetic simulation software (CST Microwave studio) indicates that this simple type of coupler achieves a coupling around 16 dB varying only by 1 dB over the whole frequency range and a directivity of about 10 dB.
Since the performance of our VNA is better in the frequency band around 230 GHz, the design was scaled down to this frequency range for the fabrication of a prototype. Results of simulation and measurements of this prototype are shown in Fig. 4. The directivity is not very good, but that should not be a problem since the input match of the mixer is not expected to be very good.

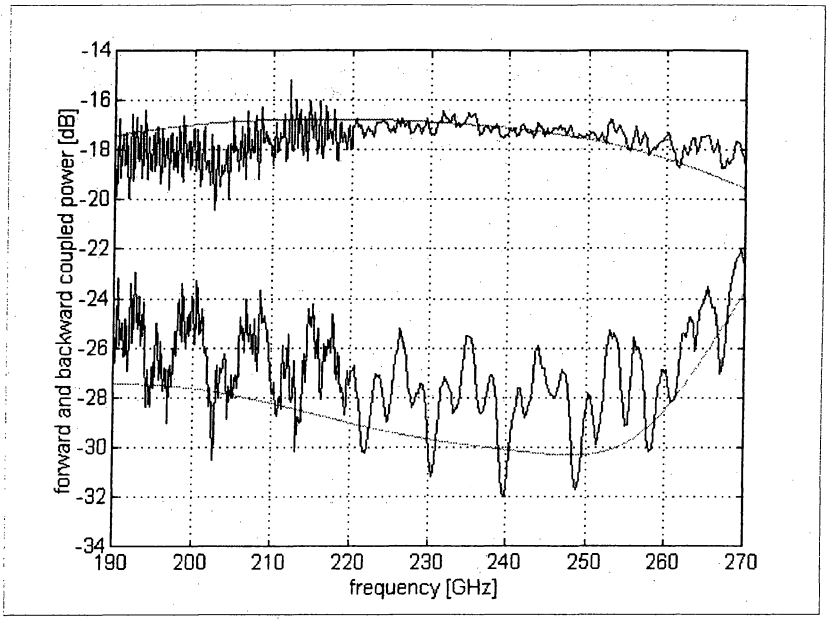


Figure 5.32 - Simulation and measurements of the crossed guide coupler, scaled to band 6

5.6.6.3.2 Mixer baseline design

A full height waveguide SIS mixer covering the 275-370 GHz frequency band has been designed. The fixed tuned single junction Nb/Al-AlOx/Al mixer will operate in Double Side Band. A ~30 % operating bandwidth can be achieved by using an "end-loaded" tuning stub to tune out the junction capacitance of 75 fF (junction size 1 μm2) followed by two quarter-wave transformer sections. All the transmission lines integrated in the mixer chip are implemented in superconducting microstrip with the exception of a section of the quarter-wave transformer, which is realized as a Capacitively Loaded Coplanar Waveguide (CLCPW). The junction is mounted on an 80 μm thick quartz that stretches only part way across the waveguide. Fig. 5 shows a three-dimensional view of the mixer including the full height waveguide to suspended microstrip transition, the low pass "hammer" type filter and the antenna probe.

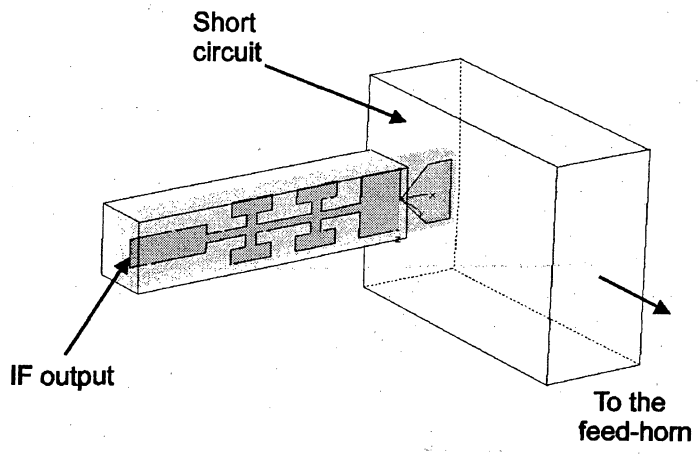


Figure 5.33 - View of the mixer substrate and the input waveguide

A detail of the mixer chip is illustrated in Fig. 6, which includes the SIS junction and its integrated matching structure. In Fig. 7, the simulated results for the embedding impedance seen by the junction are displayed as a function of frequency. The SSB noise temperature of the receiver Trec consisting of the mixer cascaded with a LNA operating at a central IF frequency of 6 GHz (TIF= 6 K is assumed) has been calculated from the complete quantum mechanical treatment. In Fig. 8, the expected value of Trec referred to the mixer input is plotted as a function of IF frequency for three different RF frequencies. SSB receiver noise temperature in the range 23-35 K is expected in the 275-370 GHz frequency band.

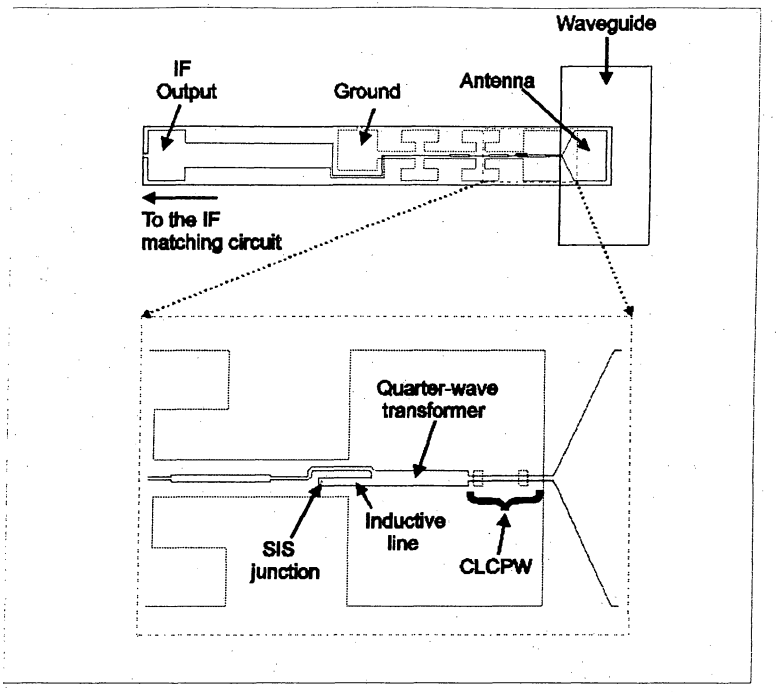


Figure 5.34 - Mixer chip layout

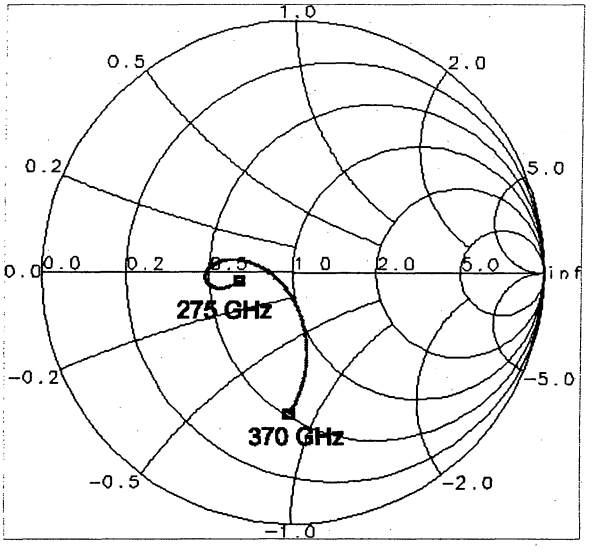


Figure 5.35 - Simulated embedding impedance of the junction in the mixer block, normalized to the RF impedance of the junction (18.7O)

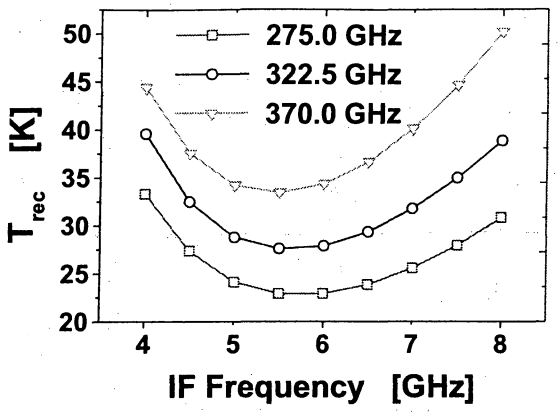


Figure 5.36 - Expected SSB receiver noise temperature referred to the input of the mixer for three RF frequencies, across the IF band.

5.6.6.3.3 Mixer future developments

In parallel with our baseline DSB solution, we are developing a 2SB (sideband-separating) mixer. The mixer integrates an input quadrature hybrid in waveguide, an in-phase LO splitter, 2 cross-guide couplers for the LO injection and 2 single-ended mixers as described in Fig. 9. Provision has been made so that the layout in the cartridge and the input optics would allow the integration of one 2SB mixer for each polarization.

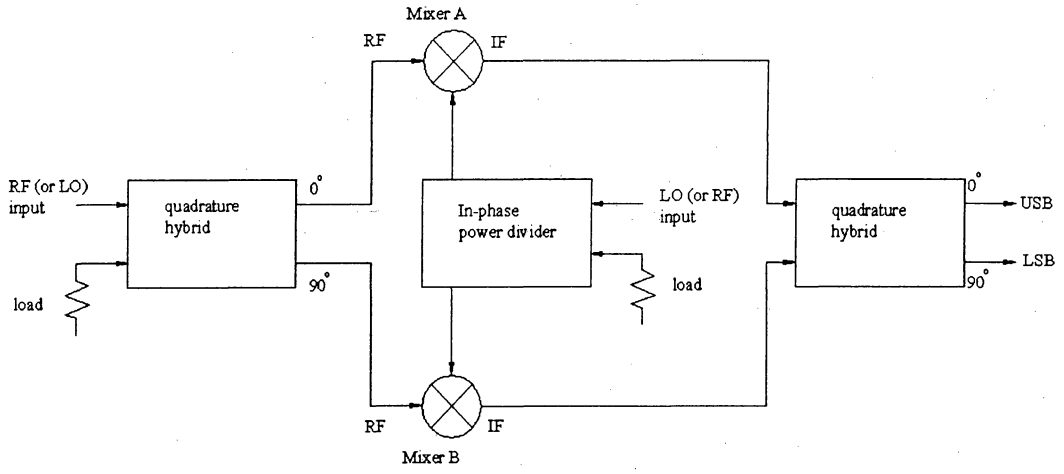
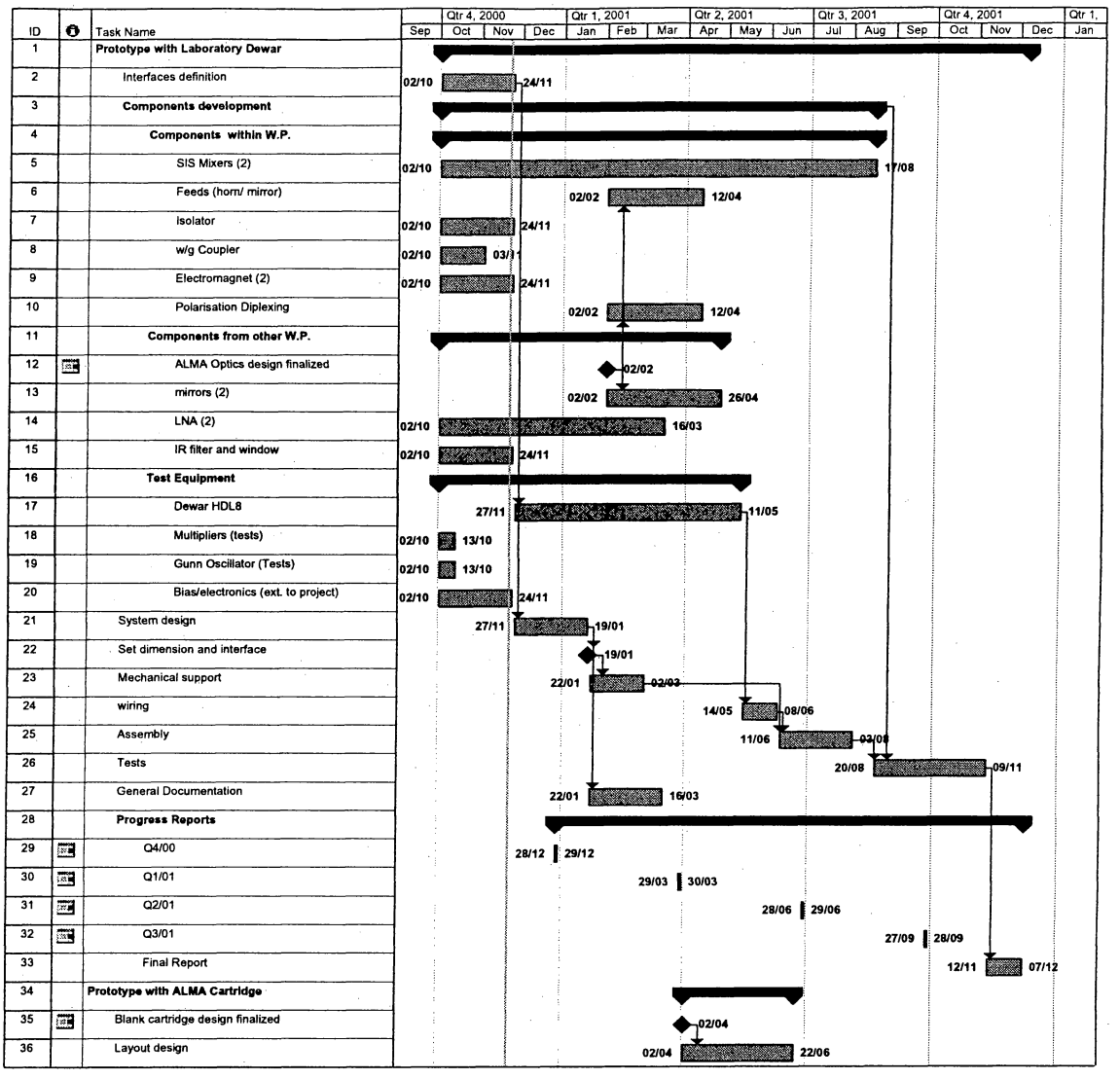


Figure 5.37 - Schematic diagram of the 2SB mixer

5.6.6.4 Timeline



5.6.7 Band 9 SIS mixer development at NOVA/SRON

A. Baryshev[1], H. Schaeffer[1], W. Wild[1], T. Klapwijk[2], T. Zijlstra[2], and R. Hesper[1]
[1] NOVA/SRON, Groningen, the Netherlands
[2] DIMES, Delft, the Netherlands

Revision History: 2000-11-22: first version

5.6.7.1 Summary

This section describes the SIS mixers for ALMA band 9, 602 to 720 GHz, developed at SRON and funded by NOVA. Starting from a design of a single-ended fixed tuned SIS mixer for the JCMT D-band (625 to 710 GHz) and a quasi-optical SIS mixer for HIFI (800-1050 GHz), we are developing balanced mixers both in waveguide design and quasi-optical design for ALMA band 9. The two quite different design approaches have been chosen in order to assess the advantages and disadvantages of each design in terms of performance, ease and cost of manufacturing and assembly. The goals for the design and development phase are to produce prototypes of each design. The goals for the construction phase are to produce large numbers of mixers of the chosen design with repeatable performance at minimum total expense.

5.6.7.2 SIS Mixer Specifications and Development Schedule

Table 5.7 shows some of the SIS mixer specifications. Although the ALMA front end specifications call for a 8 GHz IF bandwidth, this issue can only be assessed in detail after having available experimental results which prove that this wide IF bandwidth is indeed the best choice for the science to be done with ALMA. The 8 GHz IF bandwidth is scientifically driven by the desire to have maximum continuum sensitivity. However, to ensure this scientific goal, the 8 GHz bandwidth needs to be achieved without increase of receiver noise temperature and without decrease of receiver stability as compared to a lower IF bandwidth.. Intense development work to make an 8 GHz IF integrated amplifier available is being carried out at NRAO. SRON will integrate such an amplifier (produced at NRAO) into the mixer designs for band 9.

Table 5.7 SIS mixer specifications

Item	Specification
Receiver noise temperature	Noise sufficiently low to produce double sideband receiver noise (referred to the vacuum window) of 168 K over 80% of band, 250 K at any frequency
Frequency band covered	Band 9, 602 – 720 GHz
IF bandwidth	8 GHz, falling in band 4-12 GHz
Configuration	Balanced or single ended DSB operation, no mechanical tuners, waveguide or quasi-optical beam coupling

We intend to carry out the first tests of a balanced mixer in April 2001 and integrate the 4 – 12 GHz IF amplifier from NRAO into the mixer by June 2001. A front end CDR is planned for end of 2001.

5.6.7.3 Balanced waveguide SIS mixer

As already mention in Section 5.6.3.3.3 the use of balanced SIS mixers has two potential advantages for ALMA. One is the lower LO power requirement as compared to single-ended mixers (typically –17 dB), the other is the inherent rejection of AM sideband noise accompanying the LO. For the development of a balanced waveguide mixer, we start from a proven design of a single-ended fixed tuned mixer for the 650 GHz band (developed at SRON by H. van de Stadt, H. Schaeffer, J. R. Gao, L. de Jong, and W. Laauwen). Details of this design are given in Section 5.6.7.4.8. The balanced waveguide mixer will use similar end pieces (junction holders) and SIS junctions as the single-ended design, which have demonstrated a large *rf* bandwidth (on the order of 150 GHz) and low noise. These parts will be optimized for the required *rf* bandwidth and receiver noise of ALMA band 9. An advantage of this design is its simplicity and potential suitability for series production.

The balanced waveguide mixer basically consists of a magic-T with two integrated horns (one for the rf coupling and one for the LO coupling) and two junction back pieces. The principle is shown in Figure 5.38. The *rf* signal is coupled in-phase and the LO is coupled in anti-phase to MIXER1 and MIXER2, respectively. The IF output of the two mixers can be combined if the SIS junctions are biased in opposite directions. The magic-T will be fabricated in split block technique. For ease of manufacturing we chose to start with a diagonal feed horn. It is straightforward to change it to a corrugated feed horn since the horn is inserted into the magic-T block. The Magic-T has larger dimensions (and consequently simpler machining) as compared to other hybrid structures for the same band.

A magnetic field which is needed for suppressing the Josephson noise, is supplied to each junction back piece individually. This allows to compensate for a possible spread in production parameters. Figure 5.39 and Figure 5.40 show the basic design. We expect a performance similar or better to the fixed tuned JCMT D-band mixer (see Section 5.6.7.4.8).

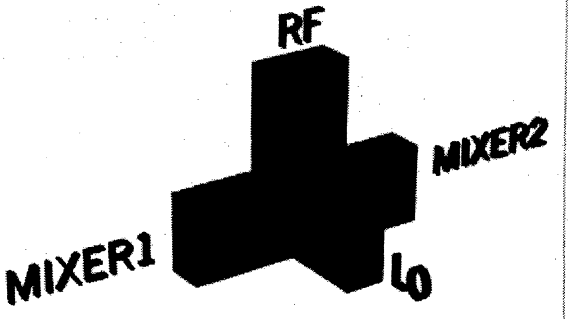


Figure 5.38 - Principle of a magic-T as waveguide hybrid for a balanced mixer.

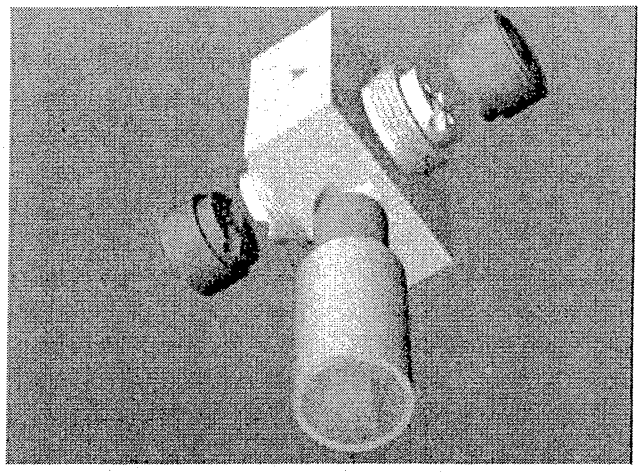


Figure 5.39 - Balanced waveguide mixer design. Clearly visible are the *rf* feed horn (here a diagonal horn) and the LO feed horn on the top of the block. The SIS junctions of the two mixers are mounted in the round end pieces.

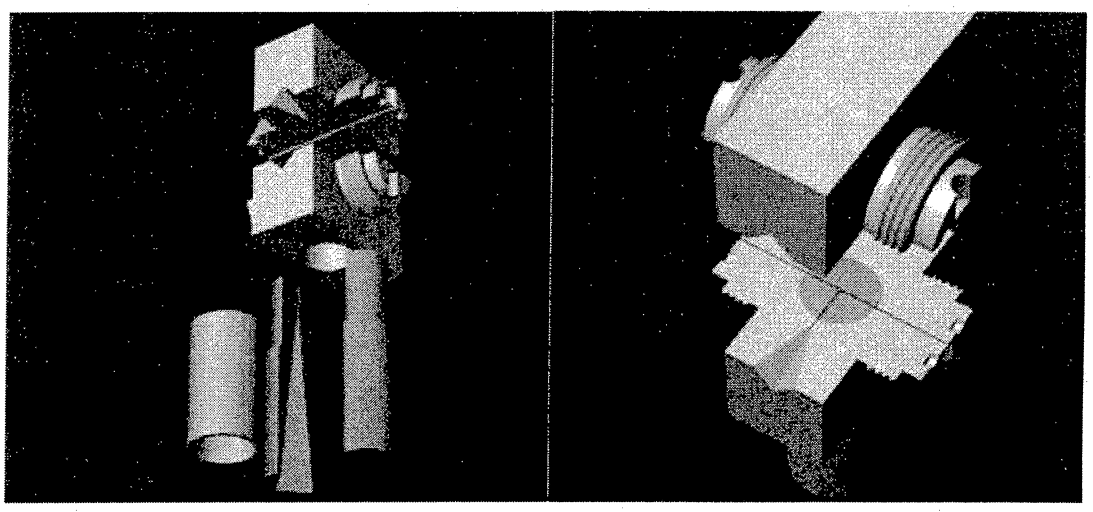


Figure 5.40 - Left: The balanced mixer consists of an inserted horn, magic-T and two end pieces (not shown here). Right: View of the front part of the magic-T with inserted *rf* feed horn and part of the LO feed horn.

5.6.7.4 Quasi-optical balanced SIS mixer

In parallel with the waveguide design, we also develop a quasi-optical SIS mixer for band 9. A quasi-optical mixer has some potential advantages over a waveguide mixer. These mixers are produced with optical lithography which allows to reproduce antenna dimensions with high accuracy. The lens can be produced quickly and in greater quantities (about 200 pcs a day). The estimated cost of the lens is much less than the cost of a corrugated horn for these frequencies. The chip is made of silicon and does not require polishing. The lens can readily produce the beam with an F-number matching the telescope beam without any intermediate optics. In the balanced mixer configuration the LO can be injected in orthogonal polarization with respect to

the *rf* signal. That allows to use only one grid for LO injection and polarization separation. Disadvantages of the quasi-optical design include the difficulty to achieve high coupling efficiency to a telescope.

Figure 5.41 shows an example of a quasi-optical mixer basically consisting of the mixer chip with integrated antenna structure mounted on a silicon lens. The lens will have an anti-reflection coating made of Stycast ™ epoxy or Parilen™ C plastic.

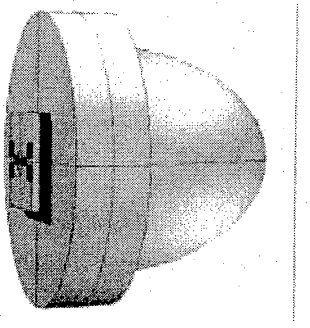


Figure 5.41 - Quasi-optical SIS mixer configuration

5.6.7.4.1 General description

The quasi-optical receiver chip will be based on Nb film technology. The losses and additional dispersion that occurs in Nb film above the gap frequency of Nb (~670 GHz) does not allow to use a simplified microstrip line model to be applied for parameter tuning. A full *rf* model including losses in Nb films has been developed during the design. It was found that it is possible to reach good receiver sensitivity at the upper part of the ALMA band 9 (602 – 720 GHz). However, the maximum sensitivity for some design has to be sacrificed in order to get a reasonably flat response across the band.
The *rf* structure of the receiver chip can be divided into four basic elements: the planar antenna, SIS junctions with integrated matching/tuning structure, *dc*/IF leads with IF on-chip transformer and magnetic field control line. The design includes different combinations of antenna structures, number of junctions per mixer, single-ended or balanced configuration and existence of control lines.

5.6.7.4.2 Antenna types

In our design two types of planar antennas are used, the double slot line antenna (DSA) and cross-slot antenna (CSA), Figure 5.42. The DSA is the most commonly used two-port antenna and the CSA is an experimental four-port antenna to be used in connection with the balanced on-chip mixer. The dimensions of the antennas are chosen to give an optimal far-field beam pattern of lens-antenna combination.
Four mixers can be connected to the CSA as shown in Figure 5.42. If the LO and *rf* signals are applied in the indicated polarizations, then the LO and rf signals appear in-phase for mixers M3, M4 and in anti-phase for mixer M1, M2. This symmetry with the proper combination of mixer IF outputs allows to use this configuration as balanced mixer.

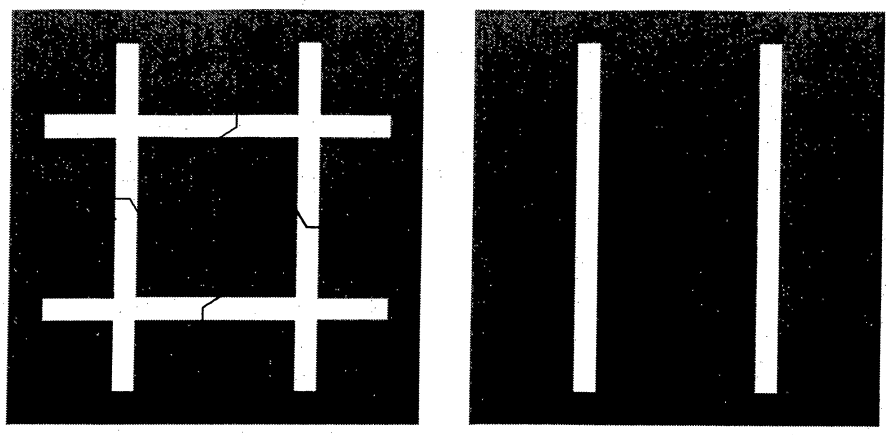


Figure 5.42 - Cross-slot antenna (left) and double slot line antenna (right). The cross-slot antenna in combination with four SIS mixers can be used as a balanced mixer.

5.6.7.4.3 Design types and *rf* properties

Nine design types are included in the mask set for the quasi-optical SIS mixer and are summarized in Table 5.8. The first four types represent a quasi-optical balanced configuration. Each of the four ports of the antenna is connected to a separate junction/tuning circuit. Depending on the polarization of the LO and interconnection of the IF output signals this type of receiver can be used as **double polarization** or **balanced** receiver. Type 5 represents a classical design that was developed also for 950 GHz at SRON. Each design type is reproduced on the mask at least 6 times. The calculated frequency response for types 1 and 2 is shown in Figure 5.43. It represents the *rf* power match from the antenna to the junction. The typical IF transient properties are shown in Figure 5.44. The additional IF tuning element improves the response in the range 2 to 12 GHz.

Table 5.8 – Quasi-optical mixer design types summary

Design type	Tuning structure	Antenna	Control line	Comments
Type-1	Single junction	Cross-slot	No	Balanced mixer
Type-2	Single junction	Cross-slot	Yes	Balanced mixer with control line
Type-3	Twin junction	Cross-slot	No	Balanced mixer
Type-4	Twin junction	Cross-slot	Yes	Balanced mixer with control line
Type-5	Virtual ground	Double-slot	No	End-point mixer (classical design)
Type-6	Single junction	Double-slot	No	End-point mixer (reference for type-1)
Type-7	Single	Double-slot	Yes	End-point mixer (reference

	junction			for type-2)
Type-8	Twin junction	Double-slot	No	End-point mixer (reference for type-3)
Type-9	Twin junction	Double-slot	Yes	End-point mixer (reference for type-4)

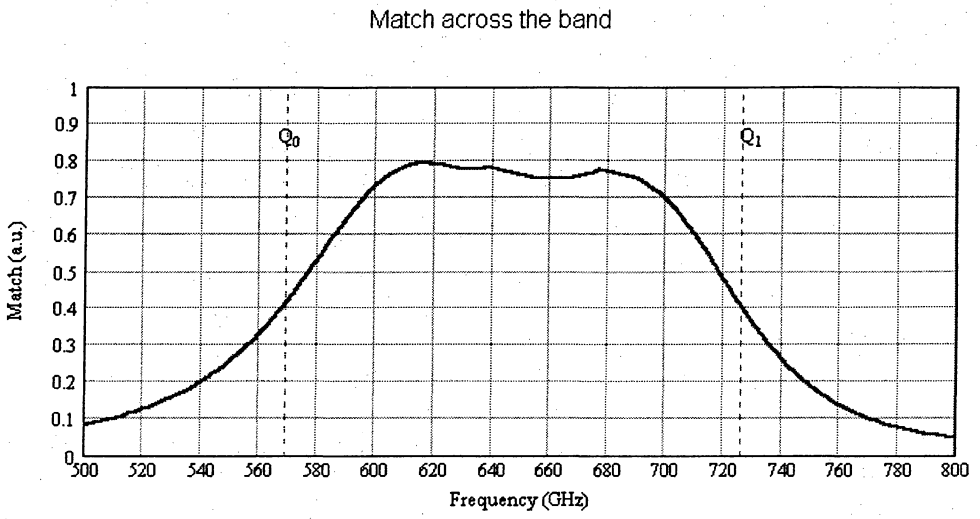


Figure 5.43 – Calculated frequency response of quasi-optical mixer types 1 and 2 (see Table 5.8).

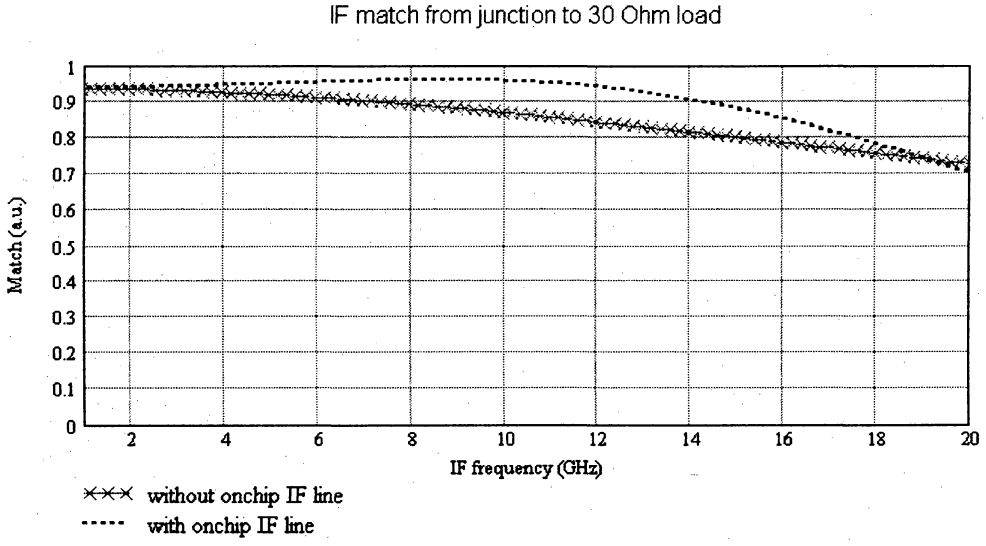


Figure 5.44 – Calculated IF frequency response of mixer types 3 and 4.

5.6.7.4.4 Mask layout

The 2" mask working area is divided into 177 3 x 3 mm square sections. Each section represents a different chip. Four places are used for alignment markers. The total amount of receiver chips is **173**. Each mask plate contains the mask set name "SIS-16" and its individual number 0…4. The ground layers of all chips are connected with each other and with the large contact pad at the edge of the wafer by means of an anodization grid. Each chip is marked with an individual number as well as with its type marker.

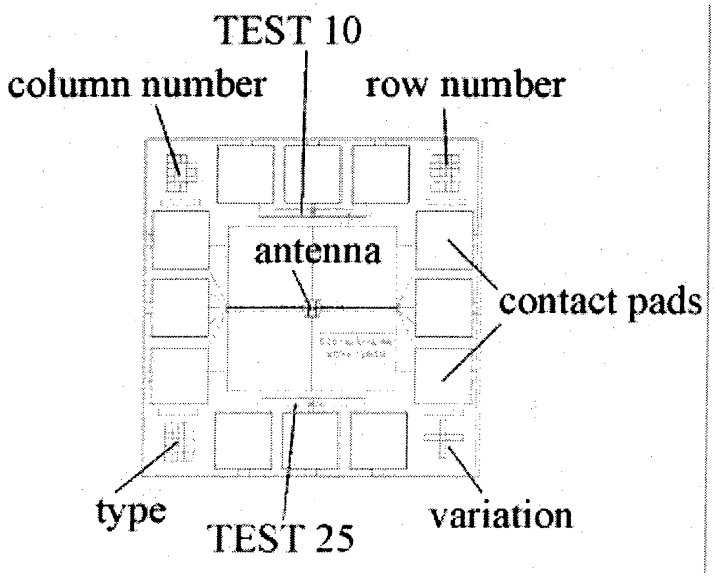


Figure 5.45 - Device chip layout

The SIS16 chip layout is presented in Figure 5.45. The antenna is situated in the center of the chip. The contact pads of size 0.5 x 0.5 mm are placed symmetrically at the four sides of the chip. Half of the designs contain test junctions of area 10 μm^2 (TEST 10) and 25 μm^2 (TEST 25). There are 9 different types of chip designs on the mask. There are three variations of junction size available for each design type. They are marked by the symbols "-", "+" and " " in the lower right corner. The junction area is 0.8, 1 and 1.2 μm^2, respectivelyly.

5.6.7.4.5 Layer sequence

Table 5.9 summarizes layering structure of the chip. The microwave properties for SIS16 were calculated assuming thickness and materials from the table and the following junction parameters:

Trilayer RnA	25…30 $\Omega \times \mu m^2$
Junction quality factor	> 15
Junction area	Set of 0.8, 1 and 1.2 μm^2.

Table 5.9 - Layer structure

	Name	Material	Thickness	Maskplate ##	File name	Definition
1	Base electrode	Nb	100 nm	Mask0	Sis16m0	Liftoff
2	Junctions	Nb/AlOx/Nb	100/1/100 nm	Mask1, Mask2	Sis16m1	Etch
3	Dielectric	SiO_2	300 nm	Mask1, Mask2	Sis16m2	Liftoff
4	Counter electrode	Nb	>400 nm	Mask3	Sis16m3	Etch
5	Gold	Al/Gold	>100 nm	Mask4	Sis16m4	Lift off

5.6.7.4.6 Tolerances

The tolerance of all structures in the mask unless specified in the following must be better than **0.5 µm** for mask0 ... mask3 and **1 µm** for mask4. Tolerances for "critical dimensions" and the smallest structure size are specified in Table 5.10.
The dimensions in mask 1,2,3 layers are corrected for technological parameter deviations. For the **counter electrode** it is assumed that all line widths are **decreased by 0.3 µm**, for the **junction definition layer** it is assumed that the final dimensions will be **decreased by 0.4 µm** as a result of all processing steps.

5.6.7.4.7 Alignment

The alignment markers (Figure 5.46) on this mask allow to align layers 1...4 with respect to layer 0. There are coarse and fine alignment elements. Nonius type structures technically allow to align layers within ± 0.05 µm. The required alignment tolerance is **± 0.25 µm**. This means that the two following layers can be misaligned by not more than **0.5 µm**. The marker for each layer is supplied with its own number.

Table 5.10 - Tolerances for critical dimensions for mask set SIS16

Layer name	GDSII file name	GDSII layer number	Smallest size (µm)	Tolerance (µm)*	Mask type
Base electrode (ground plane)	Sis16m0.gds	1	3.5 (slot)	± 0.3	"Negative"
		2	5 (slot)	± 0.5	
Junction definition 1	Sis16m1.gds	1	1.2 (line)	± 0.1	"Negative"
		2	5 (line)	± 0.5	
Junction definition 2	Sis16m2.gds	1	1.2 (line)	± 0.1	"Negative"
		2	5 (line)	± 0.5	
Counter electrode (wiring)	Sis16m3.gds	1	2.7 (line)	± 0.1	"Negative"
		2	5 (line)	± 0.5	
Gold pads	Sis16m4.gds	1	20 (line)	± 1	"Negative"

*Tolerances are specified as the absolute deviation of line (slot) width.

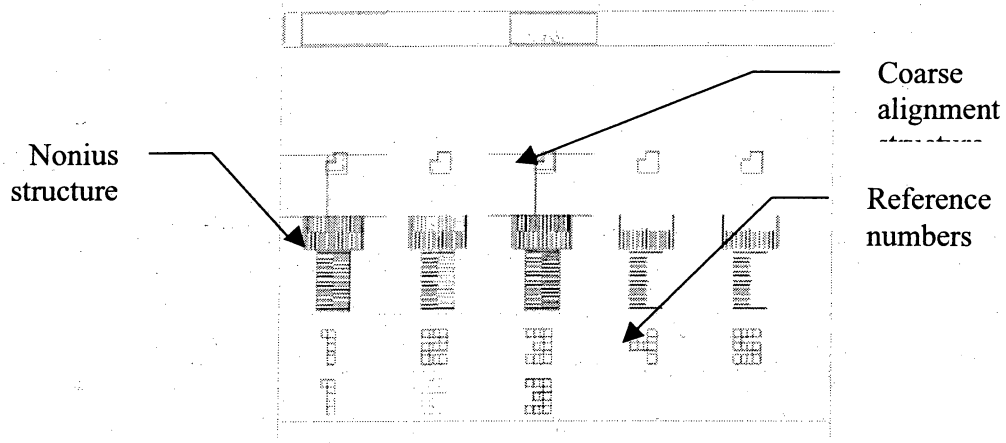


Figure 5.46 - SIS16 alignment markers structure. The base electrode is shown in red. Numbers in the figure correspond to layer numbers.

5.6.7.4.8 Other materials and technologies

The current design is tuned up for standard Nb/AlOx/Nb junction technology. The same mask set can be used without any modification with very high current density junctions (RnA=15 Oμm^2). These junctions could be made using a novel Nb/AlN/Nb process.

5.6.7.5 Single-ended 650 GHz mixer

Figure 5.47 shows a photograph of a so-called "D band" mixer, which is in use at the JCMT. The design tried to minimize the number of pieces and opted for simplicity. The mixer itself consists of a back piece, which holds the SIS junction and a corrugated feed horn (fabricated at RAL, UK), to which an Al lens holder is attached. Figure 5.48 shows the mixer noise temperatures across the *rf* band from 620 to 710 GHz for three different D-band mixers (called D5, D6, and D7). These fixed tuned SIS mixers provide a large *rf* bandwidth of about 150 GHz (Figure 5.49). The dip at around 560 GHz stems from water absorption in the atmosphere.

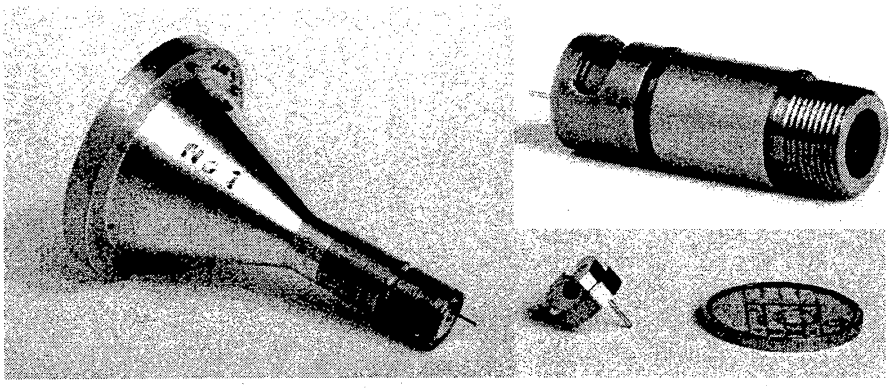


Figure 5.47 - Fixed tuned SIS mixer for the JCMT D-band (625 to 710 GHz) basically consisting of a corrugated horn (upper right) and a junction end piece (lower right). The larger aluminum piece (left) is a lens holder.

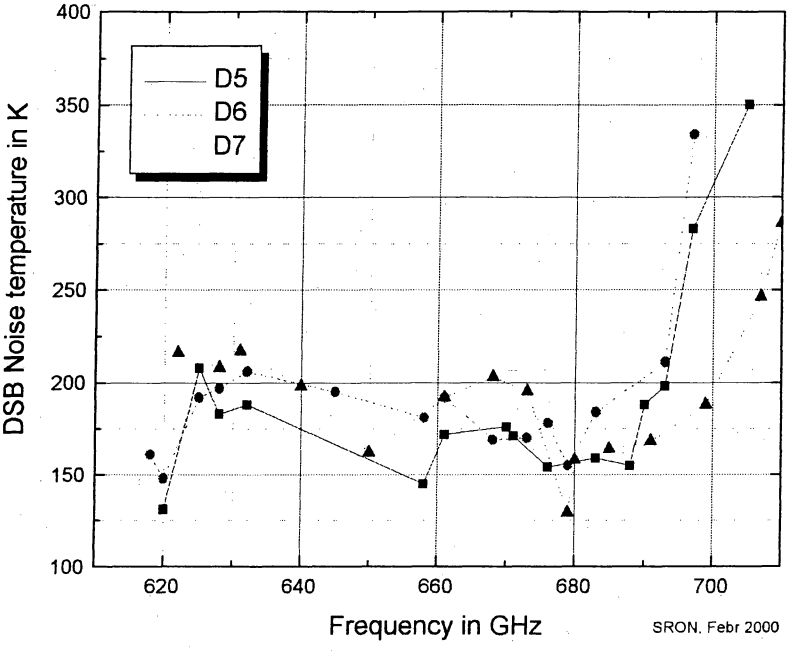


Figure 5.48 - DSB noise temperatures for three different D-band SIS mixers.

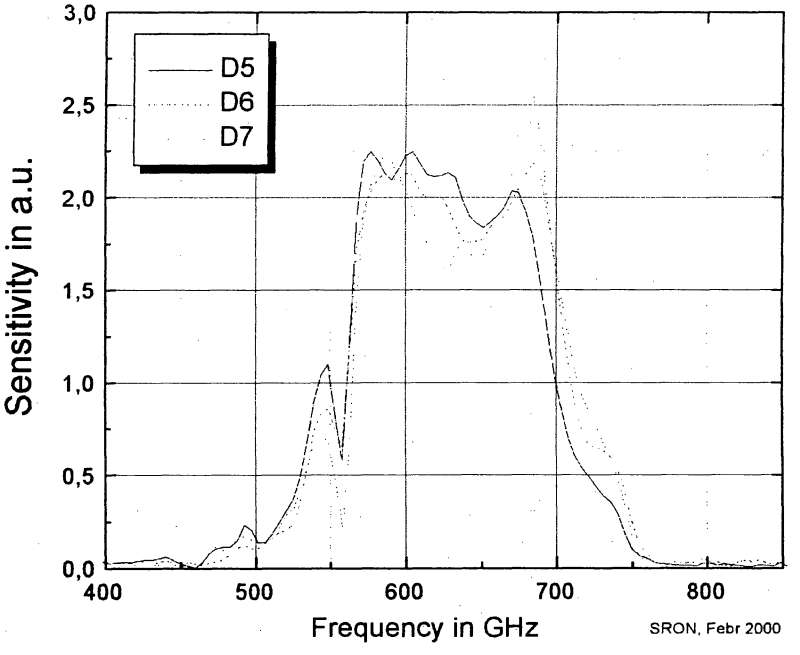


Figure 5.49 - Bandwidth of three fixed tuned SIS mixers. The mixers have been optimized for operation at 690 GHz.

5.7 The Water Vapor Radiometer.

A short version of the WVR proposal by Hills, Belitsky et al. could be inserted here. The original has 12 pages.

ALMA Receiver Optics Design

A. Baryshev (SRON), M. C. Carter (IRAM), L. R. D'Addario (NRAO), B. N. Ellison (RAL), W. Grammer (NRAO), Y. Sekimoto (NRO), J. W. Lamb (Caltech)

1 Introduction

Since the establishment of the Joint Receiver Design Group (JRDG), the requirements for the frontend have been refined and several major design decisions made, as summarized in Section 2. Principally they prescribe a single dewar for all the receiver bands, and these bands will be realized as individually testable cartridges.

A Workshop was held in Tucson in September 2000 to define a layout for the optics and dewar. Section 3 lists the design goals for the Workshop, and the configuration that emerged is detailed in this document. This includes the overall arrangement, and the details of the optics of the individual Bands. These bands are those defined by Wootten *et al.* [1] and endorsed with minor adjustment for Band 3 by the ALMA Scientific Advisory Committee (ASAC) [2]. The design goals led to a natural division of the bands into three groups. Representative designs for the optics of each of these groups are given.

2 Previous Design Decisions

During the project there have been several major decisions which are documented here for completeness. Some of these significantly influence the layout of the optics. Major decisions are

1. All cryogenic components will be in a single dewar.
2. There will be three temperature stages in the dewar at ~70 K, 12 K, and 4 K.
3. Each receiver band (frequency range covered by a single set of components) will be contained in a cartridge that may be tested as a unit and inserted into the dewar without disturbing any other cartridges.
4. The water vapor radiometer will be operated in a separate package at ambient temperature. It therefore requires a pick-off mirror to put the beam within 10 arcsec of any of the observing beams.
5. There will be no cold load for receiver calibration. However, provision will be made for a cold load for the vapor radiometer.
6. There will be no quasioptical diplexers for sideband rejection. This is based on a compromise between sensitivity and complexity (reliability)
7. The different bands will share the focal plane so that no switching mirror is required to select a given frequency.

These decisions were influenced by both performance criteria and practical considerations. Testability and mass production by separate groups dictated that the different receiver bands should be constructed in individual cartridges that can be inserted into a single dewar. The major disadvantage of this is that the dewar is larger than it would be with a more integrated approach, but the advantages of being able to change and test individual bands was deemed to be worth this sacrifice.

3 Design Goals

Because of the large number of bands (10 dual-polarization), the wide frequency coverage (> 4 octaves), and the high sensitivity requirements needed to profit fully from the excellent site, the trade-offs between some of the parameters is not straightforward. In general those parameters which affect most of the observing modes (*e.g.*, sensitivity, reliability) were favored above parameters which were for more specialized modes (such as circular polarization). The physical complexity of the receiver dictated that practical issues had to be seriously considered. These trade-offs are discussed in the text.

The design goals are divided into two categories according to whether they are fundamental to the performance of the receiver, or related to the practicality of construction and maintenance. In some cases there may be conflicts between these goals, but wherever possible the sensitivity should not be compromised.

3.1 Practical Goals

1. Receivers will be interchangeable between antennas (no on site alignment)
2. Optics part of Rx — Fixed alignment between optics and dewar
3. Antenna mounting flange should be preset within tolerances
4. Cartridges interchangeable between dewars with no re-alignment required
5. Alignment sensitivity to thermal contraction minimized
6. Alignment insensitive to dewar deflections under vacuum
7. Alignment ensured by machining
8. Optics in cartridge where practical
9. Optics cold where practical
10. No moving parts
11. Flexible for future upgrades/clearly defined interfaces
12. Standardize designs among bands
13. Maximize reliability
14. Minimize costs

3.2 Performance Related Goals

1. Minimize window apertures (IR loading, RF loss minimized)
2. Horn aperture less than 10 wavelengths
3. Minimize added noise
4. Maximize aperture efficiency: Reduction < 5 % relative to ideal corrugated horn
5. Aberrations: < 1 %
6. Truncation loss: < 1 %
7. Dissipative losses: < 1 %
8. Scattering losses: < 1 %
9. Polarization loss: < 1 %
10. Polarization: Beam squint: < 1 % of FWHM

4 Design Details

4.1 Cartridge Designs

4.1.1 Cartridge Style A

Band 1 (31.3–45 GHz) is the most demanding in terms of size. The optical elements are too large to place in the dewar, so only the feed horn will be cooled. Re-imaging optics are needed to allow a reasonable size of horn, and achieve high efficiency that is essentially frequency independent. Two mirrors, one plane and one ellipsoidal, were considered but the folded geometry resulted in optics that were too large to fit in the space above the dewar. Using a corrugated horn and a lens results in a much more compact design. Since a design for Band 1 had already been developed for the prototype antenna evaluation receiver [3] this was used for Band 1 and scaled for Band 2.

Figure 1 gives the layout and dimensions for the Band 1 optics. It comprises a conical corrugated horn and an aspheric PTFE lens. The corrugated horn was designed using a mode-matching program to optimize the pattern and the return loss (shown in Figure 2). It has an aperture radius of 30 mm and an opening angle of 4.57°. To determine the lens focal length and location, a frequency independent design was used so that the position of the waist from the horn/lens would match the required antenna waist position. Since the lens is quite thick it has a significant dielectric loss, and its diameter was chosen to make the loss due to truncation and dissipation roughly the same at 1.5–2 %.

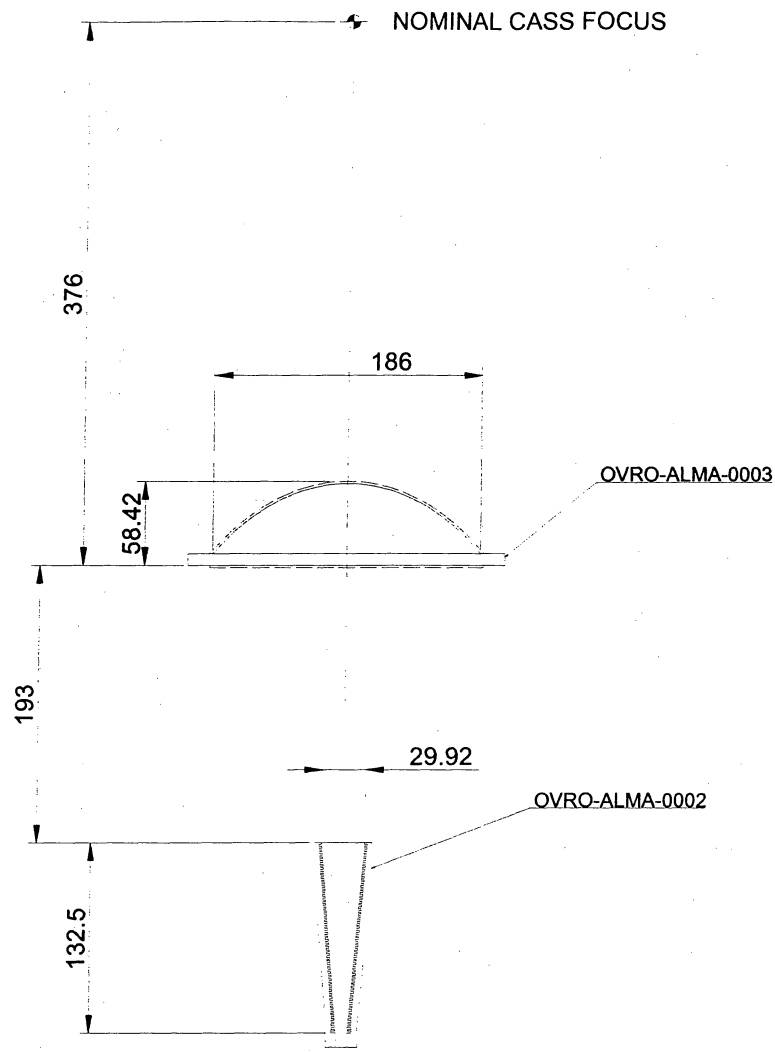


Figure 1: Band 1 optical design (will show IR filters and dewar wall).

Antireflection layers may be made by machining grooves in the surface. Straight grooves should be used on the flat surface of the lens, but this would be extremely difficult to do on the curved surface. Circular grooves can be used instead but those generate some cross-polarization and astigmatism. A viable alternative is to drill a regular array of holes into the surface, which is the choice for these lenses.

The lens is used as the vacuum window, which avoids an extra element in the optics. Since the truncation by the lens is at the level of a couple of percent, care has to be taken that the power not passing through the lens is terminated at ambient temperature.

Performance of the system was determined by calculating the antenna aperture field to find the aperture efficiency. Starting with the horn aperture field determined by mode-

matching, the field was propagated to the lens by integration. This field was traced through the lens by geometrical optics and a diffraction integral gave the distribution at the secondary mirror. Most of the contributions to efficiency are therefore accounted for, including: departures of the horn aperture field from the ideal $J_0(r)$ distribution; phase errors in the feed aperture; cross-polarization due to the feed; truncation loss at the lens, absorption in the lens; and blockage in the aperture plane. Losses which are not included are: reflection, scattering, and absorption in the filters; reflections at the lens surfaces; aberrations at the lens due to the phase center of the wave not being precisely at the focus; and losses due to the offset of the feed from the antenna axis. These were separately estimated.

Table I summarizes these various contributions. The feedhorn contribution was found by repeating the calculation with an idealized horn aperture field.

Because of the relatively small aperture of the lens, there is a perturbation of the pattern at the secondary due to edge diffraction. This appears mainly as a phase error which can be removed by refocusing the secondary mirror, which is assumed to be the case in the Table.

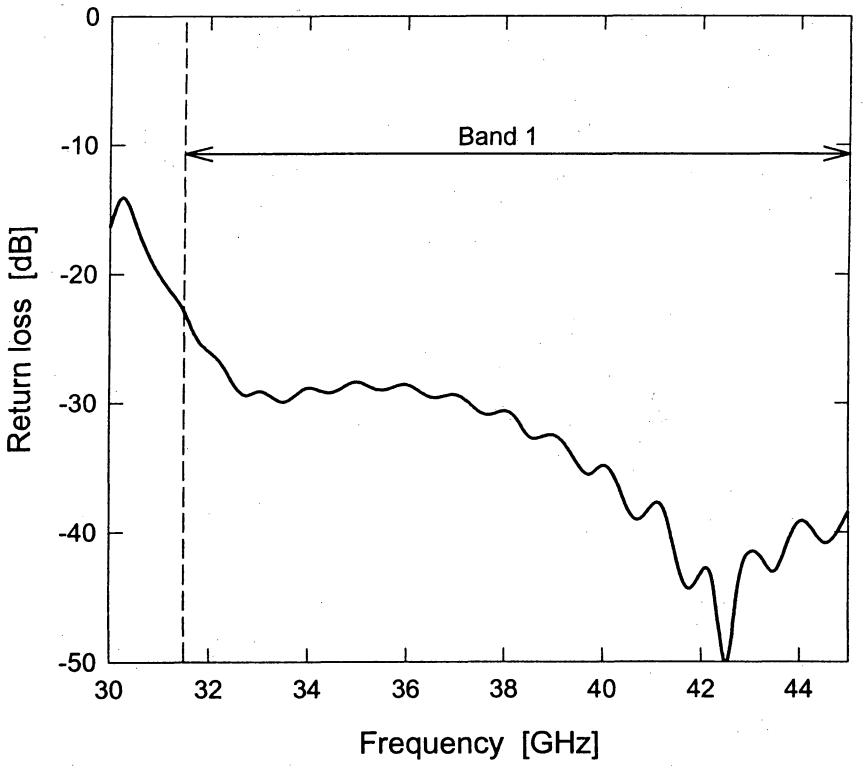


Figure 2: Return loss calculated for Band 1 corrugated horn.

TABLE I: ESTIMATED APERTURE EFFICIENCY LOSS AND ADDED NOISE FOR STYLE A CARTRIDGES

Band	Frequency [GHz]	Efficiency Loss [%]	Added Noise [K]
1	31.5	11.5	6.4
	38	8.5	8.7
	45	8.1	11.0
2	67	TBD	TBD
	81	TBD	TBD
	90	TBD	TBD

4.1.2 Cartridge Style B

Band 3 (84-116 GHz) and band 4 (125-164 GHz) are both Style B (Figure 3). External reflecting optics are used to produce a beam-waist close to the dewar wall to minimize the radiative heat load on the cryogenic system. A second focusing element is required inside the dewar since it is impractical to have the feed aperture very close to the dewar window. Since the cold optics are within the 140-mm diameter of the cartridge, it can easily be inserted into the dewar as a single unit. These cartridges occupy the outer circle at a diameter of 670 mm.

A plane mirror offset by 210 mm in the telescope focal plane reflects the incoming beam through an angle of 52° to an offset ellipsoidal mirror with a focal length of 87 mm. The ellipsoid reflects the beam through a 50° angle, making it parallel to the telescope axis, and produces a mid-band beam-waist of about 4.5 mm at the 24-mm diameter dewar window. After passing through an infrared filter anchored to the first stage shield of the dewar, the beam is imaged by a 60-mm focal-length lens on to a corrugated horn.

The corrugated horn supports both polarizations, which are separated in waveguide by an orthomode transducer.

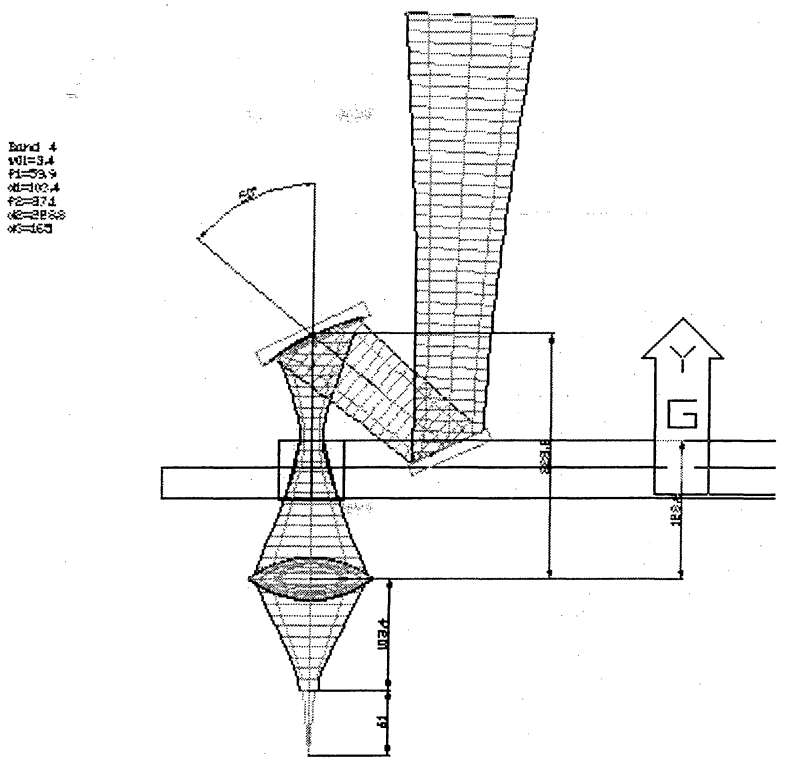


Figure 3: Layout and dimensions for Band 3 optics.

Table II lists the expected performance measures for Style B cartridges.

TABLE II: ESTIMATED APERTURE EFFICIENCY LOSS AND ADDED NOISE FOR STYLE B CARTRIDGES

Band	Frequency [GHz]	Efficiency Loss [%]	Added Noise [K]
3	84	TBD	TBD
	100	TBD	TBD
	116	TBD	TBD
4	125	TBD	TBD
	144	TBD	TBD
	163	TBD	TBD

4.1.3 Cartridge Style C

Style C cartridges include Bands 5–10. Internally, the optics vary among the cartridges according to the specific requirements, but they share some common features. For all the bands the antenna secondary focus is close to the dewar wall. All the imaging of the beam is done by cold reflective optics with 2 to 4 mirrors. Corrugated horns are the preferred beam-forming elements, but 'planar' structures such as twin-slot antennas on hyper-hemispherical will be considered. For the higher frequencies quasioptical polarizers and LO injection components may be required.

Optics for Band 6 are shown schematically in Figure 4.

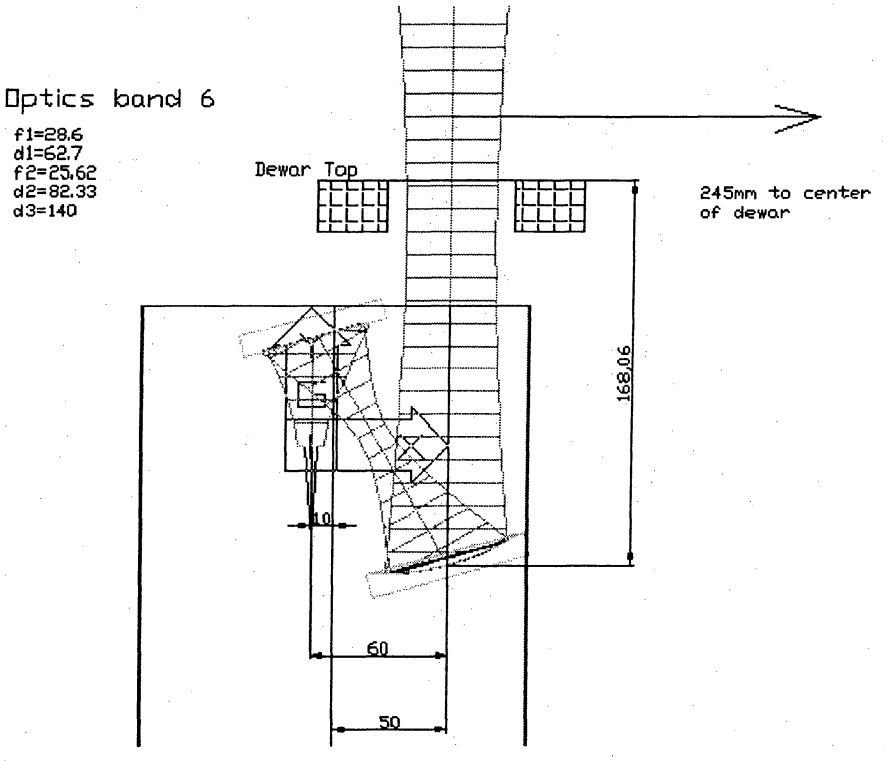


Figure 4: Optics layout and dimensions for Band 6.

Figure 5 presents the configuration for Band 7. The dewar window is radially offset by 130 mm from the dewar axis. The beam passing through the window and infrared filters is reflected through an angle of 116° by a plane mirror. It is then reflected by an ellipsoidal mirror where it is turned by 26° and focused onto a grid with a waist radius of about 2.4 mm. The grid separates the two linear polarizations, which are then incident on two identical ellipsoidal mirrors. They are then reflected down by 26° to converge into the corrugated horns by way of two flat mirrors.

Losses due to the beam distortion and cross-polarization are low due to the small incidence angles of the ellipsoidal mirrors.

Although it appears complex, this optical arrangement allows some flexibility in mixer design and LO injection.

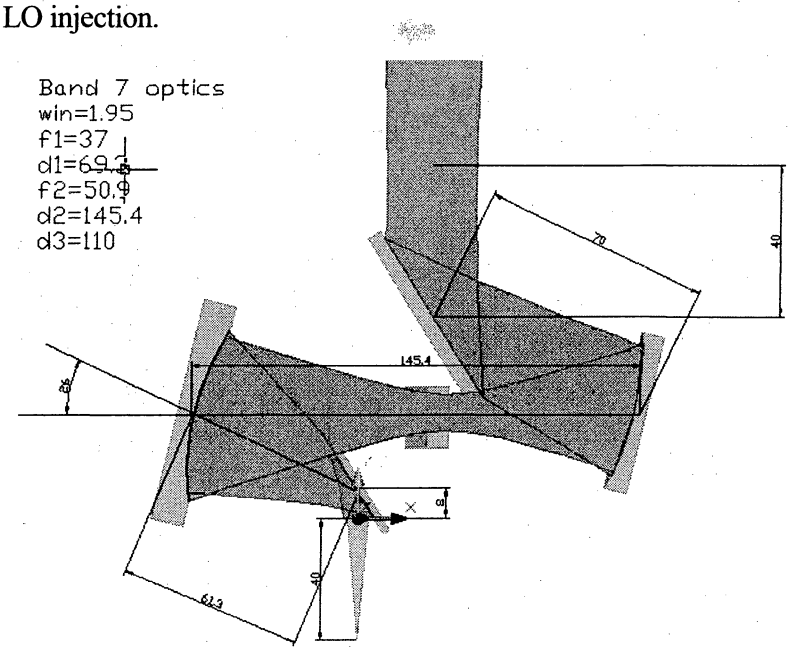


Figure 5: Schematic of the Band 7 optics configuration.

Band 9 optics are based on the optical design by van de Stadt for the FIRST HIFI receiver. The original design comprised two ellipsoidal (converging) mirrors with an intermediate hyperboloidal (diverging) mirror. The intention of this was to achieve significant cancellation of the distortion and cross-polarization of the individual mirrors according to known geometrical optics criteria. In adapting this to the ALMA receiver requirements, the hyperboloidal mirror was found to be close enough to a planar surface to omit entirely.

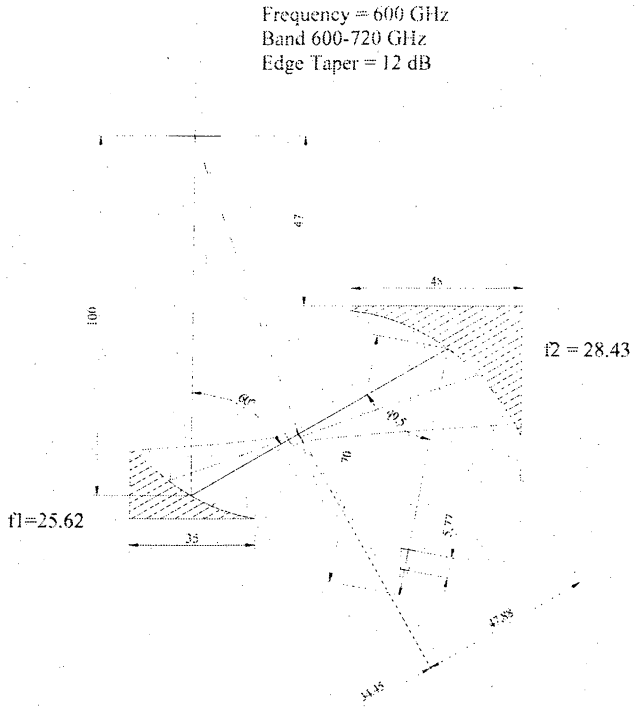


Figure 6: Optical scheme for Band 9 derived from the HIFI receiver optics concept.

For Bands 5–10 the expected performance measures are listed in Table III

TABLE III: ESTIMATED APERTURE EFFICIENCY LOSS AND ADDED NOISE FOR STYLE C CARTRIDGES

Band	Frequency [GHz]	Efficiency Loss [%]	Added Noise [K]
5	163	TBD	TBD
	187	TBD	TBD
	211	TBD	TBD
6	211	TBD	TBD
	243	TBD	TBD
	275	TBD	TBD
7	275	TBD	TBD
	323	TBD	TBD
	370	TBD	TBD
8	385	TBD	TBD
	442	TBD	TBD
	500	TBD	TBD
9	602	TBD	TBD
	661	TBD	TBD
	720	TBD	TBD
10	787	TBD	TBD
	868	TBD	TBD

4.2 Dewar Layout

A view of the cartridge locations is shown in Figure 7. The outer vacuum container of the dewar is 970 mm in diameter and 400 mm high. Ten cartridges may be accommodated in it, each containing a dual polarization receiver with an optical input, a waveguide LO input, co-axial IF outputs, and bias wiring. For cartridges with a diameter of 180 mm are located on the inner circle at a diameter of 300 mm for Bands 7–10. Bands 1, 2, 5, and 6 also utilize 180 mm diameter cartridges and are arranged on a 590 mm diameter circle. Cartridges with diameters of 140 mm located on a 670 mm circle house the remaining Bands, 3 and 4.

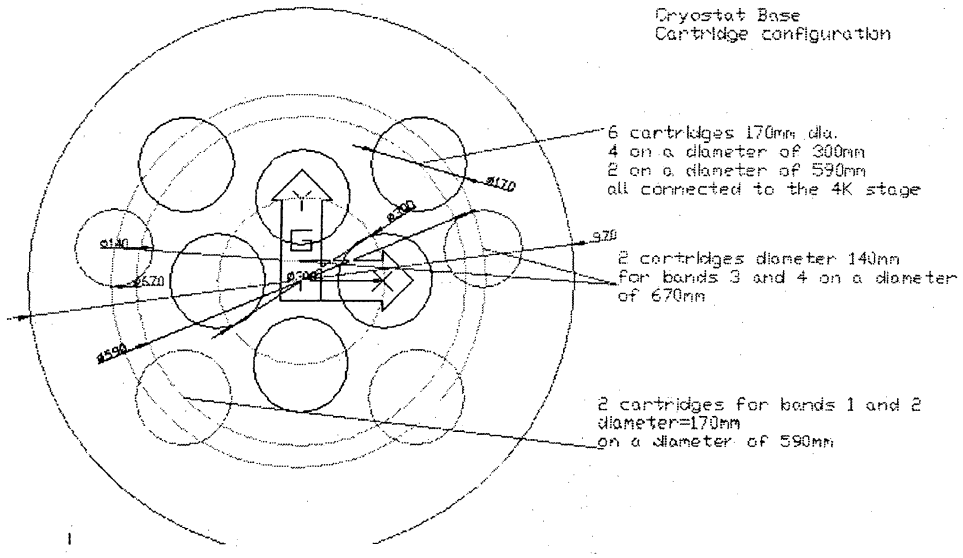


Figure 7: Locations of cartridges in dewar.

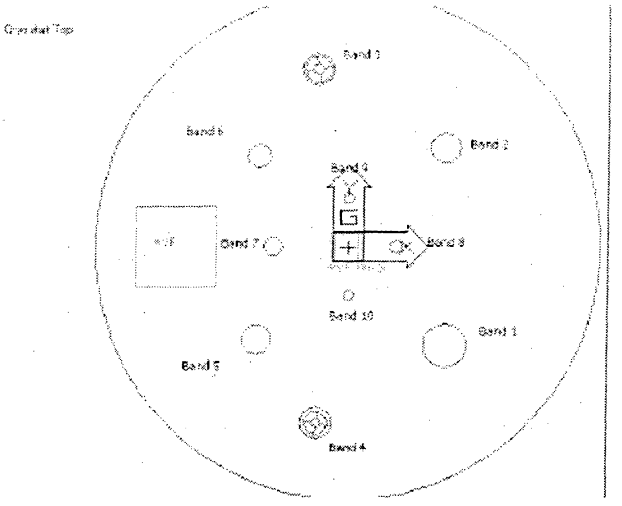


Figure 8: Arrangement of windows on the top of the dewar.

All of the cold optical elements are integral with the cartridges. As shown in Figure 8, the vacuum windows are directly above the cartridge axes for Bands 1–4. Windows for Bands 5–10, however, are offset towards the dewar axis by the folded optics in the cartridge.

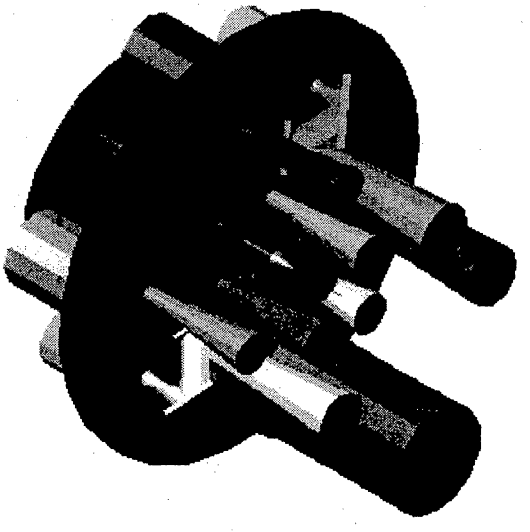


Figure 9: Layout of the optics showing the beams going into the dewar.

4.3 The Water Vapor Radiometer

The water vapor radiometer will be a self-contained unit with a pickoff mirror on the axis of the cryostat. This puts the WVR beam within 10 arcmin of any of the observing beams. This pickoff mirror may be rotated about an axis at a slight angle to the normal at its center to scan the beam and obtain information for atmospheric refraction correction. This requires a mirror which is about twice the size of the mirror which would be used to fully illuminate the aperture, but there is sufficient room for this. Note that the WVR beam can pass over the windows for other bands without problem.

5 Annex: Alignment Tolerances

Alignment tolerances for ALMA optics

Version 1.9

B.Lazareff IRAM 09-Nov-2000

Preliminary draft

1 Abstract

The purpose of this report is to define alignment tolerances for ALMA optics. A short note was circulated in April-2000 (hereafter referred to as version 1.0). It gave results for the alignment tolerances for what was then the optical train for band 10, believed to be the most critical w/r to alignment.

What is new compared with the version 1.0?

- Results are presented for bands 1–7, based upon the "final" optical design elaborated by the optics design workshop (Tucson, 25-29 Sep 2000), and the numbers provided by M.Carter (ALMA-Optics-13Oct00.xls).
- A second type of misalignment is considered, i.e. displacement of an individual element, as opposed to a "break" in the optical train.
- Refocussing off-axis mirrors are treated specifically for geometrical misalignment (were previously treated as in-line thin lenses).

What is still missing

- A check of the tolerance w/r to aberrations with the present "final" configuration. Version 1.0 of this note note showed that the tolerances w/r to aberrations were, in the case of band 10, significantly larger than those w/r to geometrical alignment in the aperture plane. It seems a safe assumption that this conclusion will hold for modified versions of band 10 optics, and a fortiori for optical trains operating at longer wavelengths.
- Bands 8–10, that were not defined by the September optical workshop, and whose optical design was left to the respective groups in charge of those bands.

Although incomplete, the present report is circulated to provide input to the dewar design.

2 Method, assumptions

In each of the 10 ALMA bands, an optical train is designed to couple the feed (horn, QO radiator) to the telescope. Its goal is to provide maximum coupling of the feed to a point source in the sky. Mechanical misalignments cause the parameters of the beam illuminating[1] the secondary to deviate from their nominal values. Such deviations can be classified as:

1) Displacement
   a) Along propagation axis
   b) Lateral shift in focal plane (= tilt in aperture plane, = pointing offset on the sky)
   c) Tilt in focal plane (= lateral shift in aperture plane, =loss of aperture efficiency)
2) Distorsion (coupling to higher order modes, if the launched beam is fundamental gaussian)

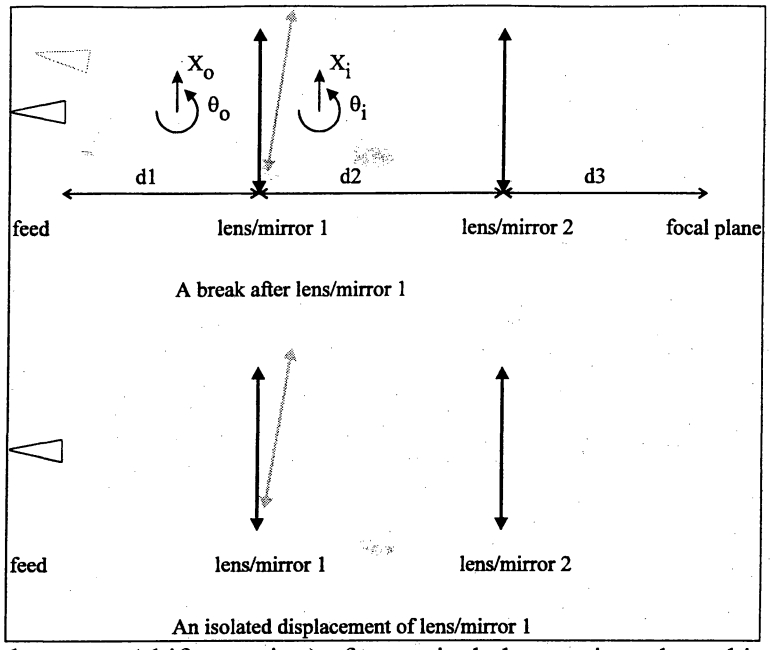
Effect 1a is not considered here, because even in the nominal design, the various bands are not constrained to have a common focus. Effect 1b is not considered either; lateral shifts of the beam illuminating the secondary would be at most of the order of a few mm, and cause a negligible loss of efficiency; they also cause a pointing offset which needs to be calibrated for each aband anyway. Effect 1c is what is considered in the present draft version, and is believed to be the main driver for alignment tolerances. A calculation of the loss of on-axis efficiency versus aperture plane misalignment was made in version 1.0, concluding to a tolerance of 6*mrad* in the focal plane for 1.3% loss of efficiency. Effect 2 was considered in version 1.0, and will be computed for the sake of completeness in the final version.

I use the ABCD equations of geometrical optics to propagate the perturbation of the chief ray; one can show that the same equations apply to the main axis of a gaussian beam. It is assumed that the unperturbed optical path is contained in a plane — which is true for the present ALMA design —, and only perturbations of the chief ray within that plane are computed, which simplifies the work while still providing a valid estimate for tolerances.

2.1.1 Perturbation matrices

Two types of perturbations to perfect alignment are considered: breaks and isolated displacements. The difference between the two is illustrated in the figure below, in the case of an inline lens system.

[1] As is common practice, I regard the optics and telescope working as a transmitter.



Note: The displacement (shift+rotation) of an optical element is reckoned in its image (**i**) space in the case of a break, and in its object (**o**) space in the case of a displacement. The center of rotation is defined in either case at the optical center of the element. These two frames are distinct only in the case of mirrors, of course.

In the case of a break, the (X,θ) displacement is propagated through the rest of the system, for example, in the case of a break after lens/mirror 1:

$$P = \text{Space}(d3)\cdot\text{Lens}(f2)\cdot\text{Space}(d2)$$

where Lens(f) is the usual ABCD matrix for a thin lens/mirror, and Space(d) the matrix for free propagation over a distance d.

In the case of an individual displacement of an optical element, we need the perturbation of the image ray as a function of the displacement of that element:

$$\begin{pmatrix} X_i \\ \theta_i \end{pmatrix} = \begin{pmatrix} 1-\cos(\alpha) & 0 & \sin(\alpha) \\ -\frac{1}{f} & 2 & 0 \end{pmatrix} \cdot \begin{pmatrix} X_o \\ \theta_o \\ Z_o \end{pmatrix}$$

in the case of a mirror (the only case where a Z-displacement along the ray path is significant); and:

$$\begin{pmatrix} X_i \\ \theta_i \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ \frac{1}{f} & 0 \end{pmatrix} \cdot \begin{pmatrix} X_o \\ \theta_o \end{pmatrix}$$

in the case of a lens (insensitive to lens tilt in the paraxial approximation); in either case, the perturbation is then propagated through the rest of the system.

The end result is a 2×2 matrix (2×3 for the individual displacement of a mirror) that relates the X,θ displacement of the beam in the focal plane to the misalignment parameters. Only the second line of that matrix (that provides the θ-displacement of the beam) is used, together with the tolerance on the shooting angle $\theta_C = 6\times10^{-3} rd$, to derive the positioning tolerance for the misalignment under consideration:

$$Tol(X) = |P_{2,1}|^{-1} \times \theta_C$$

$$Tol(\theta) = |P_{2,2}|^{-1} \times \theta_C$$

$$Tol(Z) = |P_{2,3}|^{-1} \times \theta_C$$ (only in the case of a mirror's individual displacement)

3 Results

I give for each band the tolerance (lateral shift X and tilt θ) for breaks and individual displacements. I also give the relevant elements of matrix P, that can be useful in the case of simultaneous and deterministic shift and tilt displacement, where one should combine algebraically the effects on the output beam before estimating the tolerance. Units are mm, radians. The tolerance for a break after the last element is trivial, and is shown only for completeness.

Band		Feed break	Displ element #1	Break after element #1	Displ element #2	Break after element #2
1	$Tol(x)$ $Tol(\theta)$ $[Tol(Z)]$	1.14 ∞	1.14 ∞	∞ 6×10^{-3}	N.A.	N.A.
1	$P_{2,1}$ $P_{2,2}$ $[P_{2,3}]$	-5.3×10^{-3} *0*	5.3×10^{-3} *0*			
2	$Tol(x)$ $Tol(\theta)$ $[Tol(Z)]$	0.58 ∞	0.58 ∞	∞ 6×10^{-3}	N.A.	N.A.
2	$P_{2,1}$ $P_{2,2}$ $[P_{2,3}]$	*-0.010* *0*	*0.010* *0*			

Band		Feed break	Displ element #1	Break after element #1	Displ element #2	Break after element #2
3	$Tol(x)$ $Tol(\theta)$ $[Tol(Z)]$	0.47 0.34	0.27 ∞	0.63 4.3×10^{-3}	0.63 3×10^{-3} ∞	∞ 6×10^{-3}
3	$P_{2,1}$ $P_{2,2}$ $[P_{2,3}]$	*0.013* *-0.017*	*-0.022* *0*	*-9.5×10^{-3}* *-1.38*	*-9.5×10^{-3}* *2* *0*	
4	$Tol(x)$ $Tol(\theta)$ $[Tol(Z)]$	0.437 0.749	0.25 ∞	0.58 3.7×10^{-3}	0.58 3.0×10^{-3} ∞	∞ 6×10^{-3}
4	$P_{2,1}$ $P_{2,2}$ $[P_{2,3}]$	*0.014* *-8.0×10^{-3}*	*-0.024* *0*	*-0.010* *-1.59*	*-0.010* *2.0* *0*	
5	$Tol(x)$ $Tol(\theta)$ $[Tol(Z)]$	0.35 0.052	0.20 2.2×10^{-3} 0.687	0.37 4.3×10^{-3}	0.37 3.0×10^{-3} ∞	∞ 6×10^{-3}
5	$P_{2,1}$ $P_{2,2}$ $[P_{2,3}]$	*0.017* *0.115*	*0.031* *-2.808* *-8.7×10^{-3}*	*-0.016* *-1.404*	*-0.016* *2* *0*	

Since the optical trains for bands 5, 6, and 7 have nearly identical parameters (from the point of view of geometrical optics), the tolerances for band 5 apply also to bands 6 and 7, and I have not performed separate calculations.

When the optical train comprises a final planar mirror (bands 3 and 4), an angular tolerance of 3×10^{-3} rd applies to that mirror.

If you have read so far, you have certainly noticed that, for bands 3, 4, 5, the matrix element $P_{2,1}$ has the same value for "Break after element 1" as for "Displacement of element 2", instead of having opposite signs. That is because I have been lazy: for the *propagation* of the ray, I used the matrix $Lens(d)$ for a mirror while I should have used $-Lens(d)$. This affects neither the values of the tolerances, nor the *relative* signs of matrix elements in the same box of the table, which are of concern when combining deterministic X, θ displacements in the optical train.

The symbol ∞ should not be understood literally, it just means that the considered displacement produces a pure lateral shift of the beam, and within a few mm, such a shift is negligible. Large angular tolerances are found for the feed, especially in bands 1–4; this is due to the feed being imaged approximately to the aperture plane (in fact, with the numbers supplied by M.Carter, that condition is not always met exactly). There again, that large tolerance should not be interpreted literally: beyond a certain point, the beam from the feed might spill off the finite aperture of the

mirrors; but with easily achievable angular tolerances (like 0.01rd) for feed placement, this should not be a concern.

References

[1] A. Wootten, L. Snyder, E. van Dishoeck, F. Owen , "Frequency band considerations and recommendations," *ALMA Millimeter Array Memo Series*, No. 213, May 2000.

[2] ASAC

[3] J. W. Lamb, "ALMA evaluation receiver optics design," OVRO-ALMA Design Report, Apr. 2000.

ALMA Memo #303
Water Vapour Radiometers for ALMA

Richard Hills and John Richer

Introduction

This is a discussion document setting out the options for performing atmospheric phase corrections by means of radiometry. It is a lightly edited version of a paper prepared for the March 2000 meeting of the Science Advisory Committee. We have tried to take account the comments received but not yet responding to the conclusions of that meeting.

A great deal already been written on this subject. In particular, the relevant memos and other documents have been summarised on the ALMA web site at http://www.alma.nrao.edu/development/cal_imaging/phasecal.html.

Status of current 183 GHz phase correction experiments

The JCMT-CSO single-baseline interferometer was the first to demonstrate phase correction using the 183 GHz line, using equipment built by Martina Wiedner, Richard Hills and colleagues. Only a limited quantity of data were gathered but the results (ALMA memo 252) were encouraging and suggested that even an uncooled system could provide effective phase calibration at submillimetre wavelengths. Single baseline interferometry at JCMT-CSO is no longer a supported mode of operation, so further observations would be difficult though not perhaps impossible to arrange. It is however possible that two SMA antennas can be equipped with 183 GHz systems, using the radiometer currently at CSO plus a clone of it being built in Canada by Christine Wilson and the HIA. It is unclear when this experiment might produce results on Mauna Kea, but access to a large set of data in a variety of atmospheric conditions would certainly be useful in establishing the capabilities of the technique.

On the Chajnantor site, two further 183 GHz radiometers are in operation; these were built as a collaboration between Onsala and Cambridge and are very similar to the Mauna Kea systems, again using uncooled DSB mixers and three roughly 1-GHz wide filters. These two independent systems are aligned with the twin 11-GHz site testing interferometers, with their beams matched as well as possible using newly designed mirrors. The intention is to see how effectively and for what fraction of the time it is

possible to use the 183 GHz systems to correct the 11GHz atmospheric phase measurements. It is possible to estimate the height of the turbulence from the lag between the two 11 GHz phase measurements (which are obtained by looking at different satellites) together with information on wind speed and direction. This will be important in establishing how strongly the quality of radiometric phase correction depends on the turbulent scale height, both in practice and through models.

Initial results for both the lag estimates and radiometric phase correction have been obtained in the past 2 months, although operational difficulties (principally power outages, and the difficulties in performing system upgrades and receiver tests on site) have restricted the quantity of data so far obtained. Work on analysing the existing data and on improving the measurements will continue as a high priority, with the goal of producing a report in about 6 months.

Although the results from these more detailed studies will be needed in order to answer some of the questions, we need to have an initial set of specifications for the ALMA radiometer system and a baseline design for inclusion in the plans and cost estimates. We do in fact have sufficient information to provide much of this information already. The following sections summarise our current thinking on the requirements and the design choices.

Design Considerations for the ALMA water vapour monitors

1) Requirements

The first question to be decided is whether we wish to correct just the phase error in the interferometric signal or whether we should also plan to take out the tilts in the wavefront across the individual dishes which cause pointing errors. (The latter effect is sometimes called anomalous refraction, although it is only anomalous in the sense that it would not occur if the atmosphere were uniform.) The correction of such pointing errors with radiometers was discussed by James Lamb and Dave Woody in MMA Memo 224. The problem has been studied in the context of the LMT/GMT 50-metre project by Luca Olmi. (Radio Science, 35, 275-286, Jan 2000.)

In each case we then need to set detailed requirements. We need to decide the path length error allowed as a function of integration time, weather conditions, zenith angle (z) and change in z. For pointing corrections, we need to set the required accuracy (which should be a term in pointing error budget) again presumably as a function of weather and z.

The rms path error given as the goal in existing documents is 38.5 fs (femtoseconds) which is 11.5 micrometers of path. Note that, at this level, the loss of correlation from this cause is only 5% at 950 GHz and 0.7% at 350 GHz, so this is setting the goal very high. (Compare these to the transmission losses of about 70% and 20% for these same frequencies with 1 mm of water vapour.) No reference is made to whether this figure degrades in less than ideal conditions, but is clear that it can be allowed to without seriously affecting the data. A more realistic goal would be to multiply the above figure by $(1 + w_v)$ where w_v is the amount of water in the path in millimetres.

The time allowed for achieving this accuracy is also not presently specified. We have generally been assuming that this refers to a one-second timescale, but we really need to look more closely at the data to see if we are justified in going as fast as this. (Note that the question of whether the correction is applied to the phase in real time or the data taken with short dump-times and stored for later processing has only a small effect on the radiometer requirements but quite large implications for the software.)

A "systematic (avg)" error of 8.4 fs is also quoted in Larry D'Addario's "Phase Stability Specification Note". This is intended to cover things like drifts and offsets that do not integrate down with time in the way that white noise would. The relevant timescale here is the time between calibrations because anything that varies more slowly than this will be taken out by the observations of the calibrator. We will presumably observe calibration sources much less often than we would if we were using only them to remove atmospheric phase fluctuations, but a typical calibration cycle of 50 or 100 seconds seems reasonable. We can, if necessary, move further and use brighter sources than is planned for fast-switching phase correction. Presumably the same observations will generally be used to check the pointing and/or the amplitude calibration. We should try to extend the stability of the radiometers to at least a few times the calibration cycle so that the results on the calibrators give us an independent estimate of how well the correction is working. Stability over about 5 minutes is therefore the requirement. (Note that this implies that the phase stability of the rest of the system must be maintained for at least this length of time.)

It is however essential that we can measure the atmospheric term accurately as we move from source to calibrator. This is certainly more difficult if there are large changes in the total water in the path and/or ground spillover (although it is only the dish-to-dish differences in these effects that are important). At low elevations it would be beneficial to look for calibration sources that are closer to the target in zenith angle than in azimuth, i.e. to search in an elliptical patch of sky.

The value given above (8.4 fs) is an extremely tight specification: it corresponds to less than a degree of phase at 345 GHz, well below that given for the path stability on the antennas for example. We think it unrealistic to insist on this stability from the radiometer system and suggest instead that the value proposed for the short-term noise should apply for the stability over the calibration cycle as well. Obviously the goal will be to do better.

The key sensitivity number is that at the optimum frequency the change in brightness temperature is ~$15/w_v$ mK per micrometer of added path. This suggests that a radiometric precision of order 150 mK (corresponding to ~10 microns of path) would be sufficient in good conditions. Given bandwidths of >100 MHz and an integration time of 1 second, this looks reasonable, even for a room temperature mixer, for which Tsys of 1500K should be possible.

For antenna pointing corrections a suitable budget allocation is 0.3 arcsec rms (in dry conditions). This is a wavefront slope of 1.5 microns per metre, which leads to a figure of 9 microns when taken between two points 6 metres apart on the dish. The measurement is however now a difference between two numbers and it probably has

to be measured in shorter times than the interferometric phase. This looks marginal with a single uncooled mixer.

Studies of the existing site data (e.g. MMA Memo 223 and references therein) show that much of the observing time will be seriously affected by single dish pointing errors: the overall median seeing is about 1 arcsec compared to the specification for the antennas of 0.6 arc seconds. More study is needed of how fast the pointing fluctuates and how the bad seeing correlates with the other conditions. The obvious conclusion at this stage is that we do need to correct the pointing and that we should assume that this needs to be updated once per second. (With a wind speed of several metres per second and a 12-metre aperture, we can obviously expect some pointing changes on timescales as short as this, but the bulk of the power will normally be a periods of more like 10 seconds.) Note that this has to be done in real time and that we will therefore need to use an algorithm that anticipates the error for a time about one second ahead of the most recent reading.

Other requirements: Compatibility of interfaces (CANbus, etc.) Minimum interference with other systems. A special problem is leakage of the LO and its harmonics into other systems via various paths e.g. out of the feed and by reflection off the subreflector. It is unlikely that we can suppress these completely. The LO's should therefore be locked to system clock so any interference is at an accurately defined frequency. The design should use the fixed reference frequencies already provided at each antenna. (Using 2 GHz and 125 MHz would provide a satisfactory combination). We might add a requirement that the LO can be shifted by a small amount (say 125 MHz) so that any interference can be moved away from a critical line (by about 200 km/s in that case).

Suggested baseline spec: $10(1 + w_v)$ microns of path and $0.3(1 + w_v)$ arc sec of pointing over a 5 minutes of time and 1 degree in z, with 1 sec time resolution.

2) Basic technical approach.

The obvious options are line measurements at 183 GHz, 22 GHz, and in the mid-infrared (10 or 20 microns), or measurement of the (sub)mm continuum as for example used at IRAM.

The latter is unlikely to provide accurate enough path estimates and could not easily accommodate a wide range of conditions.

22 GHz is now essentially ruled out by the size of the optics. The feed would be ~250 mm diameter to measure the interferometric phase and at least 500 mm for correcting the pointing. Sensitivity would in any case be problematical - a cooled system would certainly be required.

The use of infra-red radiometers is a new suggestion from Dr David Naylor (Lethbridge, Canada). The principle is essentially the same as with the millimetre radiometers but uses water vapour emission bands in the mid-IR. The system uses detectors cooled to 77 K. We could not use the telescope's optics so to measure the pointing corrections we would probably need either several detectors per dish or some

optical relay system to give an appropriate spreading of the beam. The initial report on sensitivity and stability looks encouraging, but questions such as how much the results are affected by the temperature and pressure in the fluctuating layer and the effects of cirrus clouds have yet to be investigated. This needs to be done before we can judge whether this might be a viable option for ALMA.

Meanwhile the baseline should remain 183 GHz.

3) Mixer or HFET?

183 GHz HFETS will probably be available but will be expensive, noisy and with poor short-term stability. The baseline should be to use mixers.

4) Cooled or uncooled?

The main advantages of cooled systems are sensitivity and stability. It would also be easy to provide a cold reference load. There is however some concern about how one would calibrate out losses in the Dewar window, especially if there is a possibility of getting dirt or water on it. External optics would almost certainly still be required for the pointing system and it might be possible to introduce some additional calibration signal there. With cooled systems, the radiometer will essentially take up one complete slot in a Dewar and the development path will interact strongly with the main receiver programme. It will also take up some of the cooling power budget (IF amps, windows, connections, etc.) and there would be greater likelihood of LO power leakage.

An uncooled system is clearly simpler, and should cost less to develop and build. Uncooled Schottky mixers can be obtained commercially and are robust and stable.

We therefore believe that an uncooled system should be adopted as the baseline. Assuming, however, that the goal of correcting the pointing is confirmed, there is some question as to whether sufficient sensitivity can be obtained with an uncooled system. Until this is established the cooled option should be kept open as the backup.

Digression on cooled systems:

5a) SIS

If we use SIS mixers, these will have to go in the main Dewar and will presumably be based on the ALMA band 5 mixers. Sensitivity is then excellent and stability almost certainly acceptable given a suitable switching scheme. One can argue that no significant development effort on the mixers is required. The standard IF choice is not ideal (1 to 9 GHz would be better), but we could live with it. For example the LO could be at about 180 GHz so that the upper-sideband IF range of 4 to 12 GHz would correspond to line offsets of ~0.7 to 8.7 GHz. The lower sideband would not be used and would have to be rejected at about the 25 dB level. The mixers would provide a

certain amount of sideband rejection and this could be enhanced by having a waveguide filter at the input to the mixer, since the operational frequency is fixed. Although there will naturally be strong resistance to giving up one of the astronomical "slots" (or making the Dewar larger and more complicated), this option is sufficiently attractive that it should probably be kept open for the present. A straw-man design for it could be worked up and costed but no development work seems to be needed now.

We should also consider here the possibility of using the astronomical band-5 receiver to do the radiometry. Given the high sensitivity it might be possible to obtain sufficient accuracy from the shape of the line plus perhaps frequency switching, in which case it should not be necessary to compromise the astronomical performance of the receiver by adding additional switching components inside the Dewar. Another option would be to insert a 45-degree polarising grid into the beam when selecting this mode. This would make it possible to use the two polarisation channels as a cross-correlation receiver. This should also provide a way of doing sideband separation. This would of course mean that correction would not be available when using this receiver for astronomy. (Under good conditions, however, it might be possible to do the water vapour measurements with the band-7 receiver using the 325 GHz water line.) Some additional electronics for generating the LO and processing the IF would need to be added. Extra optics would be needed to do the single-dish pointing corrections and these would have to be inserted into the beam to select this mode.

An important additional consideration is that using an SIS mixer should give sufficient sensitivity to provide a correction for the water vapour *emission* when making total power observations with another receiver. One can see that this should be possible from the fact that, for 1mm of precipitable water, the extra emission ΔT_b for a given Δw_v is several times stronger between 181 and 185 GHz than it is at say 345 GHz.

Again these options seem sufficiently attractive that they should be explored in more detail. The interactions with the rest of the system are nevertheless a substantial negative factor. If nothing else we would be compelled to have band 5 available on all antennas from day 1, which may not coincide with the astronomical priority.

5b) Cooled Schottky

The advantage of using a cooled Schottky system is that it could be housed in a separate Dewar with the band 1 receiver (if that is the outcome of other discussions) where it could be cooled to 15 – 20 K. The interactions with the more critical part of the receiver system would then be reduced. It would however probably be necessary to undertake a new development to obtain suitable mixers and we are not clear what performance could be obtained. The IF amplifiers would probably play a major role here and it may again be best to use the ALMA 4 to 12 GHz ones. If we decide to use a Dicke switch (see below) then we would probably need to develop a suitable coolable switch. This option should be considered further if detailed planning for a band 1 Dewar is undertaken.

Finally in this section, we should note that very compact and relatively cheap refrigerators are now available which could cool a simple radiometer to say 70K. Although reliability might be an issue, it may turn out that this is the most cost-effective way of getting the necessary sensitivity if it cannot be obtained with an uncooled system.

6) Form of switching

For an uncooled system, there seems little chance of obtaining ~0.1 K stability with a total power system given a system temperature of at least 1000 K. (Note that we can get some relief because we are observing a line and are to a considerable extent only concerned with the differences between frequencies. We believe that some form of comparison with a load of known temperature will however be necessary.) We should therefore plan to use either a Dicke switch or a continuous-comparison radiometer which takes the difference between the sky temperature and a temperature-controlled load. For the pointing correction we also need to take the differences between different parts of the aperture. Many options are available but we clearly wish to select the simplest, cheapest and most reliable that can do the job.

The most basic option is a single mixer with a Dicke switch operating between the sky and a fixed-temperature load. Ideally this load should be at a temperature close to that of the sky brightness at which one obtains the best sensitivity (around 170 K). A modulated calibration signal would also be injected via a coupler on the input. An alternative to injecting a cal signal is to switch between the sky and two loads at different temperatures. This gives more flexibility in the choice of temperatures: something like 100 K and 250 K (spanning the sky brightness range of interest) would be best, but combinations like 200 K and 370 K would also be good. The existing MRAO design uses two loads and an optical switching scheme (a "flip-mirror"). This works quite well, but for ALMA it would probably be worth developing an all-electronic switching scheme, using ferrites or diode switches, for both reliability and stability reasons. With a single mixer the system would normally run in double-sideband mode and, provided the gain stability was adequate, the sensitivity would be given by the normal radiometer equation: $\Delta T = 2\, T_{sys}$ (DSB) / root(Bt).

The next level of sophistication is to use two mixers. With a hybrid before the mixers and a correlating backend one can then arrange that the output is the difference between the sky temperature and the load. (The use of a correlation receiver in this application is suggested in Luca Olmi's paper and he refers to the work of Predmore et al. (IEEE Trans MTT-33, 44, 1995) as a successful example of a millimetre-wave continuous comparison radiometer. The sensitivity improves by root 2 and with appropriate switching we can presumably separate the sidebands as well, although the advantages of doing this do not seem very great. (It would perhaps give better information about any contribution from clouds.)

To obtain the gradient in the emission, which gives us the pointing correction, we need to arrange the optics so that the radiometer illuminates a patch on the subreflector, covering about half of it. For a switching scheme the beam then has to

be moved around (most naturally as a circular scan about half way out) and the signal put through a pair of synchronous detectors to generate the required error signal. Lamb and Woody suggested a rotating prism to do this but a rotating mirror with its normal slightly tilted with respect to the axis of rotation would also do the job.

An alternative is to again use correlation (i.e. continuous differencing) receivers. The most obvious arrangement would be to have 4 horns in a square, which are optically re-imaged onto the secondary. The two diagonal pairs are connected to 4 mixers via hybrids in such a way that the outputs are the differences in the sky brightnesses required. A mechanism for switching against loads would still be needed to give the interferometric phase correction. Although these schemes sound complicated, the technology does probably now exist to build such combinations of splitters, hybrids and mixers in a stripline form at these frequencies.

More discussion of these schemes seems appropriate before a choice is made here.

7) Form of backend

In principle we could scan the LO and use a fixed and very simple IF with just one fixed frequency. Given that we are struggling for sensitivity this seems unattractive. The stability would probably not be good either. We therefore need a multichannel backend. The obvious choices are a set of filters (as in the MRAO and Onsala systems) and an analogue correlator along the lines developed by Andy Harris. This latter approach is being adopted for a 22 GHz water vapour phase correction scheme on BIMA. (See http://bima.astro.umd.edu/memo/memo67.ps)

More modelling is needed to determine the number of filters required. Studies such as those performed by Bryan Butler (http://www.nrao.edu/~bbutler/work/nraomemos/VLAwvr.ps) in the context of the VLA need to be carried out for the ALMA circumstances. The existing MRAO/Onsala design uses only 3 but it seems likely that at least 4 would be beneficial to give more information about what is going on in the atmosphere. The bandwidth should increase with increasing offset from the line to give more sensitivity where the water emission is weaker: a possible combination might be 0.5 – 1, 1 – 2, 2 – 4 and 4 – 8 GHz. It is of course possible to make a cross-correlation filter spectrometer to use with a correlation front-end, although twice as many filters are needed in a true multiplication scheme.

The analogue correlator form looks attractive as a compact device suitable for mass production. The existing design is limited to about 4 GHz by the analogue multipliers but faster devices are being worked on. An alternative approach using passive detectors is under development at MRAO for CMB work. Because the frequency spacing is fixed, one would need at least 16 lags to cover plus and minus 8 GHz of IF with adequate resolution. (The BIMA scheme expects to use 32 channels.)

We suggest that the analogue correlator be adopted for further investigation with filters as a safe fallback.

8) Local Oscillator

In order to use DSB systems (or a SSB one with modest rejection) we need to put the LO at the line frequency, 183.31 GHz. First harmonic mixers would require 91.155 GHz, which is easy with a fixed-tuned Gunn. Alternatively it may be more economical to adapt components from the standard ALMA LO system even though the tuning flexibility and pahse stability are not required. Fundamental mode mixers are better because there would be fewer LO harmonics and somewhat lower noise. These could be driven with a Gunn plus a doubler, but would need quite a lot of power, especially for several mixers. Using biased mixers rather than self-biasing ones would reduce this problem.

No tuning is needed, except possibly a step of a few MHz to move it out of the way of a particular line. Although with an SSB system one can in principle fit for the frequency, phase locking the LO to the system clock is clearly advisable, so that all the interference spikes are at accurately known frequencies (and with zero fringe rate).

9) Beam Offsets

It is clearly important that the radiometer samples the same path through the atmosphere as the incoming astronomical signal. It is in fact not possible for these to match absolutely perfectly. (For one thing the radiometer signal is incoherent emission from the water molecules and is therefore sampled by the intensity pattern of the antenna, which is always positive. The path length change is a coherent effect and therefore depends on the amplitude pattern. Molecules in certain locations will not contribute to the phase delay and some will even produce an advance!) The question of how well the beams need to overlap depends on how much small-scale structure there is in the water vapour and how far away it is in front of the aperture. We need more data on the height of the fluctuating layers to make quantitative statements on this.

It is however clear that it is desirable to keep the radiometer close to the astronomical feeds but this is not likely to be a very critical parameter because most of the phase fluctuation is in scale that are considerably larger than the beam. If we can place the radiometer feed in the centre of the ring or cluster of feeds, then the beam offsets are likely to be in the range 3 to 10 arc minutes. This corresponds to distances of 1 to 3 metres at a distance of 1 km, i.e. a modest offset compared to the dish diameter. To illuminate a suitable area on the subreflector to be able to do the pointing correction would require a feed about 75 mm in diameter. It is more likely that a much smaller feed (or group of feeds) would be used which would be reimaged onto the subreflector by an optical relay. The final mirror of this could then be in the central position and it would be advisable to allow about 100 mm clear diameter to accommodate it.

The baseline should be to keep the radiometer beam within 10 arc minutes of the astronomical ones and, if it is practical to do so, make this offset smaller than that for the higher frequency channels.

Conclusions

The critical issue at this stage is to decide whether we should aim to correct the single-dish pointing errors or not. Once that is determined more detailed specifications can be drawn up and design choices made. It is also important for the SAC to consider the issue of whether options involving use of the astronomical receivers should be kept open or ruled out now as an undesirable approach.

Receiver Cryogenic System

A. Orlowska, M. Harman, B. Ellison
Last changed 2001-Jan.-29

Revision History
2001-Jan.-29: First ALMA version

Contents:

List of Figures:

6.1 Introduction

The ALMA cryogenic system performs a critical role within the array operational infrastructure by providing the necessary cooling for all low noise receiver front-ends associated with the instrument. It is vital for successful operation of the array that the cryogenic system provides appropriate thermal cooling capacity and stability, mechanical robustness and a high degree of reliability. Furthermore, the system design must offer a degree of flexibility for planned and future receiver technology development and yet be sufficiently simple to ensure minimum and straightforward construction and maintenance. Achieving these requirements is a demanding development task and one that requires the application of novel design and construction methods coupled with the selection of the most suitable cryogenic cooling engine.

This report provides a summary of the design progress to date and although preliminary in a number of aspects, for example thermal heat loads, mechanical deflections tolerances and dimensional constraints need to be refined, it represents the envisaged structure and design methodology.

6.2 Design Summary

The present preliminary design is able to accommodate ten receiver observational bands operating in the millimetre to submillimetre wavelength range. In addition, sufficient room has been allowed for inclusion of an atmospheric water vapour monitor (for either cooled or room temperature operation) and, should it be necessary, a receiver calibration cold load. The radio frequency and other electronic components that form an individual receiver band are integrated into an autonomous support structure known as a *cartridge assembly*. All ten cartridges are inserted into a single large vacuum vessel (see Figure 6-1) that provides thermal insulation, radiation shielding and cryogenic heat lift, the latter via a close cycle cooling system. Further, the outer vacuum container (OVC) supports external optical components associated with the receiver-antenna optical interface: internal optical components are supported on individual cartridge assemblies. Thermal connection to each cartridge assembly heat sink stage is provided via a novel low resistance thermally activated link arrangement that requires no permanent mechanical attachment, i.e. it does not need to be physically bolted to a stage. This mechanism provides a significant operational advantage in that withdrawal of a cartridge, and hence a complete receiver band, can be simply performed at room temperature and ambient atmospheric pressure without disturbing the rest of the receiver and cryogenic system. This minimises risk of damage to the remaining receiver bands, reduces maintenance time and avoids a potentially lengthy and difficult readjustment of the external optical assembly since this need not be separated from the vacuum vessel. Furthermore, adoption of the cartridge principle allows construction and test of individual receiver assemblies to be performed at separate development facilities prior to final integration into the main vacuum vessel. We believe that this approach will provide a significant advantage to the ALMA receiver development community and is consistent with the likely multinational distribution of receiver development tasks.

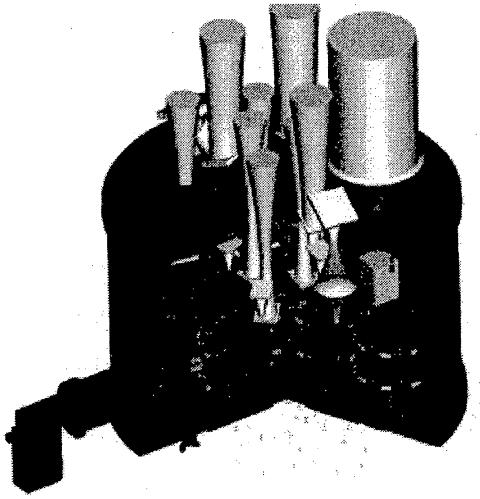


Figure 6-1: ALMA cryogenic system concept showing optics, cartridge and cryocooler arrangement

Individual elements of the cryogenic system conceptual design are described in greater detail within the following sections.

6.3 Cryostat Design Requirements and Objectives

The basis of the cryostat design can be divided into a number of key requirements as shown below. Associated with each requirement heading are a series of design objectives that, when met though the design process, will ensure the construction and operation of a cryogenic system that is well suited for the ALMA instrument.

1. **Efficient millimetre and submillimetre wavelength observation band operation.**

   - *Configuration*: - System must be flexible and able to accommodate various observation band component configurations, including:
     - Optical interface (off-axis mirror, lenses etc.)
     - Superconducting-tunnel-junction (SIS) mixer front-ends
     - IF amplifiers.
     - Local oscillators
     - Low noise amplifier front-ends.
     - Associated cabling and ancillary electronic components.
   - *Signal input windows*: - Accommodate various window sizes and infrared filter designs.
   - *Alignment:* - Maintain RF optical alignment between observational band components and external optic units during cool down and operation.

2. **Sufficient cooling capacity.**

   - *Baseline front–end cooling:* - Cool SIS mixer front-ends to 4.0 K (maximum) and low noise amplifier front-ends to 15 K (maximum).
   - *General cooling:* - Intercept radiative and conductive heat loads from windows, wiring and wave-guides etc.
   - *Stability:* - Provide a 4 K stage short term temperature stability of < 2 mK (rms). Provide a long term temperature drift that does not exceed 0.2 at the 4 K stage for 1 year.

3. **Physical mass and size compliance.**

   - *Mass:* - Total receiver package, including vacuum vessel, cartridges and coolers, must be less than 750kgs (mass estimate suggested by the ALMA Antenna Group) including receiver components and optics.
   - *Size:* - Cryostat OVC must be able to fit though the antenna receiver cabin door. This limits the total receiver package (including external optics) to an envelope size of 1.1m width x 1.6m height.

4. **Suitability for production phase manufacturing.**

   - *Technology:* - Cryostat design should baseline existing and proven technology whenever possible. Novel technologies should be properly evaluated and tested prior to commencement with the production phase.
   - *Mass production:* - Cryostat design should give a high degree of consideration to construction and assembly methods and techniques. Complexity of the design and mechanical structures should be simplified wherever possible.

5. **Observatory operational issues**

- *Evacuation and cool down:* - Total cryostat evacuation and cooldown time (from room temperature) should not exceed 5 days. The cryostat must have the ability to be evacuated to a suitable pressure to allow cooldown consistent within this time frame and to ensure long term operation (> 1 year) on a given antenna.
- *Vacuum integrity:* - Achieve reliable operation = 1 year. This will require a total system vacuum integrity (including signal windows and other feedthroughs) that corresponds to a leak rate of order 10^{-8}mbar litres/sec. We anticipate that implementation of a suitable gettering material will maintain the vacuum level to an acceptable level, typically < 10^{-6} mbar, during normal operation and will minimise the contamination of cold surfaces.
- *Maintenance:* - Cryostat must have a minimum service interval of 1 year. Anticipated servicing to include examination of cryocooler systems (cold-head and compressors) and evacuation of vacuum chamber.
- *Reliability:* - Repeated thermal and vacuum cycling of the cryostat structural components; including cartridges, vacuum vessel, windows etc., should not cause catastrophic system failure. A qualifiable limit of 100 thermal and vacuum cycles is specified.
- *Transportation:* - The cryostat must be capable of being safely transported, e.g., from construction location to observatory site or from a given antenna to operational maintenance centre) either at room temperature or cold. The system should be able to withstand and survive a 3g-shock loading.

6.4 Detailed Cryostat Design

The ALMA cryostat must fulfil the cryogenic requirements of the front-end receiver technology and, in addition, must operate with high degree reliability, efficiency and be compliant with large-scale production. To ensure that the cryostat satisfies these objectives, it is essential that contributing factors that could limit the system performance or be a cause of reliability concern are identified and resolved.

Our design procedure has established the main criteria that affect both system performance and reliability to be associated with:

- Maintaining adequate alignment of the internal cooled receiver components with the external optical assembly.
- Provision and maintenance of sufficient cooling capacity and thermal stability.

Alignment issues are predominantly associated with the ability of individual cartridges to remain co-aligned with the room temperature external optical assemblies as the cryostat is tipped on the telescope. Furthermore, the cartridge structure must be sufficiently rigid to resist distortion during receiver component integration, general handling and subsequent thermal cycling.

Provision of appropriate cooling capacity is dependent upon selection of the most appropriate cryocooler technology and careful evaluation of parasitic heat loads due to radiative, conductive and power dissipating sources. The cartridge structure provides a significant contribution to the conductive load and, as a result, its design is crucially important to successful cryostat operation.

The following sections describe of the major features of the cryostat design by splitting the whole cryogenic system into two main parts namely, the vessel structure, including OVC, radiation shields and cartridges, and the cryocooler. Detail and drawings are those available at the current time of issue of this report..

6.4.1 Outer vacuum container

The conceptual design of the OVDC is shown in Figure 6-2. It essentially consists of a metal cylinder (mid-section) capped at both ends with metal plates that support the signal input windows at one end and the receiver cartridges at the other. With the exception of the cryocooler system, all porthole access ports (window, cartridges, vacuum vale and vacuum gauges) are placed on the surface of the two endplates.

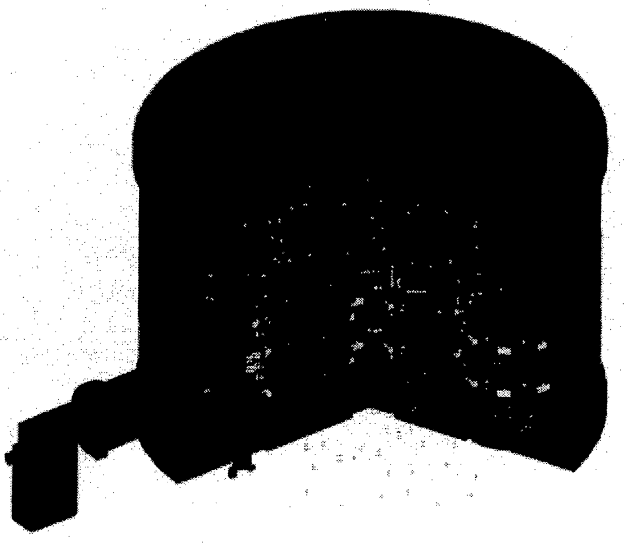


Figure 6-2: Cut-away view of outer vacuum container (OVC)

The 'top' endplate has 12 portholes with a suitable 'O' ring seals that provide locations for the input signal windows, water vapour radiometer (signal input window if cooled or cold load window if warm) and, if required, a generic receiver cold load. The base of the vessel has 10 'O' ring sealed interfaces for the cartridges and four additional ports for vacuum pump and vacuum gauge attachment, thermometry and heater electrical cabling feedthroughs. The envelope dimensions for the OVC are approximately ø1m x 0.7m high, and are consistent with the antenna cabin access door. Figure 6-3 shows the bottom end plate with the cartridges extended from the OVC.

Figure 6-3: Bottom view of cryostat with cartridges extended

The mid-section provides the radial support structure of the vessel and contains the suspension supports. It also provides 'O' ring sealed access for the attachment of the cryogenic cooler. A central internal tube has been introduced between the two end caps (see Figure 6-4) in order to

reduce the endplate deflection when the system is evacuated. Figure 6-5 shows the extent to which a simple endplate structure (holes excluded) deforms without the addition of the tube. Additional metal ribbing is also included to reduce the residual deflection to < 0.1mm. It is anticipated that addition of the window and cartridge access holes will weaken the endplates, particularly at the cartridge end, and some additional mechanical stiffening may be required in order to achieve the required deflection. This is currently being investigated.

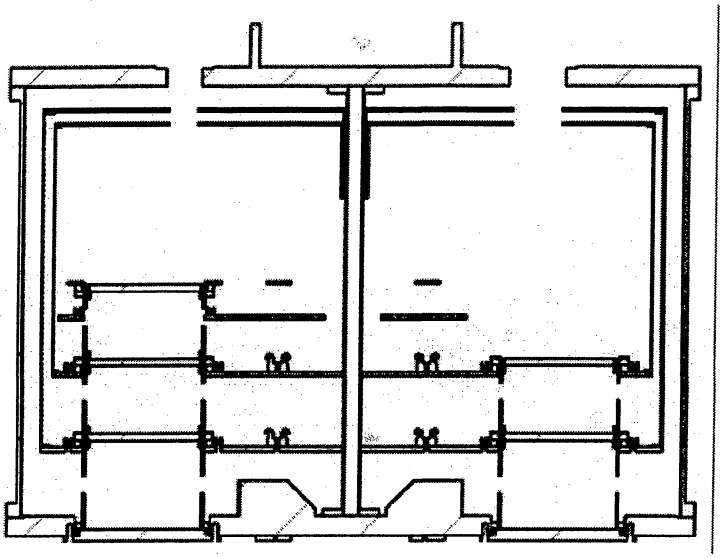


Figure 6-4: Cross-sectional view of cryostat showing internal support bar

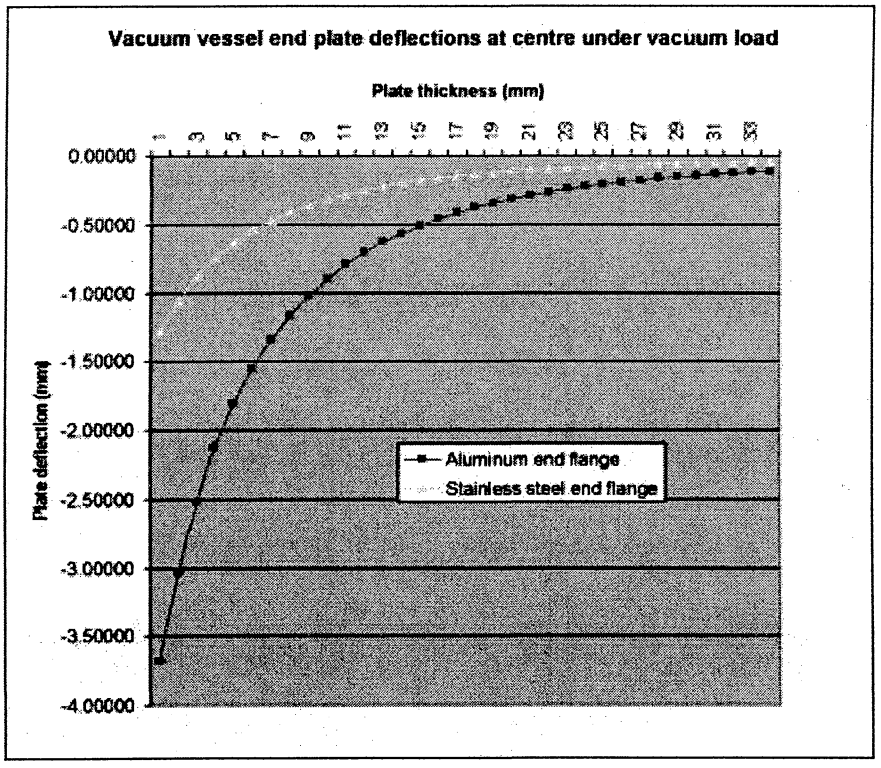


Figure 6-5: Unsupported Vacuum vessel endplate deflection

The OVC construction material can be either stainless steel or aluminium and in both cases fabricated in accordance with a pressure vessel design code BS5500. The final material choice requires detailed structural analysis that is currently ongoing, and selection will most likely be a compromise between structural mass and rigidity rather than cost. Additional consideration must also be given to the ability of the chosen material to provide a low surface emissivity, low outgassing properties and ruggedness. This normally implies a preference for stainless steel, but the total cryostat mass restriction of < 750 kg means that material thickness must be minimised and resulting deformation carefully evaluated. We have also considered the use composite materials since they may provide a significant mass advantage. However, our concerns with regard to their suitability, they are not proven in this area and therefore introduce production and performance uncertainty with little cost benefit, lead us to reject their use at the present time..

Attachment to and accurate alignment between an antenna and cryostat is achieved by use of precision registers and dowel pins located at appropriate intervals on an interface ring. Our intention is that registration between an antenna and cryostat be non-specific. That is, individual cryostats can be interchanged between any of the array antennae and external optics assemblies. The will greatly ease assembly, test and maintenance requirements though, for a cryostat of the size conceived, will not be a trivial task. Final definition of cryostat – antenna interface alignment tolerance is currently awaiting outcome of analysis from the ALMA Optics Group.

6.4.2 Internal radiation shields (70 K and 12 K)

The function of the cryostat internal radiation shields is to reduce the radiative thermal load on the 4 K stage. An addition function is to provide a good heat path between the thermal link arrangement and the cryocooler heat lift stages. The shields are constructed of pure aluminium (BS1470 grade 1200 or equivalent) and provide good thermal conduction properties, relatively low outgassing, low mass, ease of manufacture and low cost. In order to reduce the emissivity and outgassing rate further, a specialist surface treatment using chemical cleaning will be used to clean the surface. An acid etching technique, which has been proven in a production environment, is a cost-effective surface treatment method. Once clean, however, surfaces will be prone to contamination and it is therefore essential that careful procedures are employed, preferably in clean room environment, during assembly and maintenance.

A substantial load has been calculated to be incident upon the 70K shield from the room temperature OVC. Even though the cryocoolers specified in Section 6.5.0 provide a large heat lift capacity, this amount of radiation represents a significant load on the cooling system and may ultimately impact cooling effectiveness and reliability. A standard technique to reduce this load is to apply several layers of multi-layer insulation, often interleaved with polyamide netting to reduce thermal conduction, to intercept the radiative heat load. Although the use of multi-layer insulation will increase the evacuation time, our initial estimate of the reduction in radiative heat load onto the 70 K shield (typically to by 20 W) leads us to believe that its introduction is well worth while. A similar reasoning applied to 12 K shield indicates that a single layer of aluminium foil 0.08mm thick would be sufficient to reduce the radiative heat load from the 70 K shield to acceptable levels. The aluminium foil is a cost effective method of providing a surface with a very low emissivity without the use of more expensive polishing or plating processes.

The effectiveness of the radiation shields is also a function of how they are supported. For example, it is essential that the support structure implemented does not introduce an excessive thermal conduction path since this would raise the temperature of the shield either throughout or in local areas. However, it is also essential that any support structure be sufficiently rigid to prevent excessive displacement of the shield as the cryostat is tipped on the telescope. Although the shield structure is connected to the cartridge via highly flexible links (see Section 6.4.4.3), alignment must still be maintained within ±0.5 mm in order avoid to the introduction of additional cartridge deflection. Clearly, achieving tolerance of this order require the minimisation of the shield mass.

6.4.3 4 K heat sink stage

The 4K plate construction is of machined OHFC copper plate. OHFC copper has been selected primarily due to its properties as a good thermal conductor at 4K but also because of its low outgassing properties, ease of manufacture, availability and cost. The grade selected will be to

BS2870 grade C103/110 or equivalent. The plate will have features for attachment of thermal links, low conductivity supports, and suitable cooling engine connections. The overall shape must be optimised in order to provide adequate thermal conductance at 4K. Consideration also needs to be give to the mass of the plate and this should be minimised in order to provide increased cooldown efficiency.

6.4.4 Cartridge design

6.4.4.1 Benefits

Adoption of a receiver cartridge philosophy provides the following benefits:

- Relatively large working volume that can accommodate a variety of receiver configurations.
- Standardisation providing reduced production and assembly costs and minimising production lead times since a substantial fraction of the structure can be subcontracted to industry.
- Ease of receiver assembly, integration, and maintenance.
- Individual RF receiver bands cartridges may be assembled and tested independently to main cryostat. This will reduce potential 'bottle-necks' in the cryostat production and assembly phase.
- Minimising internal cable and harnessing inside the OVC since all individual signal band waveguide and electrical connections are located on the base of a cartridge.
- Minimised shield thermal conduction loads. The cryostat radiation shields are not required to support receiver components. This significantly reduces the shield heat load and alignment requirements since only individual cartridges must achieve the necessary alignment tolerances.
- Flexibility for future receiver upgrades.

Maintaining the optical alignment between observing band components and external optical units during transportation cool down, and operation is of a fundamental importance to the proper operation of the receiver system. All radio frequency (RF) optical components have a requirement to be mounted on mechanically stable structures. The mechanical loads have been identified which could adversely affect optical alignment are deflections of the vacuum vessel after evacuation, thermal contraction during cool-down, tilting of the antenna, transportation of the receiver and the interchange of production receiver components during maintenance, or complete front-end replacement.

6.4.4.2 Configuration and accommodation

The cryogenic system is capable of support a total of 10 individual cartridges. Each cartridge contains all the necessary components, ancillary electronics and cabling associated with a specific front-end band. There are 8 cartridges with an outside diameter of 170mm and 2 cartridges with an outside diameter of 140mm. Ideally, cartridges would be identical in shape, but space limitation within the OVC has prevented this. In Figure 6-6 we show a example of a typical cartridge structure and in Figure 6-7 the corresponding cartridge location on the OVC endplate.

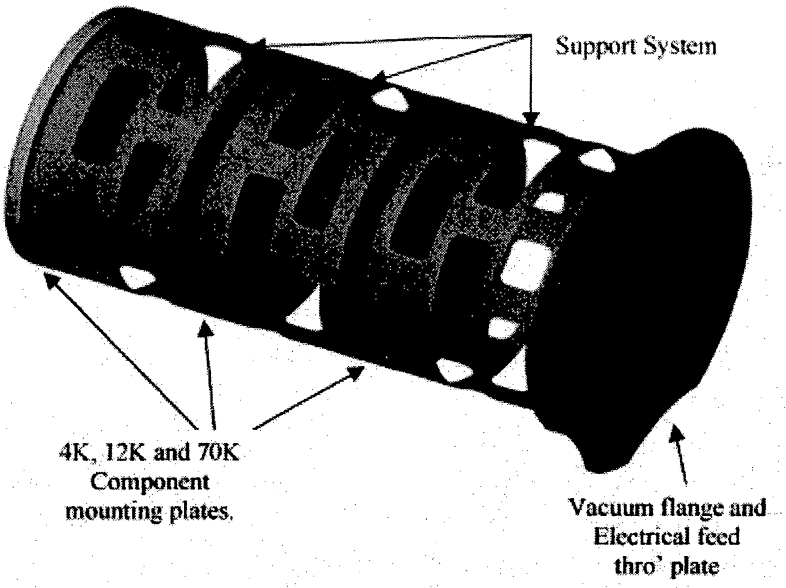


Figure 6-6: : Example of an ALMA receiver cartridge

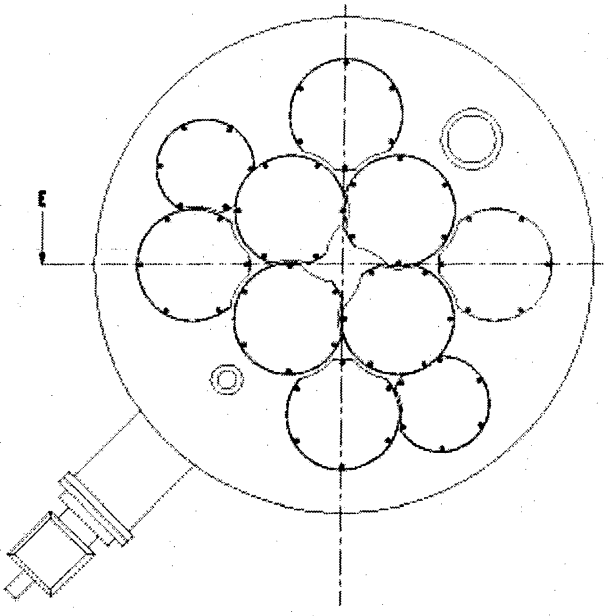


Figure 6-7: : Cartridge location on OVC endplate

The structural form of a cartridge comprises a room temperature base plate, that mates with the OVC, to which is attached a series of thermal insulators that separate three cold surfaces, referred to as the 70K, 12K and 4K stages. In the case of receiver bands 1 and 2, a 4K surface in not required and can be simply omitted from the cartridge assembly during construction. All necessary electrical connection and feedthroughs are located on the room temperature base plate making each cartridge and autonomous assembly. The thermal isolation between cold stages is accomplished by use of thin walled fibre glass tubing that has been optimised for thermal resistance and mechanical rigidity. Figure 6-8 indicates a preliminary finite element (FE) analysis of the cartridge structure: the smaller diameter indicates the distribution of the 4K end load (mirrors, mixers etc.) and increasingly blue area indicates the region of greatest deflection. Figure 6-9 indicates the corresponding conductive heat flow through the corresponding tube structure.

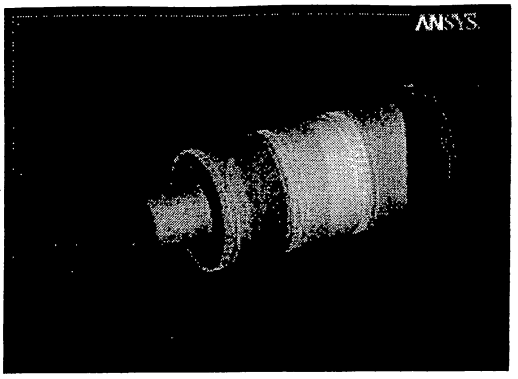


Figure 6-8: : Illustrates FE model of estimated cartridge deflections when cryostat axis is horizontal

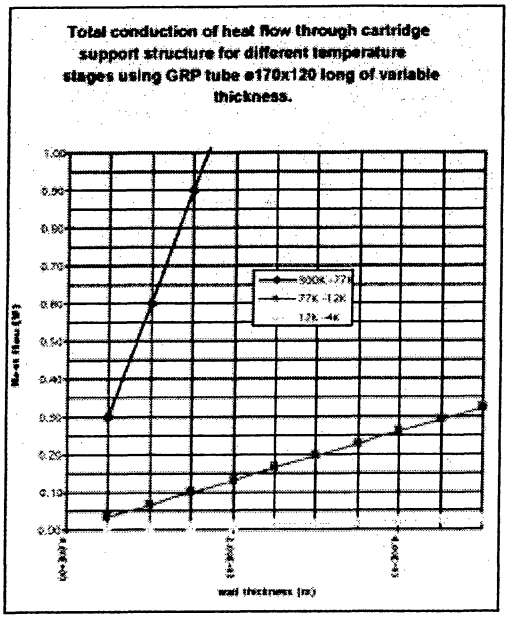


Figure 6-9: Illustration of cartridge support conductive heat flow

The concept of the cartridge assembly is very similar to that employed in hybrid helium dewars used at a number of observatories for millimetre and submillimetre wave system. Further, it benefits from the heritage gained of a similar, though smaller system, successfully developed and employed by NRAO on the 12m Kitt Peak antenna. However, the current design differs from past systems in one important aspect namely, the thermal anchor or link arrangement. In this case, a novel mechanism is proposed in which connection between a cartridge cold stage and the cryocooler heat sink point is achieved via a temperature dependent thermal link.

6.4.4.3 Cartridge thermal link arrangement

An early version of the link, which has undergone preliminary tests, is shown in Figure 6-10. The basic mode of operation involves the thermal contraction of a nylon and copper ring assembly surrounding a specific cartridge cold stage, as the cryostat temperature is reduced. The cartridge assembly passes through the ring, which is a relatively loose fit, and is self-aligned with the appropriate cartridge temperature stage. On system cool down, the nylon ring contracts at a greater rate than its metal counterpart and, as a result, squeezes the metal ring into contact with the cartridge and thereby forms a thermal link. Tolerance between the link inside diameter and stage outside diameter are selected such that when cooled operation is achieved a substantial force is exerted between the surfaces and a low thermal resistance joint is formed. An estimate of the clamping force produced by this arrangement is shown in Figure 6-11: a clamping force in excess of that obtainable from a conventional bolt arrangement is predicted. A series of radial flexible braids attach the link to the cryocooler cold plate and allow free mechanical movement of the link.

This is necessary to avoid distortion of the cartridge during cooldown and any deflections that may arise from movement of the radiation shields during use on a telescope. The clear advantage of this system is that it provides minimum mechanical intervention to insert or remove a cartridge assembly since when the cryostat is a room temperature, the link compression is relaxed and becomes free of the cartridge stage. In addition, the system also allows cartridges to be placed closer together and results in a more compact cryostat envelope.

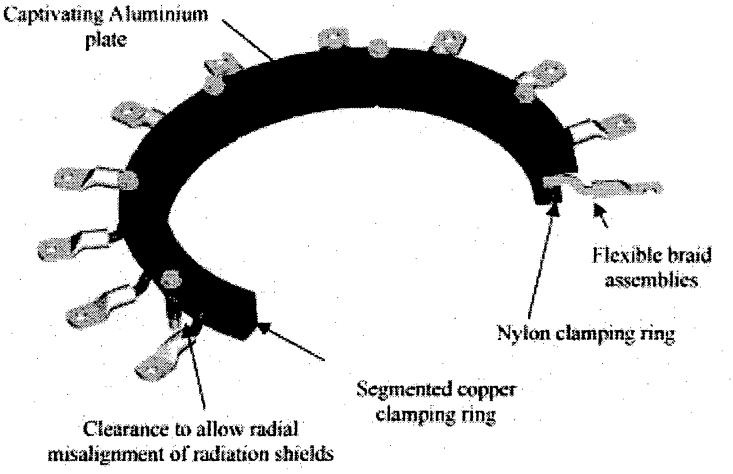


Figure 6-10: Preliminary design of the ALMA cryostat thermal link arrangement

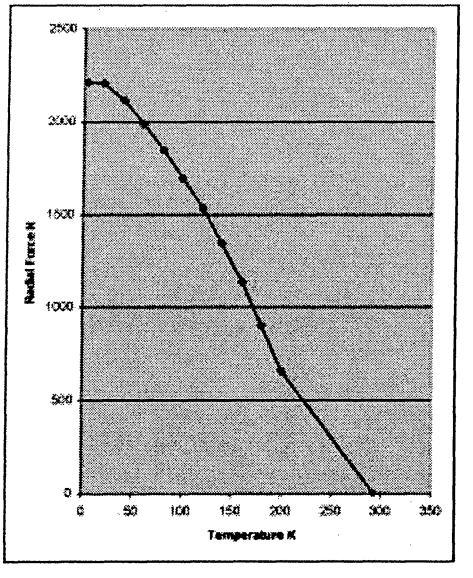


Figure 6-11: : Estimated thermal link clamping forces when cold

A thermal link similar to that shown in Figure 6-10, but reduced in scale, has been manufactured and tested. Measurements indicate successful operation and repeatable performance: the link was thermally cycled 10 times between sets of measurements with no apparent degradation in function. Table 1 shows the measured thermal gradient across the link which, although limited by the test arrangement, appear to be acceptable at 12 K and 70 K. The conductivity at 4 K needs to be improved if heat loads higher than 30 mw are anticipated. A full report on the tests is available[1] Improvements to the design have been made, reducing the number of interfaces, and a new link will be tested.

[1]ALMA Receiver Cartridge Thermal Link Test Results
A report on the experimental test results of the proposed receiver cartridge cryogenic thermal link for the ALMA cryostat
MC Crook RAL October 2000

Table 1: Thermal conductance and temperature gradients at different heat loads and stage temperatures

Temp (K)	K (W/K)	Load (mW)	?T (K)
4	0.12	40	0.33
		100	0.83
12	0.74	40	0.05
		100	0.14
80	1.3	40	0.03
		100	0.08

Achieving adequate thermal conductivity of the links at 4K is crucial. The temperature distribution across the link must be maintained to ~ 0.2K to avoid degradation of the superconducting mixer performance. This reduces the link material selection for the link to high purity metals, e.g., aluminium or copper. Since flexible copper links are readily available commercially it was felt that their use would provide the most cost-effective selection for the flexures. Furthermore, good connection of the link to the radiation shield base is essential is cartridge stages are to be properly cooled. Because of the large number of links that are required on each stage, space is at a premium and must be therefore used very efficiently. An arrangement shown in Figure 6-12 is proposed for link distribution on the radiation shield (and 4K stage) plates.

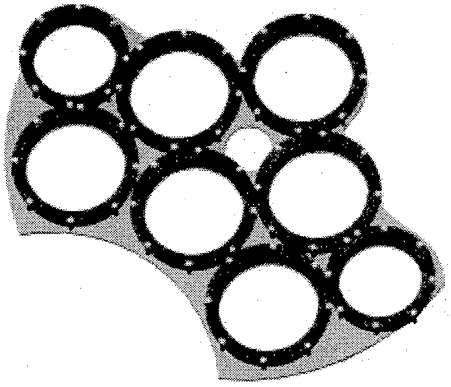
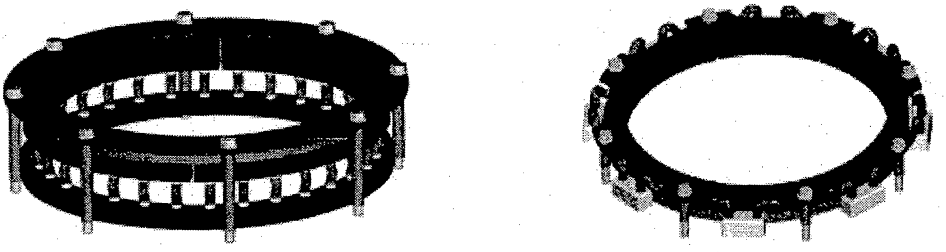


Figure 6-12: Distribution of thermal links on the 4K base plate A similar arrangement is used for the radiation shield links

Due to the number of thermal link assemblies required (provisionally 28 per cryostat) the links must be compliant with existing commercial production techniques and large-scale manufacture. For example, all components in the link assembly should be capable of manufacture using conventional milling, turning, brazing processes or laser/water cutting techniques. Several flexible thermal braids have been considered. The structure selected is provided by a local expert cryogenic company (Oxford Instruments) and has a great deal of production and cryogenic heritage (approx. 250,000 in service). Estimated cost per braid is approximately $1.5 each. It is proposed that a vacuum brazing technique be employed for attachments of the braids into a copper clamping ring. Although vacuum brazing is a relatively expensive process for 1 off production due to the high cost of heating the oven, for volume production there are significant cost savings. The flexible braid has a crimped end that eliminates the wicking effect of the solder that will occur during brazing or soldering. An additional benefit of brazing is that no expensive post manufacture cleaning is required; conventional soldering generates large oxide layers which require cleaning which can often be a lengthy and unsatisfactory process and can leave residual outgassing components.

The utilisation of laser/water cutting is proposed for the production thermal links where precision machining is not a requirement on non-critical surfaces and can achieve tolerances to approx. 0.3mm. Laser/water cutting is a cost-effective process for cutting profiles in volume component production since it is automated, quick and with a minimal amount of set-up time. It is becoming common place in most metal stockholder or machine shops.

An updated design for the thermal link arrangement is shown in Figure 6-13. We intend, through collaboration with the NRAO, to manufacture and test this new version prior to the cryostat production phase. The new design is of a larger link, suitable for all the 170 f cartridges and has eliminated one bolted joint per braid. This should lead to an increase in conductivity.



4 K link *12 and 70 K link*

Figure 6-13: Modified thermal link arrangement

6.4.4.4 Cartridge manufacture and assembly issues

As previously indicated, the cartridge is composed of 3 main sub-assembly sections (77K, 12K and 4K stages), vacuum plate, and thermal isolation support structure. The vacuum plate will be manufactured from stainless steel. Associated waveguides and electrical interfaces will be required to be integrated onto the base of the vacuum plate and stainless steel is the most suitable material for this. The grade selected will be to BS1501 grade 304S12 or equivalent. This grade has been selected for its high strength, machine and weld ability, corrosive resistance, cost effectiveness, and availability. The cold plates will be manufactured from pure aluminium The grade selected will be to BS1470 grade 1200 or equivalent as for reasons previously stated. The cold plates can be modified to suit individual RF band design configurations and if required additional fixing holes and cut-outs may be added. An appropriate surface finish will be required around the periphery of the cold plates to ensure the thermal link assemblies can make good thermal contact thereby minimising joint thermal resistance. The cold plate support structure will be manufactured from GRP (Glass re-enforced plastic). The tubes will be slotted to increase the thermal path and the slots optimised utilising finite element analysis techniques to provide a suitable trade off between rigidity and thermal resistance. It will be important to control the manufacturing process and quality of the tubes since orientation of the glass fibres and resin epoxy significantly affects the thermal and mechanical qualities of the composite material. In addition, construction and assembly of the whole cartridge will need to be carefully specified and monitored since specific tolerances must be achieved in order to ensure proper function of the thermal link and adequate alignment and interchange of the assemblies.

6.5 Cryocooler selection

6.5.1 Cooler requirements

Cooling the SIS mixers to = 4.0 K and achieving appropriate temperature stability is a critical objective if optimum receiver performance if to be attained and maintained and is a function of the overall system efficiency. In order to achieve good cooling efficiency we must first identify all sources of significant heat input. With the ALMA cryostat the areas that contribute to the system thermal loading are:

- Radiation from the surrounding environment.
- Conductive heat flux from mechanical support structures, electrical wiring, and wave-guides.
- Power dissipation from electronic components that form part of the internal receiver system.

An additional radiative heat source is from the signal input windows (see Figure 6-14). This heat load will be intercepted at the different heat stages utilising a method of infrared (IR) radiation filtering. Each IR filter corresponds to a signal observing band and will be optimised to ensure that

it does not compromise receiver system sensitivity and hence observing performance. At present the IR filters designs for the ALMA IR filters are unavailable and thus, only budgetary estimates based on the past experience of receiver development groups are utilised.

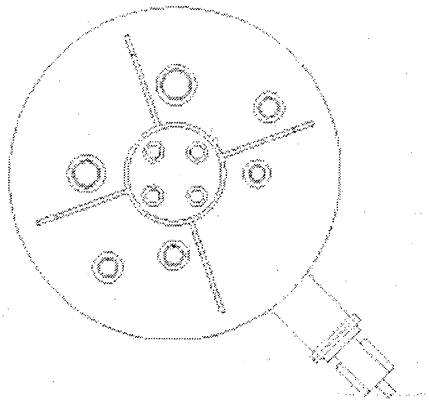


Figure 6-14: Top view of cryostat showing signal input window layout

The prime design driver for the support structure is to ensure that RF optical alignment is maintained but as for the reasons stated earlier it is beneficial to minimise the heat leak transfer onto the cold stages. Due to the conflicting nature of these requirements optimisation of the supports will be required and a suitable trade off between rigidity and thermal conductivity made. A more detailed analysis will be conducted using finite analysis techniques that give high degree of accuracy in predicting both mechanical stability and thermal conductivity and so providing a high level of optimisation

In addition to satisfying the cooling requirements for ALMA, selection of the cooler must include the additional following considerations:

- Reliability.
- Low vibration.
- Low input power including ease of cooling the compressor at high altitude.
- Serviceability, including clearly defined service intervals straightforward serviceability, the latter preferably without significant disassembly of the cryostat.
- Be affordable.

6.5.2 Heat loads

The cryostat thermal models are preliminary and further analysis is underway in respect of some parameters such as filtering, mechanical supports, local oscillator dissipation etc., are indicated in Table 2 (all TBC).

Table 2: Estimated maximum cryocooler heat lift requirement

Temp. Stage	Loading	Total Stage Load (Max.)	Comment
70K	26 W + support + LO +IR loads + margin	40 W	Plus extra if JT used
12K	5 W + support + margin	8-10 W	Plus extra if JT used
4K	0.6 W + support + uncertainty in IR loads	0.75 - 1 W	Cooling required at T = 3.8 K

6.5.3 Number of stages

Depending upon the type of cooler selected, two configurations can be considered, (each with several possible means of realisation)

- A 3 stage with cooling stages at 70 K, 15 K and 4 K.
- A two stage with cooling at about 40 K and 4 K.

The cryostat will be large with extended surface areas incurring large radiative loads. Many decisions, yet to be confirmed, will affect the thermal loads (number/size of windows, use of Multi Layer Insulation, optical design etc) and a margin on cooler performance will be required. It is much more thermodynamically economical to intercept heat loads at about 70 K where cooling is relatively cheap, rather than at a lower temperatures. Although input power is not the primary consideration in selection, the problems of heat dissipation at high altitudes have to be borne in mind. For these reasons the 3 stage system is preferred.

6.5.4 Cooler type

A 3 stage system could either be implemented by one cooler type, e.g. a three stage Gifford-McMahon (GM) or pulse tube, or by a two stage cooler with the coldest stage provided by a further stage of a different type, e.g. a two stage pre-cooler with a helium JT stage. Both systems have some advantages; simplicity and ease of integration in the case of a single cooler, distributed cooling and efficiency in the case of the JT option.

Both GM and pulse tube coolers operating at low frequencies suffer considerable loss of cooling when run in orientations more than about 30 degrees from optimum (cold finger pointed down). This is a problem at all temperatures but more significant at 4 K. GM coolers lose about 10% of their cooling power when horizontal, pulse tubes probably more.

6.5.5 Temperature

A maximum temperature of 4 K at the receivers is a requirement. Some margin in the temperature is necessary in order to provide for heat transfer at the receivers and hence sub 4 K operation is required at the cryocooler cold finger. To achieve a temperature of 3.8 K, however, a JT system would require an exhaust pressure of less than 0.657 bar, lower temperatures require lower pressures, but the temperature remains constant with heat lift (until the maximum heat lift is reached). This is a challenging requirement.

A GM or pulse tube cooler would have a base temperature below 4 K, but the temperature would be dependent on heat lift, and possibly orientation. Currently available two stage GM coolers have base temperatures below 3 K.

Heat transfer at 4 K is an important issue. A JT stage offers distributed cooling with potentially shorter conduction paths to the receiver channels. This advantage must be weighed against the additional complexity of the cold plumbing. A 4 K or GM cooler gives single point cooling and heat must be conducted from each of the channels. For a large system this is disadvantageous but such methods have been used successfully on smaller systems.

6.5.6 Temperature stability

Temperature stability requirements at 4 K are extremely stringent, around 2 mK peak to peak over one minute (TBC). The temperature of both GM and pulse tube coolers is orientation dependent so such coolers would be used with their axis parallel to the elevation axis.

Temperature stability of JT systems is good, as the temperature only depends on the exhaust pressure at the orifice. To achieve similar stability with a 3 stage cooler high heat capacity such as a rare earth (or a helium reservoir) would have to be introduced at the 4 K plate.

6.5.7 Cycle

Both GM and pulse tubes can be considered. The GM cooler is well established commercially with a number or manufacturers and known reliability. Such coolers are produced in very large numbers for the MRI and cryopumping markets. Service intervals of about 10,000 hours for the expander and 20,000 hours for the compressors are quoted although anecdotal evidence points to longer life. The main source of wear problems is the sealing of the displacer unit and the valves. Vibration levels are relatively high, but such coolers have been used successfully on telescopes and in the laboratory. The vibration levels of mechanically driven expanders are significantly lower than those generated by pneumatically driven systems.

Pulse tubes are becoming established commercially although fewer manufacturers are involved. The compressor units used are the same as those for GM coolers and will have the same service intervals. The pulse tube itself has no moving parts and the valves cans be mounted at a slight distance (say 30cm) from the cold unit, simplifying servicing and greatly reducing any vibration in the cryostat. Pulse tubes are less efficient than GM coolers and a pulse tube would require a larger compressor than a GM cooler of equivalent heat lift.

The original baseline of two stage PT (70K, 12 K) and a 4 K JT appears to be receding due to the uncertainty of a commercial PT of adequate heat lift being available in time for the ALMA procurement. The major cooler manufacturers are all working on PT development to replace GMs in their largest market – MRI magnets, unfortunately no suitably large systems are yet on the market, though they may be available within the next 1-2 years. However, our design is sufficiently flexible to allow use of a variety of cooler types. These include:

- Large 3 stage 4 K GM from Sumitomo. Designed for 1 W at 4K, 10 W at 15 K and 40 W at 70 K. Complete with He pot at 4 K for temperature stability. At present few (one?) have been made and it is not clear whether it would have sufficient heat load at a sufficiently low temperature or whether it would do so if horizontal. Uses standard 7.5kW compressor.

- 2 stage 12 K GM + JT system. A number of manufacturers, certainly both Sumitomo and Leybold could provide the GM but we would have to design the JT. The GM would require a 7 kW compressor; the JT system would need a smaller one. On manufacturer (Daikin) provides a complete system.

- 2 separate cold heads, one to provide 4 K and possible a 40 K shield. One to provide the 12 K and 70 K cooling. This system would most likely require 2 compressors in order to maximise the 4 K cooling power. It would be possible to replace one or both of the cold heads with PTs, either now, or eventually, as the systems become more commercially available.

The current options for cooler selection are shown in Table 3 and depend on the priority of requirements. For example,

- is cost more important than input power,

- how much emphasis do we place on temperature stability over performance margin,

- how do we trade off cooldown time vs. cooler cost?

However there are some overriding issues that include

- Performance – The cooler has to give the required amount of cooling at 4 K – there is some leeway in the temperature of the other stages. Not only does the total heat lift at 4 K have to be met, but we must be able to maintain base temperature, and achieve 4 K on the cartridges. This implies an actual cooler temperature of below 3.8 K

- Cost does not just involve the cost of the cooler components but also the impact on the cryostat design, servicing, mass manufacture, input power etc
- There should be sufficient margin at 70 K to allow for vacuum degradation over time.
- Reliability of the cooler system must not jeopardise the project.

Table 3: Summary of available and suitable cooler technology

System	Advantage	Disadvantage	I/P Pwr	Cost*
3 stage GM	• Simple • Lowest power • Commercially procured • Only one compressor • Smallest volume in cryostat • Use He pot for temperature stability • Lowest mass	• Only one supplier • Little performance margin • No reliability/heritage	7.5 kW	£30k?
2 stage GM+ JT	• Several GM manufacturers • Medium power • JT to our design • Good temperature stability at 4 K • Easy to distribute cooling • Complete system (Daikin) available • Lowest vibration at 4 K	• Complex • Little margin in JT heat lift • JT head loads impose heat penalty on other stages • 2 compressors • Risk of JT blockage • More servicing (compressors) • Heat switches required • High mass	~ 9kW	£30k + JT devel. costs Possibly more expensive cryostat
2 GMs (or PTs)	• Performance margin • Intermediate shield • Commercially available • Several manufacturers ∴ known heritage • Fast cool down	• 2 Large compressors • More servicing • Large volume of cryostat required • Needs He pot or other temperature stabiliser • Higher vibration levels	=12 kW	£ 35 K?

* Note all prices are approximate, and we anticipate a reduction for bulk orders

6.5.8 Cryocooler selection summary

The option of two cold heads offers the most margin in cooling power and is the system about which we have most information on reliability. We would have an extra cold stage to provide shielding and relieve the pressure on the other system. The additional cost of coolers would be outweighed by the simplicity of the system. This option would have by far the fastest cool down time

The single 3 stage GM option is attractive for its simplicity and low input requirements but at present the actual performance is not confirmed to meet the ALMA requirements.

The JT option would offer the best temperature stability, but the performance may become marginal if the 4 K heat lift requirement rises. Lower input powers are militated by higher complexity. Compressor availability is uncertain.

6.5.8.1 *Baseline selection*

The optimum solution would be the 3 stage GM cooler provided that this can be demonstrated to have sufficient performance with margin (and in a horizontal orientation). A viable alternative would be to use the two cold head system. Although this is less desirable from the cost and complexity point-of-view, it offers guaranteed heat lift using currently available technology and we therefore believe that it should be maintained as a cryocooler option. The use of a J-T system is not precluded from out design, following recent consultation with cooler manufacturers we have concerns about availability and reliability.

6.6 Production and construction

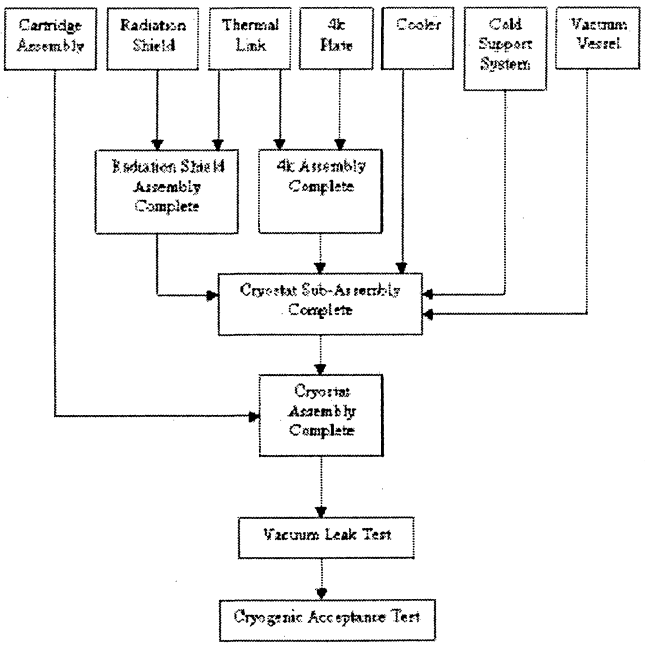
6.6.1 Issues

Issues relating to future cryostat large-scale construction and production have yet to be fully evaluated. However, some immediate points that we consider worthy of mention and consideration include:

- Selection of appropriate production site(s). Large scale facilities required including clean room environment.
- Creation of skilled production and assembly team.
- Appropriate selection and monitoring of constructional materials.
- Division of cryostat system into sub-components for outsourcing. Will require evaluation and selection of appropriate manufacturing companies.
- Construction and assembly quality control at production site(s) and within industrial sub-contractors.
- System test and evaluation plan.
- Integrated cryostat-receiver test and evaluation - where is this best performed?

It is essential that these and other potential concerns that may arise within Phase 1, are resolved speedily if the ALMA cryogenic system is to be produced within a timescale consistent with Phase 2.

6.6.2 Proposed method of assembly



The above flow diagram indicates a suggested methodology for effective production and evaluation of the ALMA cryostat system for the production phase. Completed receiver cartridge assemblies (provided by the receiver groups) would be integrated and tested with the cryogenic system.

6.7 Performance summary

Currently not available.

6.8 Compliance table

Currently not available.

ALMA Project Book, Chapter 7: LOCAL OSCILLATORS

L. D'Addario, editor. Contributors: R. Bradley, E. Bryerton, R. Sramek, W. Shillue, S. Thacker.

[Draft 2000-Dec-08. Minor revisions 2001-Feb-05.]

(Figures and Tables are numbered separately within each major section. Reference citation numbers are uniform throughout with references collected at the end.)

7.0 REQUIREMENTS AND SPECIFICATIONS

The local oscillator subsystem is responsible for establishing all of the time and phase synchronization in the array, on scales ranging from 48 msec (20.8 Hz) to << 1 psec (1 THz). It is also responsible for generating the sinusoidal signals necessary for converting the received signals from RF to IF to baseband, and for tuning these as required to establish the desired sky frequency and interferometer phase (including fringe tracking, phase switching, and other interferometer-specific features). The latter are more properly known as "local oscillator" signals, but the subsystem must also supply various coherent references to other devices so as to achieve synchronization and accurate timing. These include digitizers, computers, and the correlator. It does this by distributing periodic reference signals derived from a common master oscillator.

The LO subsystem also forms part of the array master clock, in cooperation with a computer of the monitor-control subsystem. It does this by providing an interface to an external time scale (currently GPS) and by measuring the difference between external time and array time. Measures of time larger than 48 msec are obtained in the MC system by integration. Further details are given in a later section.

Table 1: Specification Summary

Item	Specification	Goal (if different)
Frequency Range, 1st LO	1st LO: 27.3 to 938 GHz (see Table 2) 2nd LO: 8-10 and 12-14 GHz	
Output Power	1st LO: band dependent (see Table 3) 2nd LO: +10dBm ea. to 2 converters.	100 μW
Sideband Noise, 1st LO	10 K/μW	3 K/μW
Amplitude Stability, 1st LO	.03% <1s; 3% between adjustments	.01%; 1%
Phase Noise (>1 Hz)	63 fsec (18.9 μm)	31.4 fsec (9.4 μm)
Phase Drift (<1 Hz)	29.2 fsec (8.8 μm)	6.9 fsec (2.1 μm)
Tuning step size, maximum	On the sky: 250 MHz SIS mixer 1st LO: 500 MHz	
Subarrays with independent tunability	TBD (3 or more)	5
Simultaneous different sky frequencies	1 per subarray	
Time for frequency change, maximum	Within .03% (freq switching): 10 msec Otherwise: 1.5 sec	1 msec 1.0 sec
Repeatability	1. Phase-unambiguous synthesis 2. Stability specs apply across frequency changes.	

The main specifications for the LO subsystem are summarized in Table 1, and some of them are discussed in more detail below.

7.0.1 Frequency Ranges

Table 2 shows the first LO frequency range for each band. The RF receiving band specification is based on [4]. The LO tuning range specification is the range in which the LO will provide appropriate mixer drive power for the heterodyne receiver of that band for an IF band. The numbers in Table 2 assume an IF band of 4-12 GHz. For the HFET receivers, low-side LO is used for band 1 and high-side for band 2. For the SIS receivers, both sidebands are assumed to be accessible (DSB or sideband separating); therefore, the LO range is 12 GHz inside the RF range on each end. If some SIS receivers have a maximum IF less than 12 GHz or have one sideband inaccessible, the LO range must be increased to cover the full band; this may be difficult to achieve. For band 3, the range shown is for an SIS receiver under the same conditions. If an HFET receiver is used, a very different configuration will be necessary, probably involving two conversions to IF; this is not shown in the table.

Table 2: First LO Frequency Range

Band #	RF Band, GHz	Front End Type	1st LO Band, GHz
1	31.3-45	HFET	27.3-33
2	67-90	HFET	79-94
3	89-116	TBD	101-104
4	125-163	SIS	137-151
5	163-211	SIS	175-199
6	211-275	SIS	223-263
7	275-370	SIS	287-358
8	385-500	SIS	397-488
9	602-720	SIS	614-708
10	787-950	SIS	799-938

The second LO converts from either the lower half of the IF band (4-8 GHz) or the upper half (8-12 GHz) to "baseband" at 2-4 GHz. High-side LO is used in either case, so as to avoid spurious responses associated with the second harmonic of the LO. This requires 8-10 GHz for the lower half and 12-14 GHz for the upper half. There are 4 baseband channels provided for each polarization, so 4 separately-tunable second LO synthesizers will be provided, one for each polarization-pair of channels. To keep all modules the same, a synthesizer that tunes the whole range 8-14 GHz is planned.

7.0.2 Output Power

The local oscillator must provide adequate mixer drive power for both HFET and SIS based receivers. A conventional balanced mixer used in a millimeter-wave HFET front-end requires approximately 5 mW of LO power. However, 20 mW may be required if a sideband-separating mixer follows the low noise HFET amplifier.

The LO power required for SIS mixers will depend upon several factors. Based on the theory of Tucker and Feldman [7] the required LO power is given by

$$P_{LO} = \frac{\left[\frac{N_j h f \alpha}{e}\right]^2}{2R_n}$$

where N_j is the number of junctions, h is Plank's constant, f is the operating frequency, α is a parameter that characterizes the normalized level of LO amplitude across the SIS junction and is usually set to unity, e is the electron charge, and R_n is the normal state resistance taken here to be approximately 20 ohms [8]. These requirements are given in Table 3. These values are supported by measurements on practical mixers.

Table 3: First LO Power Requirements

ALMA Receiver Band	LO Tuning Range [GHz]	Type of Receiver Front-End	Number of SIS Junctions	Minimum Required Mixer Power	Required Power at Input of -20 dB Coupler of SIS Mixer	Required Power at LO port of a balanced, sideband-separating Mixer	LO Power Specification of 50% Over Worst-Case
1	27-33	HFET	---	5 mW	---	10 mW	15 mW
2	71-94	HFET	---	5 mW	---	10 mW	15 mW
3a 3b	101-104 101-104	HFET SIS	— 4	5 mW 0.10 μW	--- 10 μW	10 mW 0.40 μW	15 mW 15 μW
4	137-151	SIS	4	0.15 μW	15 μW	0.60 μW	23 μW
5	175-199	SIS	4	0.26 μW	26 μW	1.06 μW	39 μW
6	223-263	SIS	4	0.46 μW	46 μW	1.84 μW	69 μW
7	287-358	SIS	2	0.21 μW	21 μW	0.84 μW	32 μW
8	397-488	SIS	2	0.40 μW	40 μW	---	60 μW
9	614-708	SIS	2	0.42 μW	42 μW	---	63 μW
10	799-938	SIS	1	0.37μW	36 μW	0.73 μW	54 μW

In the worst-case scenario where only single-ended, two-port SIS mixers are used, a waveguide or quasi-optical LO coupler, having a coupling factor of -20 dB, will be required to combine the LO and RF signals appropriately. The LO power required at the input of the coupler is also given in Table 3. However, if a balanced mixer can be utilized, the LO power is supplied via a separate LO port on the mixer thus rendering the coupler unnecessary. Column #7 in Table 3 lists the power requirements for a balanced mixer configuration that is both sideband separating and balanced. The last column is a suggested *specification* per RF band based upon a 50 percent overhead for the worst-case conditions. The LO power *goal* will be 100 μW per band to ensure adequate power to overcome losses within the mixer block.

SIS mixer LO power requirements have also been studied by Belitsky [10], whose results are close to the above below 550 GHz, but considerably larger at higher frequencies due to assumed higher losses. We are reasonably confident that the 100 μW goal, if achieved, will be adequate through band 9. If balanced mixers or LO diplexers

are implemented for bands 9 and 10, then the specifications listed in the table should provide substantial margin at all bands.

7.0.3 Sideband Noise

Sideband noise refers to noise accompanying the LO at frequency offsets within the IF band of the mixer, and thus in its RF sidebands. The sideband noise should not be a significant portion of the receiver noise at any band. In Figure 1, the effective sideband noise is plotted against frequency for LO SNRs of 3 and 10 K/μW, where the LO power at the mixer is calculated from the formula in the previous section. We assume the following configurations, which we believe to be realistic: four junctions in a balanced mixer with 10 dB of LO-IF isolation through 275 GHz (bands 1-6), two junctions unbalanced from 275-720 GHz (bands 7-9), and one junction unbalanced for band 10, all with a normal resistance of 20 ohms. Above 275 GHz, the lack of LO isolation due to unbalanced mixers is compensated by the smaller power requirement due to using fewer junctions. The reduction above 750 GHz is due to assuming a single-junction mixer for band 10.

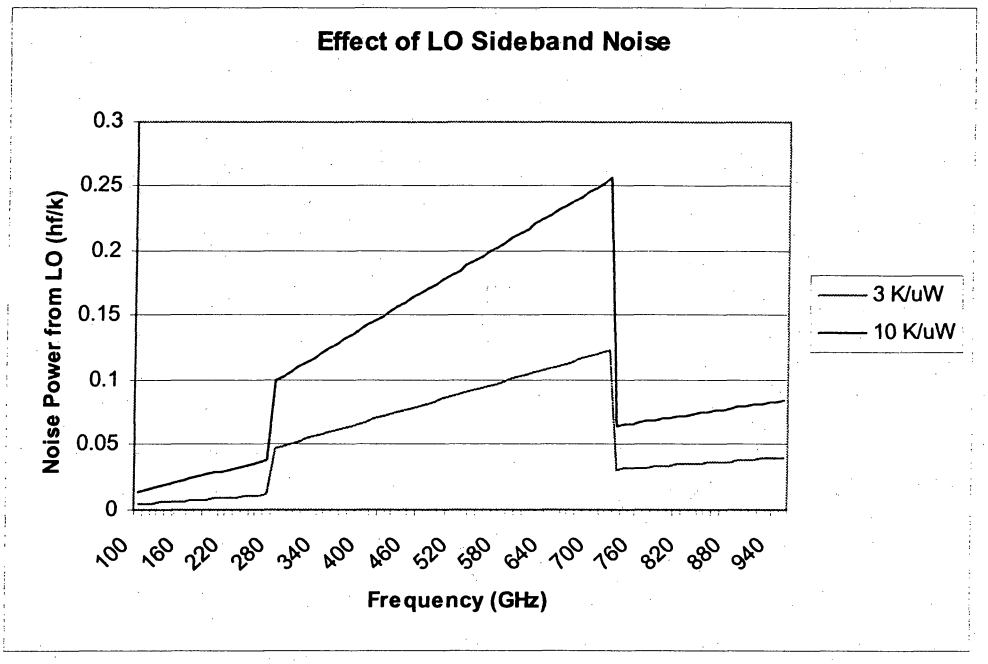


Figure 1: Effective LO sideband noise vs. frequency. Balanced mixers assumed below 275 GHz; 4-junction mixers for 100-275 GHz, 2 junctions for 275-750, and 1 for above 750; 20 o hm junctions.

These results form the basis of the specification of 10 K/μW maximum LO SNR in the RF sidebands, as listed in Table 1.

7.0.4 Tuning Resolution (step size)

The maximum tuning step achievable by a combination of first and second LO settings is limited by the desire to place any selected sky frequency near the middle of a 2 GHz baseband channel. A separate requirement is placed on the first LO for the SIS mixer bands so as to ensure that a line of interest can be placed near the middle of a selected sideband.

The present design exceeds the requirements. The second LO operates with steps of about 62 MHz (62.5 MHz

average). The first LO has steps of 5 MHz times the total multiplication factor, which ranges from 3 (band 1) to 88 (band 10), resulting in step sizes from 15 to 440 MHz. In addition, the first and second LOs each allow very fine tuning over a small range (10-40 MHz times the high-frequency multiplication factor, exact range TBD) around each nominal lock point.

7.0.5 Fringe Tracking, Phase Switching, Sideband Suppression

These issues are considered in detail in [3] and are briefly summarized here.

The second LO is required to support 180d phase switching at intervals of 250 usec so as to suppress spurious signals and d.c. offsets between the Downconverter and the Digitizer. It is desirable, but not required, that the first LO also support this feature. The present design allows either or both LOs to include such phase switching.

To allow sideband separation after correlation, 90d phase switching would be needed in the first LO. However, this feature is not required for ALMA. Sideband *suppression* will be supported by offsetting the first and second LO frequencies by the same amount within each antenna, but different among antennas.

Fringe tracking can be provided in either LO. The present design provides hardware to support fringe tracking in both places, so the choice can be made in software.

All of these features are achieved through direct digital synthesizers. A common Fine Tuning Synthesizer assembly, based on a DDS chip, will be built for use a component of the first LO Controller and of each of the four Second LO Synthesizers.

7.0.6 Phase Errors

The goals for phase accuracy and stability include:

- Greater than 90% interferometric coherence at 950 GHz (77 fsec rms), after all calibrations and corrections, on all time scales from 1s to 1e4 sec.
- Absolute visibility calibration to 0.1 radian at 950 GHz (16.8 fsec).

Assuming that phase errors are independent among antennas, we allocate half of the *squared* error budget to each antenna. Of this, we allocate half to the atmosphere, one-third to electronics, and the balance to the antenna structure. This gives:

Table 4: Phase Error Goals

	Atmosphere	Electronics	Structure	Total per antenna
systematic (avg), fsec	8.4	6.9	4.8	11.9
random (rms), fsec	38.5	31.4	22.2	54.5

These allocations are somewhat different from those assumed by Woody *et al.* (MMA Memo 144). There it was planned to achieve 90% coherence only 50% of the time at 300 GHz; here we use 950 GHz, and do not specify the time distribution. Of the error sources, only the atmosphere is non- stationary, and it is uncertain whether the above goals for it can really be achieved; if not, then the goals for the other components can be relaxed somewhat.

The time scales relevant to the stability goals include:

- shortest correlator integrating time ~.001 sec
- astronomical calibration cycle time ~10 (typ fast sw.)
- u-v cell crossing time (config dependent) ~100
- source mapping time ~1e4 (3h)

In addition, the following considerations apply to the need to support fast switching for phase calibration:

- Calibrator is observed at a different frequency and/or mode (e.g., bandwidth and frequency resolution), maybe even a different receiving band.
- Unambiguous phase synthesis in all LOs, with repeatable phase after changing frequency, is required.
- Stable phase in all electronics *common* to calibrator and target (e.g., LO reference) is required for intervals > cal cycle time (typically 10s). This allows correction of phase *changes* on switching time scale, but not absolute calibration.
- Stable phase in electronics not common to calibrator and target is required for much longer, since absolute cal must be done with same setup on calibrator and target.

In the following discussion, phase fluctuations are addressed as phase noise on time scales less than one second and phase drift on time scales greater than one second. This distinction is arbitrary, yet useful for analytical purposes. Additional information on phase noise can be found in ALMA Memo #311 [5]. For more information on phase drift see [6].

Phase Noise (random phase errors)

Although the phase noise budget above assigns 31 fsec to the electronics, it is a goal that will be difficult to reach. As a *specification*, we recommend that twice this value, or 63 fsec, be adopted for the electronics. This gives 85% coherence at 950 GHz when the atmosphere and antenna phase noise contributions are as given above. A proposed allocation among components based upon this goal and specification is given in Table 5

Table 5: Phase Noise Allocation for LO Electronics

Component	Goal [fsec]	Specification [fsec]
Reference Source	14	35
Fiber Distribution Subsystem (laser, round-trip fiber corrector, fiber, photomixer)	14	35
YTO Driver, inside PLL bandwidth	14	25
YTO Driver, outside PLL bandwidth	10	15
Multiplier #1	10	15
Multiplier #2	10	15
Multiplier #3	10	15
Total	31.4	63

In each of the components listed in Table 5, the noise contribution is ascertained by integrating the phase noise power spectral density over an appropriate bandwidth determined by the component's location within the LO system. The bandwidth of the reference source and fiber distribution sub-system is bounded by the array coherence time at the lower end and by the PLL corner frequency at the upper end. "YTO driver inside" refers to the noise of the YTO driver integrated from the array coherence time to the *effective* PLL cutoff frequency (noise attenuated by the loop plus noise of the locking circuitry). This upper bound includes the non-negligible noise contribution from within the transition region between the passband and the stopband of the loop filter. "YTO driver outside" refers to

the noise of the YTO driver integrated from the *effective* PLL cutoff frequency to infinity (noise of the oscillator, multipliers, and amplifiers not attenuated by the loop). The phase noise power spectral density for the multipliers outside the loop is integrated over all frequencies.

Phase Drift (systematic phase errors)

Adopting the 6.9 fsec goal of Table 4, we identify the components of Table 6 as the most likely significant contributors to overall phase drift. The most difficult case, band 10, is considered. As a preliminary guess, each is allocated an equal contribution at 950 GHz to the RSS total (2.18 fsec for each of the 10 components). It is likely that some of the components will be better than these allocations and some will be worse.

Table 6: Phase Drift at 938 GHz, Preliminary Allocations

Component	Drift (degrees)
Line length corrector and fiber	.09 @ 119 GHz
Cable (fiber) to photomixer	.09 @ 119 GHz
Photomixer	.09 @ 119 GHz
Lock Loop (PLL IF, phase det., etc.)	.09 @ 119 GHz
Fine Tuning Synthesizer (DDS)	.09 @ 31.25 MHz
Cold Multiplier #1	0.19 @ 237 GHz
Cold Multiplier #2	0.37 @ 475 GHz
Cold Multiplier #3	0.75 @ 950 GHz
IF amplifiers	0.4 @ 12 GHz
Second LO	0.4 @ 8 GHz

7.1 DESIGN OVERVIEW

A general description of the design is provided here for reference, with details being given in later sections. We describe here the current *baseline design*. It includes, for the first LO, distribution of a reference signal at up to 122 GHz and, at each antenna, the direct phase locking of a source at the reference frequency (without frequency multiplication); the source is able to generate relatively high power, so that higher frequency first LOs can be generated by cryogenically cooled multipliers. The reference is transmitted as the difference between two infrared-wavelength carriers from lasers. In an alternative design, described separately in a later section, the first LO for all bands is generated directly from the two-laser signal by a photodetector at the antenna; no other antenna-based equipment is needed, and the central building equipment is nearly the same as the baseline. Implementation of this option awaits improvements in photodetector technology.

7.1.1 Central reference generation and transmission

Nearly all time-dependent functions in the array must be coherent with a single master oscillator from which reference signals are derived and distributed. This is expected to be a hydrogen maser.

As shown in Figure 1, we begin by generating from the master a set of fixed-frequency signals covering the range 20 Hz through 2.0 GHz. Some of these signals are used at the central building, and some are distributed to the

antennas.

Figure 1: http://www.tuc.nrao.edu/~ldaddari/pb7-1fig1.pdf Block diagram of central building portion of the local oscillator subsystem.

A mm-wavelength reference is then synthesized for the first LO. This is the only variable-frequency signal that is distributed to the antennas. The process uses a microwave synthesizer to produce 8.62-11.08 GHz in 5 MHz steps (using primarily the 2 GHz and 5 MHz references from the master), followed by synthesis of 27-122 GHz as the difference between two laser-generated optical frequencies. The lasers are designated as "master" and "slave," with one master required for the array and one slave for each subarray. The two-frequency optical signal is sent to each antenna on a single-mode fiber. For each antenna separately, the optical signal passes through a line-length stabilizer based on two-way optical phase measurement of the master laser signal.

To support operation of more than one subarry with independent frequency selection, a separate laser synthesizer must be provided for each, along with appropriate switching. The master laser may still be common to all antennas, but each subarray needs a microwave synthesizer, slave laser, and phase lock circuitry. The number of these to be provided is TBD.

Meanwhile, the 2 GHz reference is transmitted to each antenna as intensity modulation on the master laser carrier; and the 20 Hz and 25 MHz references are multiplexed and transmitted on a dedicated optical carrier (by a modulation method TBD), probably on a separate fiber.

At each antenna, the Reference Receiver assembly demodulates the 20 Hz, 25 MHz, and 2 GHz signals from their carriers; frequency-shifts the master laser carrier and transmits it back on the same fiber to the center for line length stabilization; produces an additional reference at 125 MHz; and distributes all the fixed references to various devices in the receiver cabin.

Figure 2: http://www.tuc.nrao.edu/~ldaddari/pb7-1fig2.pdf Block diagram of the antenna portion of the First LO subsystem.

7.1.2 First LOs

The first LO electronics at each antenna is divided into two main parts (see Fig. 2). First, there is a set of "drive" hardware that generates a high-power signal at 31 to 122 GHz that is directly phase locked to the reference. All of the drive hardware operates at room temperature. Second, depending on band, there is a set of zero to three cascaded frequency multipliers for producing the higher frequency LOs. These operate at cryogenic temperature. The second part is not needed for bands 1 through 3.

The mm reference is recovered by photomixing and used to phase lock a VCO. The VCO is either a YIG tuned oscillator (YTO) or YTO followed by a doubler and power amplifier. A separate VCO assembly is used for band 1, where it provides the LO directly to the front end at 27.3 to 33 GHz. Another VCO at 34-52 GHz drives multiplication chains for all other bands.

For each band, a "warm multiplier assembly" (WMA) is attached to the bottom of the cryogenic dewar so as to minimize the transmission loss between it and the front end, and also so that the phase locked loop can be closed as close as possible to the receiver. Each WMA contains the photodetector for recovering the reference and a mixer to compare it with a sample of the LO output. Except for band 1, the WMAs also contain a frequency doubler or tripler and a power amplifier; the output frequency range is different for each band, but is always in the range 68-122 GHz.

The PLL is closed at the driver output frequency by mixing a sample with the photonic reference at an offset approximately 31MHz. The PLL IF is brought to the First LO Controller module, where the offset reference is

provided by a direct digital synthesizer (DDS). The DDS signal includes fringe rotation and phase switching, as well as fine-frequency tuning over a narrow range (at least 10 MHz, possibly more).

For bands 4 and higher, additional frequency multiplication is provided by a cooled multipier assembly at 80K. Either a doubler, tripler (band 6), or a cascade of two or three devices is used to produce multiplication factors of 2, 3, 4, 6, or 8, depending on band.

7.1.3 Second LOs

The second conversion (IF to baseband) requires LOs at 8-14 GHz. Four second-LO synthesizers are provided to allow independent tuning of each polarizaton-pair of 2 GHz baseband channels.

The design covers the range in 62.5 MHz steps, with the possibility of finely-adjustable offsets of several MHz from the nominal frequencies. The offset includes fringe rotation, sideband suppression, and phase switching capability.

7.2 TIMING AND CLOCK INTERFACES

Throughout the ALMA telescope, including the central building and each of the antennas, a set of fixed-frequency, periodic reference signals shall be available to any device. These signals will be coherent across the array, and will all be derived from a single master oscillator at the central building as described earlier. The nominal frequencies are:

f1 = 2.0 GHz
f2 = 25.0 MHz
f3 = 20+5/6 Hz (20.833.. Hz, 1/.048 Hz exactly).

The actual frequencies, with respect to the SI second, will be very close to these values (see below). Note that ratios of the nominal frequencies are integers; these are exact. (The value of f3 has been changed from that given in earlier ALMA documentation, where it was 20.0 Hz. The new value is intended to avoid small-harmonic relationships to the power line frequencies.)

An important concept is that of "array time," which is the continuous measure of time on which the array operates internally. It is determined by the phases of the three reference signals at a particular place in the central building, up to an ambiguity interval of 48 msec. At other locations (especially at the antennas), the distributed versions of the reference signals represent the array time plus the propagation delay of the distribution system; if true array time is needed to high accuracy, then the user must calibrate the delay time and subtract it. The f1 reference is the most accurate measure of time, but it has an ambiguity interval of 0.5 nsec. Therefore, at all distribution points f2 is required to be accurate to << 1 cycle of f1, so as to extend the ambiguity interval to 40 microsec. Similarly, f3 must resolve a cycle of f2, extending it to 48 msec. For larger time intervals, array time is determined by the monitor control subsystem in a computer at the central building which integrates the phase of f3 by counting cycles. Time within the computer must be accurate enough to resolve a cycle of f3, thus extending the ambiguity interval to any desired extent. The MC system is responsible for distributing this time to all devices that need it (including other computers) while maintaining the ability to resolve a cycle of f3. It does this by sending a message to the device and guaranteeing its delivery time to be within a known cycle of f3.

In accordance with specifications (see Table 1 of Requirements and Specifications, above), it is intended to keep array time very close to International Atomic Time. Nevertheless, it is important to understand that all devices in the ALMA telescope are synchronized to array time, and only indirectly (and less accurately) to any external measure of time.

Many of the considerations affecting timing and synchronization are covered in ALMA Memo 298 [1]. Detailed discussion of the implementation of the master clock is given in Specification 09001NX0002 [2], and more details are given in connection with the description of the Central Reference Generator in a later section. Some of the main principles are listed below.

1. Reference frequencies will be maintained within 1 part in 10^11 of their nominal values with respect to the SI second. (This is feasible if the master oscillator is a hydrogen maser that is regularly checked against distributed international timing signals.)

2. References f1 and f2 will be sinusoidal and will be delivered to users on coaxial cables. Reference f3 will be a logic waveform delivered on a bus conforming to RS485; it will spend a minimum of 1 microsecond and a maximum of 10% in the logic-1 state during each cycle.

3. At any one location, the phase stability of f3 on all time scales shall be less than 0.1 cycle of f2; and the phase stability of f2 shall be less than 0.1 cycle of f1. (Phase stabilities much better than this are actually expected.) The phase stability of f1 shall be better than 1psec with respect to the master oscillator. (This requires stabilization of the transmission path length to each antenna.) However, there is no specification for the absolute phase of any signal nor for the relative phases among them, which may vary from one location to another.

This specification means that, at any point in the telescope, it is possible to know the time to an accuracy of 1psec and an ambiguity interval of 1/f3 = 48 msec by means of these signals alone (i.e., without maintaining any local clock).

4. At one location in the central building, the phase of f3 will be adjusted so that it has a known relationship to external measures of time. In particular, the 0-to-1 transition of each 125th cycle of f3 shall coincide with the UTC second within 10 microseconds. Knowledge of the difference between those transitions and the UTC second shall be maintained to <100 nsec.

5. Any device that requires synchronization of its timing to the rest of the telescope shall accomplish that synchronization using one or more of these signals. A device should use only those reference signals needed to achieve its accuracy and ambiguity requirements.

6. Any frequency that is a multiple of f3 may be synthesized with unambiguous phase from the references. Devices may synthesize such frequencies for internal use. If any device-internal signal requires synchronization to other parts of the telescope, then it must be synthesized in this way; thus, it must have a frequency that is a multiple of f3.

7. If a frequency not in the given set is needed by multiple devices within one room, then that frequency may be synthesized once and distributed locally. Care must be taken to maintain a phase-stable distribution network. In particular, 125 MHz will be synthesized and distributed at each antenna.

8. The array master clock is driven by the same reference signals, and it maintains knowledge of the complete time by counting. (A full specification of the master clock implementation is given in [2].) The reference signals provide knowledge of the array time with an ambiguity interval of 1/f3=48 msec. Any device which requires synchronization on longer intervals must do so using commands received via the Monitor-Control (MC) system. Further details are given in [1].

7.3 FIXED REFERENCE DISTRIBUTION

7.3.1 Master Oscillator

Hydrogen Maser frequency standard. Specifications TBD

7.3.2 Central Reference Generator

The Central Reference Generator will supply the low noise, phase coherent RF and timing signals. It will produce 20.833 Hz, 25 MHz, composite 25 MHz + 20.833 Hz, 125 MHz, and 2 GHz signals to be distributed throughout the array. All signals will be coherent to the master reference (maser) and distributed according to other module requirements. (i.e. coaxial, twisted pair, fiber optic, etc.) The module will have an internal microprocessor to monitor PLL control voltages, signal levels, lock indicators, and maser vs. GPS drift rate. The drift rate will be measured by a 64 bit counter and can be accessed and reset through the ALMA Monitor Bus (AMB).

Specifications

Inputs:
DC Inputs
+15 VDC @ 2.6A at turn-on 1.8A after warm-up
+5 VDC @ 310mA
5 MHz input from hydrogen maser @ +10 to +15 dBm
1 Hz (1 PPS) from GPS receiver at TTL level

Outputs:
(All output signals are phase coherent with the 5 MHz input.)
20.833 Hz timing signal (48 mS)
6 msec pulse width (positive going; 12.5 % duty cycle)
Rise time ≤ 3 nsec
0.8 nsec rms jitter (max)
RS485 signaling
25 MHz reference signal
+7 dBm output power into 50Ω
Harmonics < -55 dBc
SSB phase noise:
-136 dBc @ 10 Hz offset
-156 dBc @ 100 Hz offset
-163 dBc @ 1 kHz offset
-165 dBc @ 10 kHz offset
-165 dBc @ 100 kHz offset
25 MHz + 20.833 Hz (48 mS) composite signal
Modulated onto 1550 nm laser source
Output through single-mode fiber optic system at 1550 nm wavelength
125 MHz reference signal
+13 dBm output power into 50Ω
Harmonics < -55 dBc
SSB phase noise:
-121 dBc @ 10 Hz offset
-131 dBc @ 100 Hz offset
-155 dBc @ 1 kHz offset
-175 dBc @ 10 kHz offset
-177 dBc @ 100 kHz offset

2 GHz reference signal
+10 dBm output power into 50Ω
Harmonics < -55 dBc
SSB phase noise:
-98 dBc @ 10 Hz offset
-112 dBc @ 100 Hz offset
-133 dBc @ 1 kHz offset
-142 dBc @ 10 kHz offset
-144 dBc @ 100 kHz offset

Monitor and Control:
The CRG will interface to the AMB through an interface board
The microprocessor in the CRG will monitor and control the following functions through the interface:
All RF output signal levels
Lock status of the maser to 5 MHz VCXO PLL
PLL control voltages
Maser vs. GPS drift rate
64 bit counter will allow ≅ 5849 years uninterrupted counting
Control remote reset of drift counter through CAN interface

Miscellaneous:
Operating Temperature: -20 to +70°C, forced air cooling
Enclosue: Double-wide AT module
Connectors:
50Ω OSP for all RF signals
FC/APC or E2000 optical signals
DB-50 style electrical connectors

7.3.3 Reference Receiver

The LO Reference Receiver will be installed into each antenna to receive and distribute the reference RF and timing signals required by the various antenna modules. An optical detector will receive the composite signal from the Central Reference Generator and the 20.833 Hz timing will be demodulated and distributed. The 25 MHz signal will be used to phase lock a low noise LO chain using VCXO's and PLL's to generate 25 MHz, 125 MHz, and 125 MHz comb outputs. The 25 MHz and 125 MHz signals will be reference signals distributed throughout the antenna. The 125 MHz comb will be used by the Second LO Synthesizers only. The 125 MHz VCXO will be phase locked to the 2 GHz reference from the High Fiber receive system using the 16th harmonic from the 125 MHz comb. This will allow the reference signal to the Second LO Synthesizer to remain phase coherent to the maser, and track the fiber optic, line corrected reference signal. An interface board will be used to communicate between the microprocessor and the ALMA Monitor Bus (AMB).

Specifications

Inputs:
DC Inputs
+18 VDC @ (?) A
+5 VDC @ 310 mA
Optical Input
1550 nm wavelength
Composite 25 MHz + 20.833 Hz RF modulated on to carrier
2 GHz input from High Fiber receiver @ (?) dBm

Outputs:

20.833 Hz timing signal (48 ms)
- 6 msec pulse width (positive going; 12.5 % duty cycle)
- Rise time ≤ 3 nsec
- 0.8 nsec rms jitter (max)
- RS485 signalling

25 MHz reference signal
- +10 dBm output power into 50Ω
- Harmonics < -55 dBc
- SSB phase noise
  - -115 dBc @ 10 Hz offset
  - -136 dBc @ 100 Hz offset
  - -160 dBc @ 1 kHz offset
  - -171 dBc @ 10 kHz offset
  - -174 dBc @ 100 kHz offset

125 MHz reference signal
- +10 dBm output power into 50Ω
- Harmonics < -55 dBc
- SSB phase noise
  - -121 dBc @ 10 Hz offset
  - -131 dBc @ 100 Hz offset
  - -155 dBc @ 1 kHz offset
  - -175 dBc @ 10 kHz offset
  - -177 dBc @ 100 kHz offset
    - 1 kHz to 100 kHz (offset) = 91.8 fs
    - 100 kHz to 1 MHz (offset) = 67.7 fs

125 MHz comb output
- -40 ± 5 dBm per line into 50 Ω
- Noise (per line) at 14 GHz
  - 1 kHz to 1 MHz (offset) = 9.8 fs

Monitor and Control:

The LO Reference Receiver will interface to the AMB.
The microprocessor in the receiver will monitor the following functions:
- All RF output signal levels
- PLL lock indicators
- PLL control voltages

Miscellaneous:

Operating Temperature: -20 to +70°C, forced air cooling.
Enclosure: double-wide AT module
Connectors:
- 50Ω OSP for all RF signals
- FC/APC or E2000 for optical signals
- DB-50 style electrical connectors

7.4 FIRST LO

7.4.1 Laser Synthesizer

Overview

The requirement of mutual coherence between each element of the array has led to the practice of a reference tone being generated in a central location and then distributed to each antenna. Typically a microwave reference frequency is generated at the central location and used to intensity-modulate the lightwave, which is demodulated at the antenna. It is advantageous to use the highest possible reference frequency so as to minimize reference frequency multiplication at the antennas. In recent years, commercial optical transmitters and receivers have been developed for frequencies up to 10 GHz, while the highest performance modulators (in limited availability) promise operation to 40 GHz. ALMA has adopted a different technique that allows the distribution of reference frequencies up to 120 GHz and possibly higher. Rather than modulating a single optical carrier, it is based on phase locking two lasers to a frequency difference equal to the desired reference frequency. Both optical signals are transmitted on a single fiber, and at the antenna a high frequency photomixer detects the beatnote.

This method relies on new developments in the field of photonics, including developments within the ALMA project. Some of this work has been discussed in ALMA memos [15–17].

This technique is identical to that which would be used in the direct photonic local oscillator option. Clearly, there is a need to weigh the benefit of distributing a higher reference frequency versus the risk of adopting a new method. The benefit-risk decision was made in favor of the new method partially because it will allow for the adoption of the direct photonic LO option if further technological breakthroughs permit it [24]. For instance, if suitable photomixers were developed for the range 300-950 GHz, then it would be possible to substitute the photomixers for the entire multiplier chain at great cost savings. The central generation of the LO frequency would be the same except that the tuning range would be 27-950 GHz and the required optical power level would increase.

Specifications

Frequency Range: 27.3 to 33 and 71 to 122 GHz
Frequency Step Size: <60 MHz
Switching Speed:
For switching less than 0.03%, 10 msec or less.
For larger frequency changes, 1.5 seconds or less.
Phase Noise: 25 fsec DC to 1 MHz offset
Phase Drift: 5 microns
Transmission distance: 25 km
(See also discussions in the Requirements and Specifications section of this chapter.)

Basic Description

The laser synthesizer consists of three major elements: a master laser, a slave laser, and a microwave synthesizer. The master laser and microwave synthesizer will be briefly discussed and then the main discussion will concern how the slave laser is locked to the master in the laser synthesizer.

Master Laser

The master laser forms one half of the LO reference. Primarily for purposes of round trip correction, discussed in the next section, the master laser is a highly stable narrow linewidth laser source. The linewidth must be less than 6 kHz and the frequency drift less than 100 kHz over an instrument cal cycle. Very few lasers can meet this specification. The laser that has been used for prototyping and for the test interferometer is a fiber ring laser, which

used to be commercially available from MPB [24]. This laser uses a very high Q fiber cavity of 22-m length and a sophisticated technique for achieving single mode operation and staying locked to that mode. Nevertheless, for ALMA the frequency drift performance will not be good enough even if the laser were still available, as the drift is specified as 10 MHz/hour. The master laser is such a key element in the ALMA system design that it should probably be developed in-house, rather than relying on a commercial product appearing. There is really no commercial application for wavelength stabilized lasers, at least on the stability scale that we require. However, there is at least one commercial product that we are aware of that might work. It is a solid-state laser available from Lightwave Electronics for about $75k. ALMA currently does not have a plan for the master laser.

Microwave Reference

The laser difference frequency is locked to a harmonic of a variable frequency reference from a microwave synthesizer. Its frequency range is chosen to minimize the largest harmonic number needed while keeping the total range well under one octave. This is partly because one method of harmonic generation (discussed later) involves photonic devices of limited bandwidth that operate well around 10 GHz. The selected range, 8.62 to 11.08 GHz, allows coverage of the full output range using harmonic numbers of 3, 9, 11, and 13 only.

The step size of the synthesizer is a compromise between obtaining the necessary resolution at the final LO relative to the first IF bandwidth (500 MHz spec) and achieving good phase stability and phase repeatability. To avoid phase ambiguity, the synthesizer must receive a reference from the master oscillator at a submultiple of the step size. The overall multiplication factor of the laser synthesizer is then the output frequency (to 122 GHz) divided by the master reference. The chosen value of 5 MHz achieves these objectives.

Many commercial laboratory synthesizers are available which provide the required tuning range and resolution, but none is phase-unambiguous across frequency changes and most are subject to large phase drift with temperature and aging. They are also expensive and provide many features unneeded in our application. For these reasons, the ALMA project will develop a custom module for this synthesizer.

Slave Laser and Phase Locking

The slave laser is phase locked to the master laser with a difference frequency given by a multiple of the frequency of the microwave synthesizer. The master laser can theoretically be split and shared by all of the laser synthesizers so that there is only one master laser for the entire array. This depends somewhat on the available optical power and noise added by optical amplifiers, but nevertheless the baseline plan is to have a single master laser. A separate laser synthesizer with slave laser is required for each *subarray* that needs independent frequency control. There will be at least three and perhaps five such subarrays (exact number TBD). An alternative is to have a separate laser synthesizer for each antenna, which would simplify the required switching. Thus, the cost and reproducibility of the slave laser is important. In either case, it is desirable to allow for the possibility of a single laser synthesizer to supply the LO reference for all antennas simultaneously, as this will maximize the cancellation of common-mode phase noise and drift. What this all implies is that there will be multiple laser synthesizers and a switching network between the synthesizers and the antennas that allows for a single laser synthesizer for the whole array.

For simplicity, the rest of this discussion will consist of description of a single laser synthesizer. As previously mentioned, the laser synthesizer consists functionally of three elements: master laser, slave laser, and microwave reference. ALMA nomenclature has led to the slave laser module being called the laser synthesizer because it does the phase locking and because the master laser can be shared among several synthesizers.

Fig. 1 shows a simplified schematic layout for the laser synthesizer. The red-lined section is optical fiber. For the prototype, the slave laser will be an external cavity diode laser (ECDL). This type of laser has single-mode, narrow linewidth and ease-of-tuning which makes phase locking convenient. The tuning range required, 27-122 GHz, is a small fraction of the available tuning range of the laser (6 THz). For production, the slave laser will be either an ECDL, a tunable DBR laser, or other suitable type. A small amount of light from the laser is coupled off and

combined with a portion of the output of the master laser. The combined light is then detected by a photomixer in the appropriate frequency range. Dual reference frequencies are then used to perform the phase lock. A variable reference frequency from 8.6-11.1 GHz is used to beat the signal down to an intermediate frequency (IF) using a harmonic mixer. The IF is then phase compared to a reference at 125 MHz, and the resulting phase error is used to correct and phase lock the slave laser to the master laser. The loop bandwidth is expected to be on the order of 1 MHz. The laser synthesizer has a much wider tuning range than a typical RF oscillator and is much more sensitive to temperature and environmental effects. The module will be shock mounted and temperature stabilized, but it is expected that regular frequency calibration will be required to account for laser drift and aging.

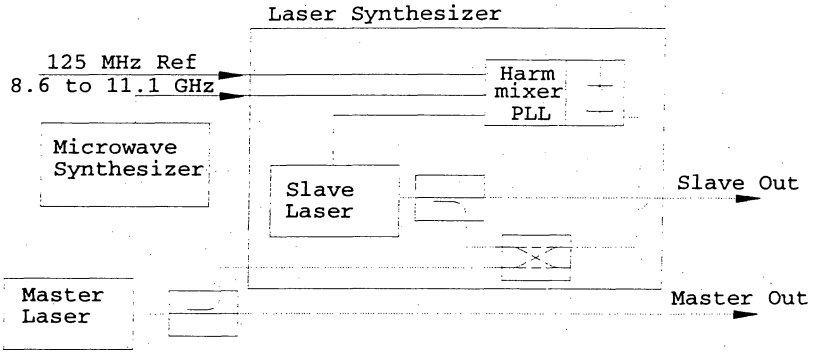


Figure 1: [synth1] Simplified schematic of laser synthesizer.

Fig. 2 shows an alternative technique that in which the harmonic generation is done photonically, eliminating the harmonic mixer. This approach is the current baseline for production, but additional development is needed. The box marked 'OFS' is an optical frequency shifter, wherein the frequency of the master laser is shifted by a fixed multiple of the microwave reference frequency. This technique has the advantage that the difference frequency gets shifted down to less than 1 GHz, where a standard, cheap off-the-shelf photodetector can be used instead of a millimeter-wave photomixer. In addition, much higher harmonic numbers can be used, allowing extension to frequencies above 1 THz. The technique or creating a comb of optical frequencies from a single laser and a microwave reference is variously referred to as an optical comb generator [19], a fiber comb generator [21], or an actively mode-locked laser. The technique based on [21] is being developed by ALMA partners at the University of Kent. Some aspects of the technique are discussed further in the section on Direct Photonic LO.

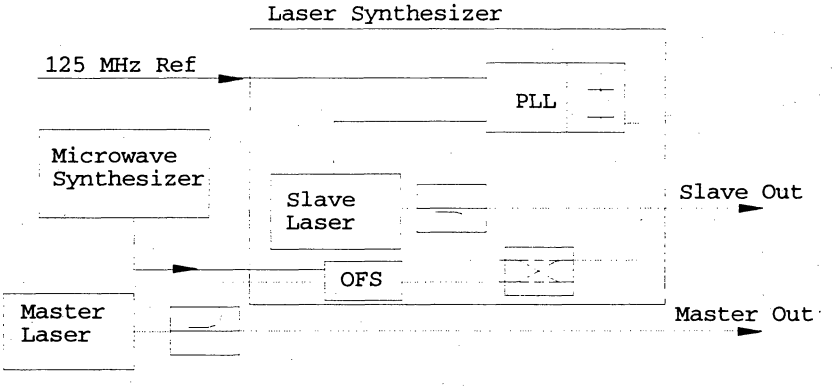


Figure 2: [synth2] Simplified schematic of laser synthesizer with optical phase locking.

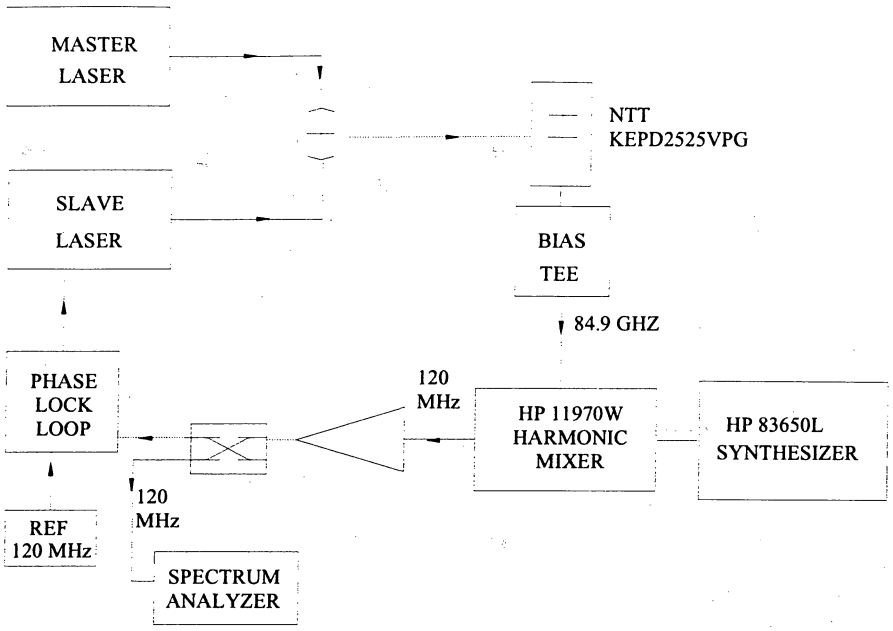


Figure 3 [schematic]: Schematic for Phase Lock at 84.9 GHz

Figure 4 [spectrum]: http://www.tuc.nrao.edu/~ldaddari/pb7-4fig4.pdf Spectrum of Phase Lock at 84.9 GHz.

Fig. 3 shows a slightly different phase locking arrangement using a harmonic mixer to achieve phase locking above 50 GHz. A measurement was made of a beatnote at 84.9 GHz, with the spectrum appearing as in Fig 4. The loop bandwidth is a few hundred kHz, and the phase noise is about 0.24 rad (450 fsec) in a 20 MHz bandwidth. Some performance improvement is still required to meet the specifications given earlier, and that effort is underway.

Slave Laser requirements

One of the biggest challenges is the selection of a laser to satisfy the tuning and phase locking requirements. Tuning range of available commercial lasers is as high as 100 nm, or 12 THz, well in excess of our requirement. However, the resolution, accuracy, and repeatability of the tuning are about one thousand times worse than for millimeter-wave oscillators, simply because the free running frequency of a 1550 nm laser is 195 THz. The phase locking technique can be similar to that of a millimeter-wave oscillator, but the free-running slave laser is not as well behaved. It has much greater frequency jitter, wider linewidth, much higher frequency sensitivity to any perturbation, and less predictable tuning characteristics. In addition, commercial lasers are not made with phase locking in mind as an application, and the laser specifications usually leave out jitter and noise characteristics that are important for the ALMA application. Good calibration and intelligent tuning control will be required to tune the laser to the phase locking range, and robust, wide bandwidth phase locking with compensation for slow drift of the laser is required to remove the laser jitter and noise.

Development Goals

The main development goal at this time is to significantly improve the phase noise result. Intensive effort is ongoing in this area. We have reason to be optimistic based upon results by outside groups [20]. Also, it will be necessary soon to decide upon a type of laser and method of implementation for the ALMA laser synthesizers. This will quite likely be different than the ones that are built for the test interferometer. This is mainly because the test

interferometer decisions are being made partly due to the expedience of time. For ALMA, considerations such as cost, reliability, packaging, and availability of components will weigh more heavily.

7.4.2 Line Length Correction

Overview

The path of the first LO to each antenna is a source of phase noise that must be continuously corrected. The fiber itself is electromagnetically highly linear below certain maximum power-distance relationships. Thus, the LO reference beatnote looks the same at either end of 10 km of fiber. However, environmental effects such as vibration, acoustic and temperature effects cause slow variations (> 1 msec) of phase to occur. The specification for the electronic contribution to this type of slow phase noise is 21 fsec, or 6.3 microns. Allowing for phase variation due to the photomixer and multiplier chain, we might take 4.5 microns as a reasonable target for phase variation due to the fiber link. The longest distance to an antenna will be 25 km, so the optical path must be stabilized to within a factor of 1.8e-10. Previously, microwave carriers modulated onto optical fibers have been used in a round trip configuration to continuously stabilize fiber length to well within the microwave wavelength, but that frequency was on the order of 10 GHz, and the stabilized line length on the order of 1 psec or 300 microns [23]. ALMA requirements are almost two orders of magnitude less than this. Thus an alternative technique was proposed.

It was decided to stabilize the fiber length by using a round trip correction using the optical wavelength as the standard rather than the microwave carrier. This is depicted schematically in Fig. 5. A very stable laser is used for this application, in which light from the laser is propagated from the central building to the antenna, and then back again. The returned signal and a sample of the original signal are input to a photodetector which gives an output proportional to the phase difference between the two signals. [In this way, the two-way path is kept at a constant length, and since the outgoing and return paths are on the same optical fiber, the one way path is also kept constant.

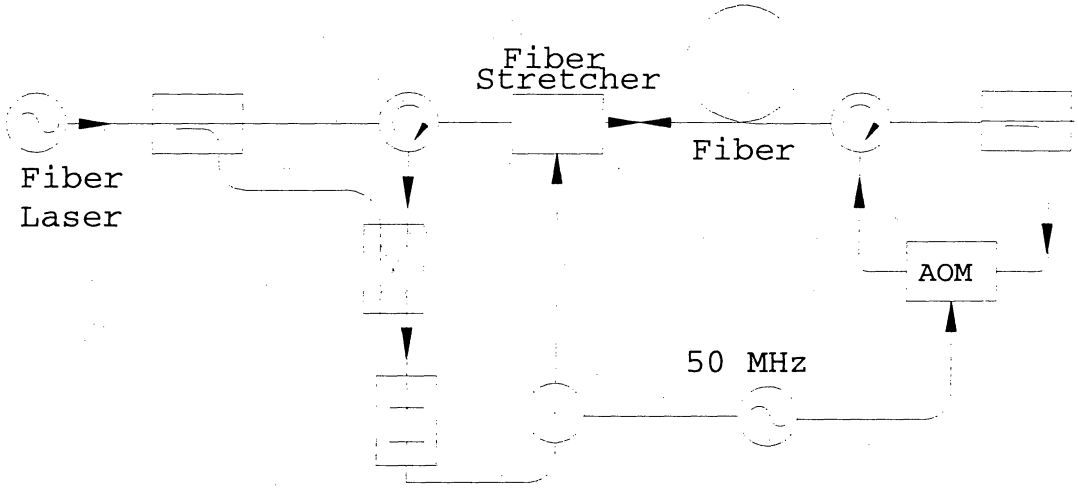


Figure 5 [opt-interferometer]: Simplified Schematic of Round Trip Optical Interferometer

Specification

The great benefit of doing the phase comparison at the optical wavelength is that the residual RMS phase is less than 1.55 microns in the ideal case. One drawback is that the stability of the round trip correction is much more difficult to achieve. If a perturbation occurs on the order of 2 microns that the control mechanism cannot correct for, then a cycle slip will occur. This means the exposed sections of fiber, particularly the fiber going up the antenna, will be critical to keep stabilized. Another drawback has to do with the required stability of the master

laser. Keep in mind that this implementation is really a 25 km long optical interferometer. The laser must then have a coherence length of at least 50 km, which implies a linewidth of less than 6 kHz. The optical interferometer does not actually keep the round trip fiber length constant, that is only a byproduct of the fact that it keeps the number of optical wavelengths contained in the fiber constant. However, that means that if the laser frequency drifts, then the fiber line length will be stretched in a proportional manner. For this stretched length to be kept to within 4.5 microns for a 25 km fiber, the laser drift should be less than 35 kHz over an instrumental cal cycle. This is a very difficult specification. If the fibers to all of the antennas were made the same length, then this phase drift would be common mode and the specification on laser drift could be relaxed proportionally. That may not be practical for the longest runs, but perhaps all antennas within a few km of the central station could have the same length of fiber.

Description of Fiber Line Stretcher

As mentioned previously, the frequency offset optical signal from the round trip is mixed with the source laser and the phase of the resulting difference frequency is used to adjust the length of the fiber within a closed servo loop. This fiber line stretcher consists of:

- A piezoelectric line stretcher to adjust for the fastest variation in the phase of the signal. This inexpensive commercial product has a range of 50 microns.
- An air-gap stretcher driven by a linear motor to account for slow changes in fiber path length. This can easily have a range of several mm. The assembly consists of an air-gap between two lensed-fibers, through which a collimated beam passes with low loss.

The fiber line stretcher is shown schematically in Figure 6. The piezo stretcher corrects for the fast variations and has a time constant of 2 kHz and the air gap stretcher adjusts slowly with a bandwidth of 2 Hz.

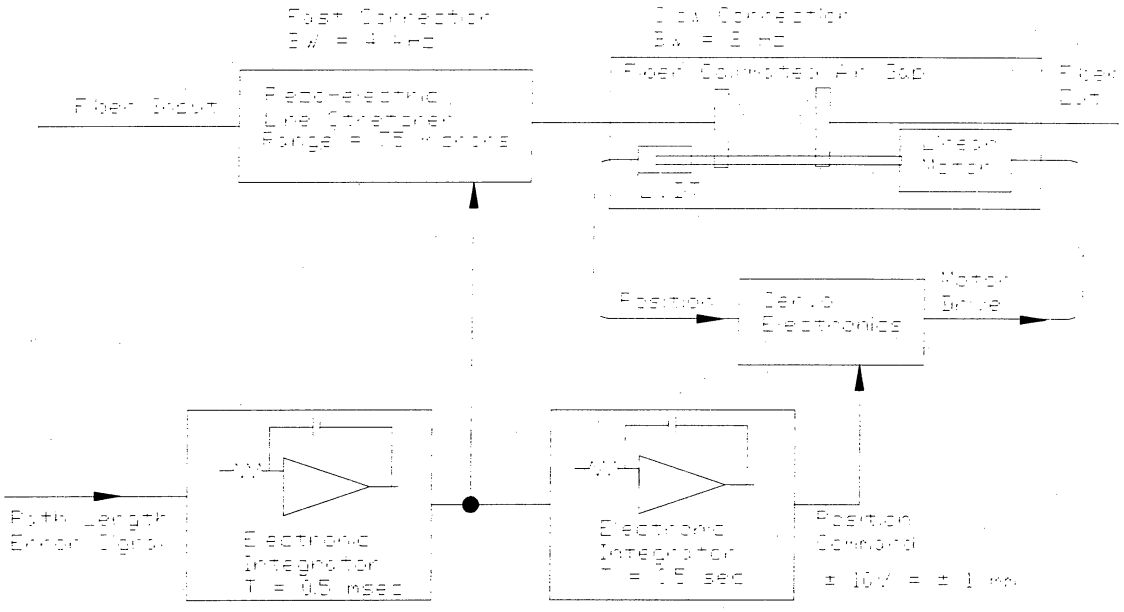


Figure 6 [line-stretcher]: Schematic of Fiber Line Stretcher

Round Trip Correction Tests

The fiber line stretcher was tested in a round trip correction configuration and results reported in [16]. That memo reports test on a 1 km fiber, showing an 11.1 GHz beatnote both with and without the phase correction as a proof that the concept worked.

Subsequently, further tests have been done on a 25 GHz beatnote using a 25 km spool of fiber. In this case the

round trip phase correction also appeared to work in that the phase drift was reduced when the correction was turned on. However, it was also apparent that there was a phase drift of ten degrees over one minute (at 25 GHz) while the correction was actively on. This is thought to be due to the frequency drift of the master laser, which is specified as 10 MHz/hour. The ten degrees corresponds to a line length error of 0.33 mm, which in turn would correspond to a frequency drift of 2.6 MHz on the master laser. As mentioned previously we need to improve this stabilization to better than 100 kHz.

Expected path length changes

A recent ALMA memo concerns underground temperature fluctuations at the Chajnantor site[18]. From this we can deduce the likely path length change of the buried fiber. At 1-m depth, the temperature fluctuation is reduced to about 0.0002 of the surface temperature fluctuation. We conservatively estimate the maximum rate of change of surface temperature at 15 deg C/hour. For a 20-min cal cycle, this becomes 5 deg C. So for a 25 km fiber with a temperature coefficient of 10 ppm/deg C, the underground path length change is 250 microns. For the fiber going up the antenna, assume 12 meters of fiber from the ground up through the cable wrap to the receiver. If it is totally uninsulated, the path length change over the same 20-minute period is 600 microns. Thus the path length change going up the antenna should exceed the buried path length change even for the furthest antennas. The fiber line length corrector should easily handle these variations. The air-gap stretcher can be configured for several-mm path length change.

Bandwidth limit of round trip correction

The round trip correction will easily handle slow variation due to temperature changes. What about fast variation due to mechanical or other perturbations? These types of perturbations in the worst cases are impulses and have fast frequency components. Thus the bandwidth of the round trip servo is significant. The piezo elements that stretch the fiber can be driven up to 5 or 10 kHz. A different design could probably be made to work to up to 100 kHz if needed. However, remember that the round trip over fiber at 25 km takes 2*n*L/c or 166 usec. Changes that take place on the order of one kHz will undergo about one radian of phase shift during the round trip, and this will limit the round trip correction bandwidth. The bandwidth will then be about 25/L kHz, where L is the fiber length (km) to the antenna.

Round Trip correction on the antenna

As mentioned previously, the fiber path on the antenna is expected to have more thermal variation than the entire buried run of fiber. Also, the antenna environment is harsh, with mechanical vibrations, continuous gusty winds, and thermal gradients. The baseline plan is to put the fiber in the bend-only axis wraps with other cables; but, if necessary, the antenna has a provision for on-axis, twist-only paths through each axis for this fiber, and this might produce better phase stability. A decision on whether to take this option will be made after experience is gained with the Test Interferometer.

7.4.3 Drivers

Design Considerations

As described in the Overview section, a fundamental choice in the baseline design is to partition the antenna portion of the first LO electronics into a part producing signals up to 122 GHz and a part producing higher frequency signals. Units of the first part are known as "drivers" and their outputs are phase locked to a photonically distributed reference. The second part consists of a power divider and cascaded frequency multipliers which are then outside the phase locked loop, but they are inside the vacuum dewar and are cryocooled to a stable temperature near 80K. These are known as the "cold multiplier assemblies." The drivers are the subject of the present section. A separate cold multiplier assembly is required for each receiving band above 122 GHz (bands 4 through 10), but the drivers are designed so as to have as many components as possible in common among several bands.
A simplified block diagram of the present design is shown in Figure 7. Considerations that led to this arrangement are discussed below.

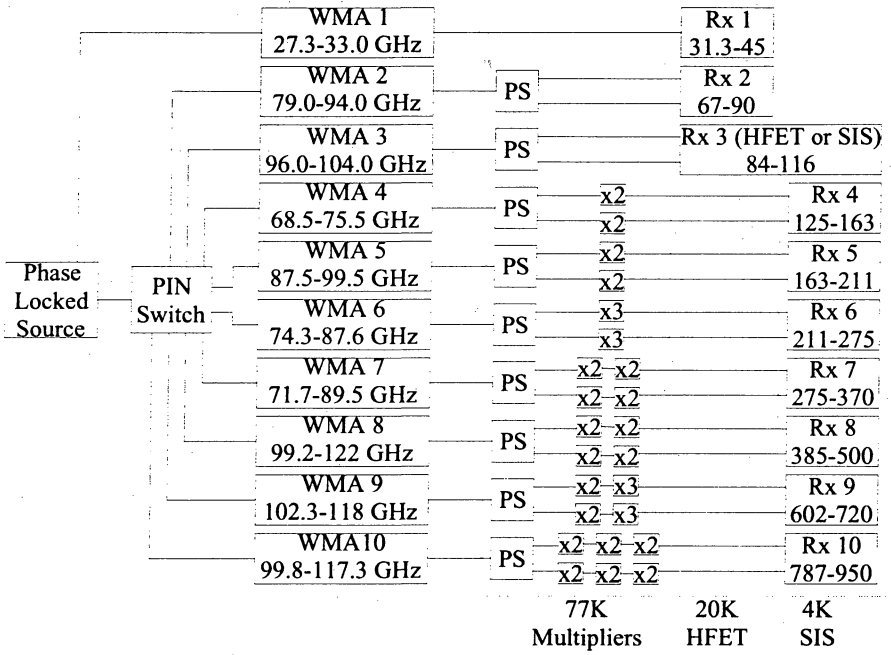


Figure 7: Simplified block diagram of the antenna portion of the first LO. Not shown are the reference signal input on optical fiber and its switching; the offset reference (DDS); and the phase locking circuitry.

The cold multiplier assemblies require drive power of 20 to 200 mW, depending on band. If substantial losses intervene between the driver and the multipliers, such as in waveguide runs or switches, then the power that would need to be generated by the driver might be substantially more, and this will be difficult to achieve. To avoid this, the present design provides for the final amplifier of each driver to be mounted just outside the vacuum dewar and to be connected to each front end assembly without switching. This requires a separate final amplifier for each band (except band 1, where sufficient power is expected to be available from a YTO without further amplification).

Furthermore, it is desired to close the phase locked loop with respect to the distributed reference as close as possible to the final LO to the receiver, so the design also places a coupler, mixer, and photodiode close to the dewar wall and duplicates these components for each band. This choice is relatively expensive, since otherwise many of these

components could be common; but it avoids some switching and some waveguide runs that might be as long as 400 mm; the tradeoff is subject to re-consideration. These band-specific parts are called "warm multiplier assemblies" (WMAs).

Finally, in order to minimize the frequency transmitted to the dewar-mounted circuits from separated components, and thus to minimize loss and cost, most bands include a frequency doubler or tripler among the band-specific, dewar-mounted components. Again, this is subject to re-consideration.

In order to achieve high reliability and to achieve the requirements for speed of frequency change, the design avoids all mechanical tuning. The widest range electronically-tuned signal sources (VCOs) are YIG tuned oscillators (YTOs), and these are also have nearly the lowest phase noise of all microwave oscillators. However, such oscillators are readily available commercially up to only 26.5 GHz. They can be built up to at least 40 GHz, but our own studies have shown [5] that phase noise of currently available units increases rapidly above about 26 GHz. We have therefore chosen to use a YTO at 17-26 GHz as the basic signal source; this range is sufficient, after appropriate multiplication, for all bands except band 1, we will use a separate YTO at 27-33 GHz. (Phase noise is much less critical at band 1 because no multiplication is required.) Since bands 2-10 all need multiplication of the YTO by at least 2, a frequency doubler is made part of the common "VCO assembly," as shown in Figure 2 of the Design Overview section.

The resulting frequency plan is shown in Table 1. Each driver requires a multiplication factor of 1 (no multiplier), 2, or 3 after the VCO, resulting in three types of dewar-mounted warm multiplier assembly.

Table 1: Driver Frequency Plan

Band	YTO Band (GHz)	VCO Band (GHz) WR-22	VCO Band (GHz) Coax Option	Driver Band (GHz)	LO Band (GHz)
1	27.3-33	27.3-33	27.3-33	27.3-33	27.3-33
2	19.7-23.5	39.5-47	19.7-23.5	79-94	79-94
3	25.25-26	50.5-52	25.25-26	101-104	101-104
4	17.1-18.9	34.2-37.7	17.1-18.9	68.5-75.5	137-151
5	21.8-24.9	43.7-49.7	21.8-24.9	87.5-99.5	175-199
6	18.5-21.9	37.1-43.8	18.5-21.9	74.3-87.6	223-263
7	17.9-22.4	35.8-44.7	17.9-22.4	71.7-89.5	287-358
8	16.5-20.4	33-40.7	16.5-20.4	99.2-122	397-488
9	17.0-19.7	34.1-39.4	17.0-19.7	102.3-118	614-708
10	16.6-19.6	33.2-39.1	16.6-19.6	99.8-117.3	799-938

Another consideration that affected this plan was a desire to avoid mechanical switches (coax or waveguide), primarily because of concerns about their long-term reliability. The availability of PIN diode waveguide switches up to the WR22 band then allows the common circuitry to extend to about 50 GHz, but not beyond. If mechanical waveguide switches are found to have acceptable reliability and repeatability, then the doublers and triplers of the WMAs could be moved to the common circuitry, saving space and cost.

An alternative design moves the common doubler from the VCO to the WMAs of bands 2-10, allowing the

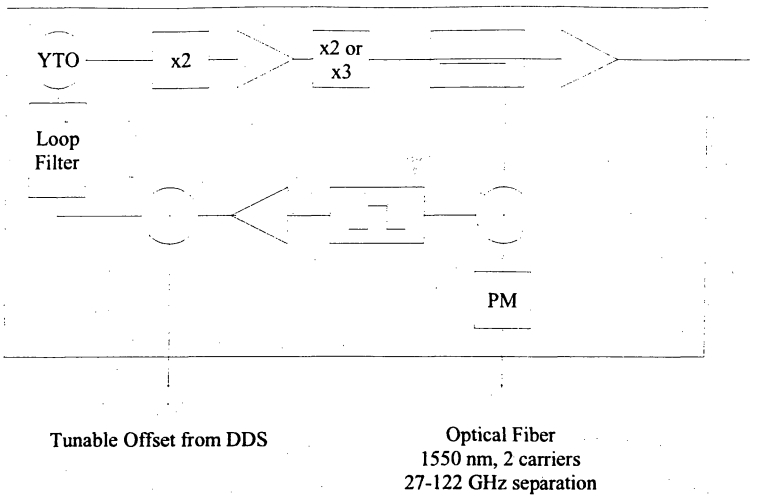


Figure 8: Block diagram of a complete driver for one band (other than band 1).

switching to take place at lower frequencies and allowing coax rather than waveguide interconnections. This is under consideration, and is included as an option in Table 1.

Design Details

The block diagram of a complete driver for one band (combining the VCO and the WMA, but omitting all switching) is shown in Fig. 8. It consists of a YTO followed by a doubler , a power amplifier, a doubler or tripler, a second power amplifier, and a directional coupler. Preliminary measurements indicate that this second power amplifier can be placed after the coupler (outside the loop) with little effect on phase drift [6]. The coupled signal is mixed with the variable millimeter-wave reference signal acquired from the photodetector. The resulting IF is filtered, amplified, and compared in a phase detector with an offset reference supplied by a direct digital synthesizer (DDS). The resulting error signal passes through the loop filter, thus closing the loop into the fine tuning port of the YTO. Fringe rotation, phase switching, and fine tuning are implemented via the DDS.

The power required to drive the cold multiplier assemblies is band-dependent. Based on estimates of efficiencies of the high-frequency cold multipliers, we predict that the power requirement per polarization is: 10 mW for bands 3-6, 50 mW for bands 7-9, and 100 mW for band 10. More power than this would make the multiplier development easier. A single MMIC power amplifier can be expected to provide 100mW to around 100 GHz (needed for bands 3-7) and 50mW to 122 GHz (bands 8-10). To obtain more power, it is planned to power-combine two amplifiers using a waveguide hybrid, as has been demonstrated by a JPL group; it is expected that four amplifiers can be combined with reasonable (>75%) combining efficiency. Another option for supplying more power would be to place the power splitter in the driver, followed by a power amplifier for each polarization. This is probably necessary for band 10, advantageous for bands 7-9, and unnecessary for bands 3-6.

The doublers and triplers will be one of three types: (i) FET using a multi-stage MMIC amplifier with the first stage biased in pinch-off and appropriate external filtering. This design has already been proven with a wideband 20->40-GHz doubler. (ii) varistor type using a zero-bias Schottky diode with appropriate embedding circuit. Prototype doublers and triplers of this type are presently being tested. (iii) varactor types, similar to those proven in [12]. See the Multipliers section below for further information.

A collaboration is underway with JPL on millimeter-wave power amplifier development. They are developing wideband power amplifiers for the FIRST mission with requirements [13] similar to ours, although they need higher power with less bandwidth. After one design iteration, they have amplifiers covering, with >100mW output power, 69-84 GHz, 88-106 GHz, and 99-114 GHz. These are microstrip MMICs using 0.1μm GaAs PHEMT process.

There is some concern that GaAs MMICs cannot extend all the way to 122 GHz. A design to cover the 99-122 GHz band is currently being fabricated in the InP process; chips should be available in 2001-Jan.

The LO power into each SIS mixer must be separately adjustable. This will be accomplished by setting bias voltages in the drivers and in the high frequency multipliers. It is recommended that this be done at the beginning of an observation only, since the performance of the SIS mixer is relatively insensitive to small fluctuations in LO power and the multipliers outside the PLL will produce a phase shift with bias voltage. The actual leveling can be accomplished in several different ways. We can a) adjust bias on varactor multipliers (there is concern that this approach may decrease the signal-to-noise ratio of the LO signal), b) adjust drain bias on one of the power amplifiers (this probably will have less effect on the signal-to-noise ratio), or c) a combination of the above. Some adjustment of the high-frequency multiplier bias will be needed in order to set separately the power to the mixers of each polarization channel.

Phase Noise

Figure 9 shows the phase noise measurement of the full YTO driver LO chain consisting of the components described in the previous section. The PLL is closed at the YTO frequency, 19.903 GHz. We have also made measurements with the PLL locked at 79.615 GHz. The results, in terms of phase noise, are basically identical since the doublers and amplifiers add little if any phase noise. The primary advantage in locking at the higher frequency is in phase *drift*.

Labeled on the plot is the integrated phase noise from 1-100 kHz, 12.65 fsec, and the calculated phase noise from 100 kHz to infinity, 17.84 fsec. The calculated phase noise assumes a 20 dB/decade drop in phase noise outside the phase lock loop as represented by the dashed line superimposed on the plot. The flattening of the phase noise spectrum near 500 kHz is not due to the noise floor of the YTO, but rather due to measurement equipment. The total noise in these two regions is 21.87 fsec, which is less than the total of the specification for these two ranges, 29.15 fsec. The loop bandwidth should be close to 300 kHz with a better reference, rather than the 100-kHz bandwidth used here. For more information, see [5].

Amplitude Drift

It is likely that the LO power leveling sub-system will set the power only at the beginning of an observation. This means that significant long-term fluctuations in LO power after leveling could translate to receiver power fluctuations in a manner similar to gain fluctuations in the IF amplifier, affecting total power measurements. The proposed specification for amplitude stability of the LO on one-second time scales is 0.03%.

Long-term amplitude stability (time scales greater than one second) is important for two reasons. First, we would like to keep the receiver noise temperature at its lowest point for highest sensitivity with a minimum number of adjustments. Secondly, the capability to maintain the amplitude stability of ALMA at the level of one percent is needed to combine imaging information from one array configuration to another [9]. The proposed specification for LO amplitude stability on time scales greater than one second is 3%. If the LO amplitude cannot remain this stable over an entire observation, than periodic levelings will be required during an observation. LO amplitude stability is being measured and will be reported in [6].

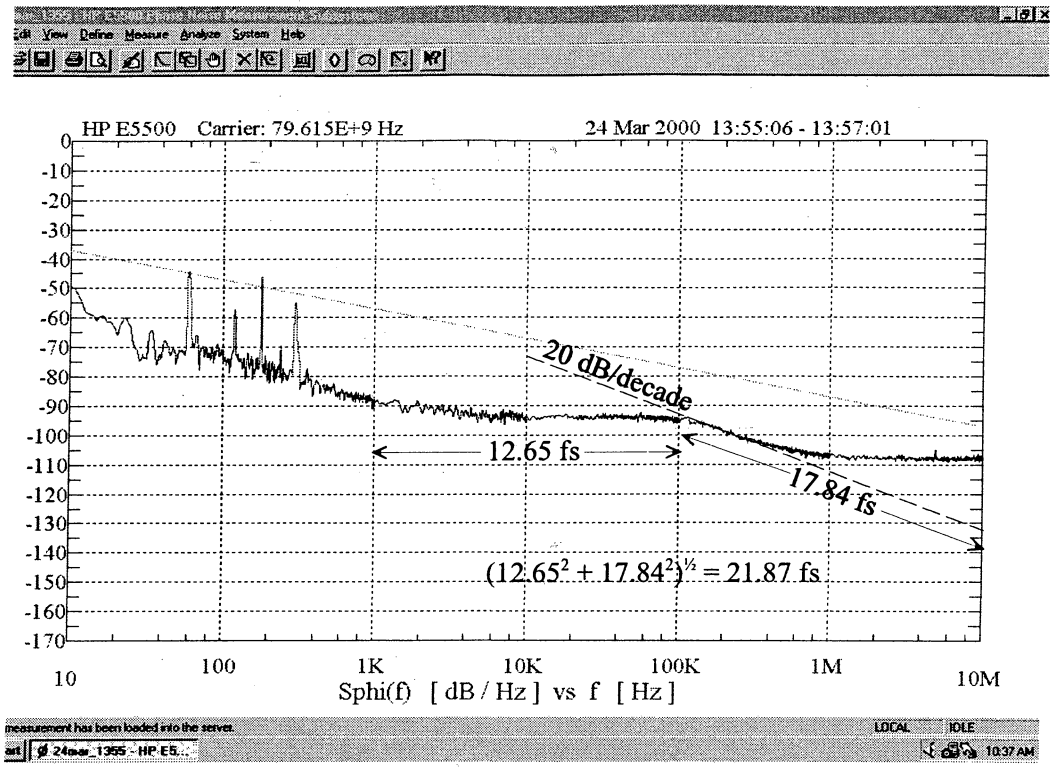


Figure 9: HP E5500 phase noise measurement of 79.615 GHz YTO driver LO chain.

7.4.4 Photomixers

Overview

The baseline plan for providing a millimeter-wave reference frequency in the range 27-122 GHz for the first LO is dependent upon the availablity of photomixers in that frequency range that will generate sufficient RF power for phase locking the drivers. This section will discuss this requirement and the devices that are either available off-the-shelf or are being developed within the project. Photomixer technology for the direct photonic LO has more difficult requirements and is discussed in a later section.

Specifications

A photomixer will be required for each LO driver. There will be two types of photomixer, one for 27-33 GHz (band 1) in WR28 waveguide or coax, and the other for 75-122 GHz (other bands) in WR10 waveguide. Each will produce at least 1 microwatt of RF power across its range.

Device Description

The photomixer requirement for the lowest frequency band, 27-33 GHz, can be met by many commercially available devices. These devices consist of a single mode fiber input and a coaxial K-connector output. Vendors include New Focus, Discovery Semiconductor, u2t, NTT, Newport, and OptoSpeed. For the higher frequency bands, it is desirable to have a photomixer with an output in fundamental mode waveguide. These are not available

commercially, but only because there is no market for them. The best photomixer chips have some response above 100 GHz. A test of a commercial chip was made at NRAO [11]. RF output power of as high as 40 microwatts at 110 GHz was measured. Fig. 10 shows the output power versus frequency for an input power level of 2.3 mW from each laser. This is well below the peak power level that the device can handle. Fig. 11 shows the output power versus total device current at a frequency of 110 GHz, with a maximum output power of 40 microwatts. These measurements were done with a commercial probe directly on the chip. As a commercial module, only a coaxial output is available.

The ALMA design is to use a packaged commercial device (coaxial K connector out) for band 1 and to use a commercial chip-level device in a custom-designed waveguide mount (WR10 out) for the other bands. Development of a waveguide mount for the u2t chips is underway at Rutherford Appleton Laboratory.

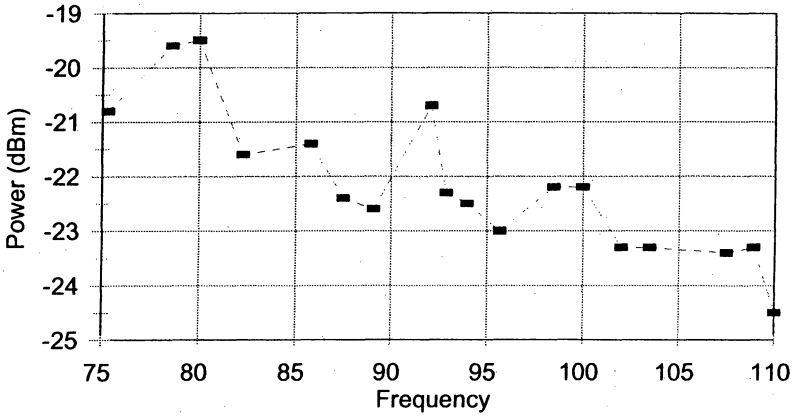


Figure 10: Measured output of u2t photomixer vs. frequency.

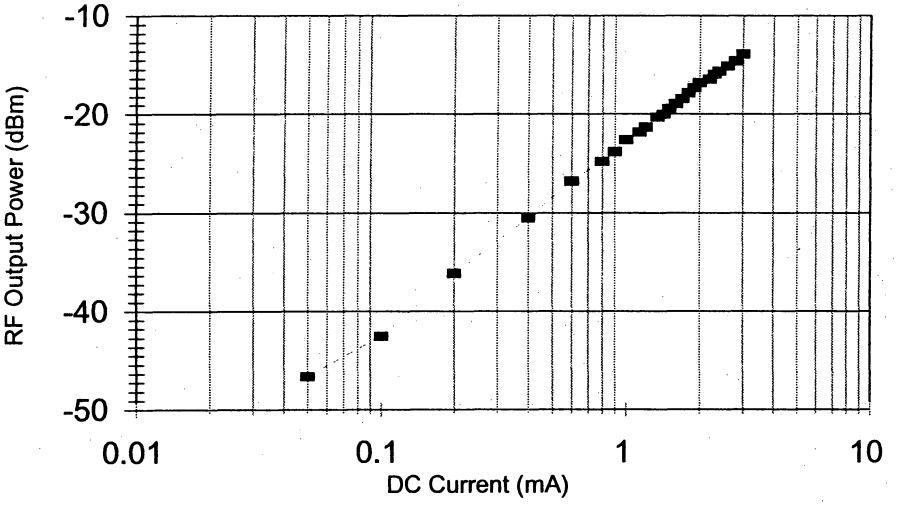


Figure 11:: Measured output of u2t photomixer vs. optical power (as measured by d.c. current) at 110 GHz.

7.4.5 Frequency Multipliers

The multiplier development plan is divided into two parts: frequency multipliers using discrete planar varactors, and frequency multipliers using monolithic circuitry. The first deals with broad-band, fixed-tuned frequency doublers used to extend the phase-locked loop LO system to cover the 137-163 GHz and 187-233 GHz bands. Frequency doublers for these bands will be based on the highly successful 40/80 GHz design [12] which uses a balanced planar varactor chip from the Semiconductor Device Laboratory of the University of Virginia. The measured results will be shown in [14] for room temperature operation. The peak efficiency increased to more than 60 percent upon cooling the doubler block to 20 K. The current status of the two new designs in progress, 55/110 GHz and 110/220 GHz output doublers, will also be discussed in [14]. These designs are the first iteration of the ALMA designs. Future iterations will be concerned with increasing the output power of the doublers and increasing the operational bandwidth as well as making the designs easier to fabricate.

Designs using discrete planar varactors are limited to about 250 GHz because the size of the chip package becomes electrically large and therefore the multiplier circuit becomes more difficult to tune properly over a wide bandwidth. Monolithic varactor multiplier designs for frequencies above 250 GHz will also be described in [14].

The multipliers required to satisfy the present ALMA frequency plan are listed in Table 2, including those that are part of the drivers. There are six basic doubler topologies and two basic tripler topologies. Multipliers of the same topology do not require separate development. Further details will be given in [14]. It is expected that at least two iterations (Levels One and Two) of the prototype multipliers will be required to achieve ALMA specifications. Tasks for evaluating cryogenic operation and device lifetime will be administered along with the continued development of the metalized polyurthane block molding process. The design and fabrication of the tripler needed for the ALMA test interferometer will also be performed. Evaluation of ferrite isolator technology for the WR-4 (170-260 GHz) band, and the construction of special bias supplies will commence late in the development period.

Table 2: Frequency multiplier list sorted by topology

Design Code	Type	Input Frequency [GHz]	Output Frequency [GHz]	Bandwidth [Percent]	Development Status
D-1B	FET	17.0 - 23.5	34.0 - 47.0	32	Proven
D-2A	Balanced	21.8 - 26.0	43.6 - 52.0	18	Low priority
D-2C	Varactor	34.2 - 37.8	68.4 - 75.6	10	Similar type proven
D-2D	(Discrete)	37.0 - 47.0	74.0 - 94.0	24	Similar type proven
D-2E		43.5 - 52.0	87.0 - 104.0	18	Similar type proven
D-2F		49.0 - 61.0	98.0 - 122.0	22	Similar type proven
D-2G		68.0 - 76.0	136.0 - 152.0	11	Low priority
D-2H		87.0 - 100.0	174.0 - 199.0	13	Low priority
D-3A	High Power	71.7 - 89.5	143.5 - 179.0	22	Prototype completed
D-3B	Varactor	99.0 - 122.0	198.0 - 244.0	21	Prototype evaluation
D-4A	Hybrid	143.5 - 179.0	287.0 - 358.0	22	Not Yet Designed
D-4B	MMIC/Varistor	198.5 - 244.0	397.0 - 488.0	21	Low priority
D-5A	High Power Varactor	199.7 - 234.5	399.5 - 469.0	16	Low priority
D-6A	Varistor	399.5 - 469.0	799.0 - 938.0	16	Low priority

Table 2: Frequency multiplier list sorted by topology

T-1A	Hybrid MMIC/ Varactor	74.3 - 87.7	223.0 - 263.0	17	Design underway
T-2A	Varistor	204.7 - 236.0	614.0 - 708.0	14	Not yet designed

7.4.6 First LO Controller

As shown in Figure 2 of the Design Overview section, the First LO Controller is a module that contains logic to control various functions of the first LO, including coarse tuning of the VCOs and all switching. It also carries the low frequency part of the PLL which is common to all bands, from IF (nominally 31 MHz) to the VCO fine tuning voltage, including the fine tuning synthesizer (DDS assembly). In addition, it controls the LO power level, partly through attenuators in the VCOs (as shown) and partly through control of the d.c. bias to multipliers (not shown).

The controller provides a single-point interface to the Monitor-Control subsystem through the AMB. The upstream computers will maintain all necessary calibration data and control algorithms, such as what settings are needed to tune to a particular frequency. Local logic will handle equipment safety requirements, such as current, voltage, and temperature limits.

7.5 SECOND LO SYNTHESIZER

The second LO synthesizer provides the last LO for down conversion of the astronomical data before digitization occurs. This is a YIG based synthesizer designed for low noise operation which interfaces to a direct digital synthesizer (DDS) to provide fine tuning. The YIG is microprocessor controlled and is phased locked to both a 125 MHz comb signal and the DDS synthesizer.

The synthesizer has the following performance specifications:

Output frequency	8.03 to 13.97 GHz
Output frequency steps	31.5 ± 10MHz
Output Power	+13 ± 2 dB
Output Spurious signal level	<-70 dBc except for harmonics of 125 MHz
Output harmonics of 125 MHz	<-80 dBc per tone
Phase Noise	Offset Level 1 KHz to 100KHz <15 fs 100 KHz to 1MHz <15 fs Harmonics <-40 dBc
Output VSWR	<2:1
RF port impedances	50 ohm
Lock time	1 sec between any two frequencies
Supply voltages	Voltage Current +24 250 ma max ±18 1.5 amps each max +5 2 amps max
Power Dissipation	50 watts max
Operating Temperature range	17° to 25°C
Operating Altitude	5500 meters max
Operating Humidity	0% to 95% RH @40°C
Monitor Points	Synthesizer freq 1 MHz Resolution

	Synthesizer output power
	Fine tuning synthesizer power
	Lock condition
	FM tuning voltage
	Main coil voltage
	Power supply voltages
Synthesizer out to comb in isolation	<-40dBc measured to nearest combline
Rear panel connections	20.83Hz in: with <0.8ns of jitter on rising edge
	125 MHz in: with following noise
	1 KHz to 100 KHz <92fs
	100KHz to 1MHz <68fs
	2 AMB node connections
	125MHz comb in: noise on each line
	1 KHz to 1MHz <10fs
	Comb extends from 8GHz to
	14GHz each line is -40±5dBm
	Synthesizer out
	Power supply voltages
Front panel requirements	Frequency display to 1 MHz Res
	BNC output of fine tune syn
	SMA output of 125MHz ref
	BNC output of FM tune voltage
	Lock condition indicator
Fine tune synthesizer drive	10 ma into 50 Ohm
Noise level	1KHz to 100KHz <92fs
	100KHz to 1MHz <68fs

7.6 FINE TUNING SYNTHESIZERS (DDS assemblies)

The Fine Tuning Synthesizer (FTS) uses direct digital synthesis techniques to achieve its high frequency and phase resolution. The device is based on the Analog Devices AD9852 DDS integrated circuit (DDS). The AD9852 provides the infrastructure to obtain very sophisticated control over its output. Below is a short discussion of the features of the device that apply to the FTS.

The DDS contains a 48 bit digital phase accumulator which is truncated to 17 bits. This 17 bit phase word is then converted to a 12 bit sine function which drives a high speed Digital to Analog Converter (DAC). There is a programmable 48 bit frequency accumulator. At each system clock cycle, 125 MHz in this application, the frequency accumulator contents are added to the phase accumulator. The result is that any frequency between DC and half the clock frequency can be generated to a resolution of 0.4 microhertz. In the FTS the lower frequency limit is established at about 8 MHz by the output coupling transformer.

There is a programmable 48 bit delta frequency word register. At each system clock cycle the contents of this register are added to the frequency accumulator. This results in an FM chirp signal.

The DDS contains a programmable 11 bit phase offset register. The contents of this register are added to the contents of the 48 bit phase accumulator before conversion to the sine function. This register is used to produce the pi / 2 and pi / 4 phase switching needed in the LO system.

The phase accumulator can be reset to zero on command. This along with the two programmable frequency control registers allow the approximation of any desired phase profile function modeled by a polynomial with up to two coefficients in the hardware. Along with the11 bit phase offset word an arbitrary initial condition can be obtained

with up to 11 bit resolution. Movement to a new function with different coefficients can be accomplished in a single system clock cycle. The primary limitations are the time required for the microcontroller to load the required parameters into the DDS registers and the settling time of other components in the LO synthesizer.

The DDS has an 8 bit data bus. I/O registers are loaded through this bus a byte at a time. There is a 6 bit address bus which identifies which part of a register is to be loaded. When all I/O registers are loaded an I/O update pulse is issued to the DDS to cause all data in the I/O registers to be loaded to the DDS active core on a single system clock edge. There is a system pipeline delay of 17 clock cycles from the first clock edge after the I/O update until the time these parameters take effect. They all take effect immediately on that clock edge.

Except for the case of a reset the DDS, the output is always phase continuous, i.e., the phase of the start of a new phase function begins where the previous one was at the time the new parameters take effect. This, of course, defines the phase offset register contents as part of the desired phase function.

The output of the DDS is a zero order hold sampled sine wave at the programmed frequency and phase.

All system timing events within ALMA are synchronized to a 48 millisecond fiducial which is distributed throughout the array. The FTS uses this signal to synchronize its activities with the rest of the array. There is a large Field Programmable Gate Array (FPGA) in the FTS which derives the necessary timing signals from this 48 ms tick and the system clock. This FPGA contains a set of counters and state decoding logic to derive the I/O update signal for the DDS and interrupts to the microcontroller to keep all parts of the FTS synchronized. This FPGA is a Xilinx Vertex series XCV50.

Supervision of the overall module is performed by a Motorola MC68HC16 series microcontroller. The microcontroller receives data through a communication link to an AMB interface board. Commands received are interpreted and the data are prepared for loading into the DDS. The DDS registers are mapped into the microcontroller memory space.

The FTS is contained within a commercial RFI box. The circuit board is mounted using standoffs at the corners to the box bottom. If heat dissipation becomes a problem with the DDS the board is laid out in such a way that the area under the DDS can be milled out and a heat spreader plug installed between the DDS and the mounting base. If this is done a spring loaded hold down device should be secured to the mounting base through holes provided in the PC board. This is also the case with the dual voltage regulator IC.

The AMB interface board is mounted piggy back over the FTS board and interfaces by a 10 pin connector. The AMB interface is attached directly to the mounting base by standoffs which pass through provided holes int the FTS PC board.

One of the Analog to Digital (A/D) channels of the microcontroller is connected to the 5 volt power supply through a voltage divider to monitor module input power. Three more of the A/D channels are connected to external monitor points provided as inputs to the module.

External Connections

1. 5VDC power, .5 amp nominal.
2. Power return. Connected to case.
3. 125MHz clock input, sine wave, terminated with 50 ohms.
4. 48mS reference, terminated with 50 ohms.
5. Output. 50 ohm impedance, (TBD) dBm.
6. 3 external monitor inputs.
7. CAN bus connector.

7.7 DIRECT PHOTONIC OPTION FOR FIRST LO

7.7.1 Overview

The ALMA local oscillator system presents a great challenge for the instrument builders. The generation of a pure frequency with high phase stability in the frequency range of 90-900 GHz at each of forty antennas, and preserving the phase relationship between antennas for long (perhaps hours) periods of time, is perhaps the most difficult part of the instrument. The cost and complexity of using conventional microwave oscillators followed by an amplifier-mulitplier chain is daunting. Nevertheless, it has been adopted as the baseline for the project in lieu of a competing technology that has been proven to work at high frequency. (There have been recent developments in multiplier techniques that suggest that the reliability and cost of that approach may be greatly improved by the application of new beam lead diodes and MMICs, as described in the section on Frequency Multipliers.)

An alternative approach has been proposed to use a direct photonic local oscillator. This approach is currently an option and not part of the baseline plan. The proposed system generates the LO by mixing two lasers with a difference frequency equal to the desired LO frequency. The laser frequencies are transmitted over single mode fiber from a central location to the various antennas that make up the array. The distances involved vary from a few hundred meters up to 30 km for the most remote antennas, so the use of lasers in the low-loss 1550-nm fiber-window is advantageous.

By using two lasers and a photomixer, it is possible to pump an SIS mixer. This has been experimentally proven by one research group at a frequency of 650 GHz [28]. There are two very attractive advantages to this technique over conventional means of generating millimeter/submillimeter LO sources. First, the same pair of lasers can be used to generate any frequency difference from DC-10 THz! Also, due to the very low loss of optical fiber, the lasers can be at a great distance from the SIS receiver. By extension, the same pair of lasers can be used to pump many SIS receivers separated by great distances. IF ALMA were to adopt this direct photonic LO scheme, a great simplification of the LO system and great cost savings would result. The critical piece of the puzzle is the development of a photomixer at frequencies above 200 GHz which has so far eluded researchers.

The potential advantages of such a system may be summarized as follows:

0. The majority of the components needed for the realization of the proposed scheme are commercially available. The communications industry has a huge investment in optical fiber systems, and the system outlined here exploits these fairly recent developme nts. We can be certain that intense development in this area will continue.

1. All of the frequency synthesis components of the local oscillator system may be situated in a laboratory environment remote from the array. At the antennas, only some leveling electronics and a photomixer are required. In terms of serviceability an d reliability, this is regarded as a great advantage.

2. The receiver interface is greatly simplified. Due to bandwidth requirements, the usual Martin-Puplett quasi-optical LO injection scheme will not be appropriate. LO injection using conventional methods with waveguides entering into the cryogenic enc losure (for each receiver band) would involve a relatively high loss and would complicate the thermal design of the receiver. In contrast, all that will be needed in the photonic system is one optical fiber into the receiver dewar resulting in negligible heat load. Vacuum feed-throughs for fiber are fully developed commercially.

3. There is a great reduction in complexity .

4. The proposed system eliminates the need for the usual microwave harmonic mixers.

5. The real cost promises to be far less than a conventional system.

7.7.2 Photomixer Technology

Photodetectors generally consist of a semiconductor material which has a bandgap energy such that it is sensitive to light in a certain wavelength range; and a photon of light can cause the generation of an electron-hole pair which under an applied electric field causes a current to flow.

Typical commercial photodetectors cannot respond faster than 50 GHz or so because the device is too long for carriers created by incident light to travel to the device terminals. Also, the device capacitance limits the high frequeny response. There are several groups working on millimeter-wave photomixers that take advantage of certain techniques to overcome these limitations.

One of these techniques the use of traveling-wave or velocity-matched devices [27]. In the traveling-wave photomixer, the active area is elongated in one dimension and the output current is collected in a traveling-wave structure such as coplanar waveguide. In the velocity-matched device, the output power from many high-speed photodetectors is combined coherently. An ALMA development project is ongoing under the auspices of Max Planck Institute for Radio Astronomy with the goal of developing this technique to create devices that will be able to pump an SIS receiver at 650 GHz.

Another technique is to tailor the device junction with a doping profile that allows for carriers to travel at overshoot velocity over short distances and thereby achieve faster response. Development of this technique for use in photomixers up to 1 THz is ongoing in a joint project between Nobeyama Radio Observatory and NTT Labs.

The highest reported RF power generated by a photomixer above 100 GHz is 0.500 mW [25]. A plot of the highest measured RF power versus frequency is shown in Fig. 1. Various research groups and types of photomixer are

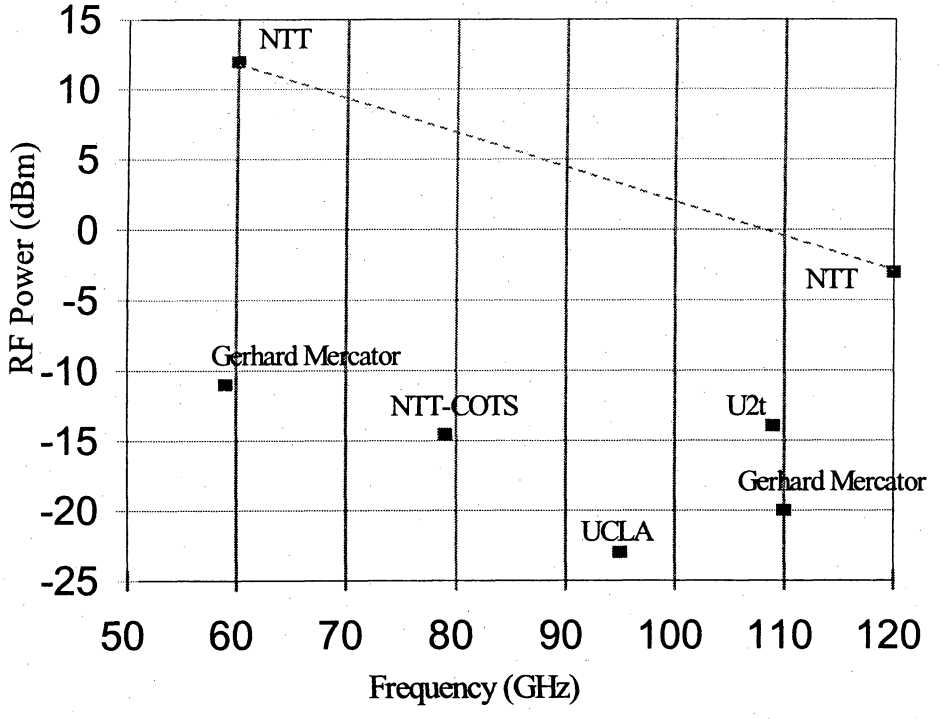


Figure 1 [photomixerpower]: Compilation of measured RF power vs. frequency for various photomixers. Top line is NTT research group.

represented in this plot, but the highest power levels are those reported by NTT using their Uni-traveling carrier photodiode [26]. This figure only contains results of photomixers that work at 1550 nm. New results are hoped at least by the first quarter of 2001.

In all of these research endeavors, measuring 10-100 microwatts of RF power is the most important goal. However, there is also the non-trivial problem of how to couple the power from the device into the SIS-mixer. Two approaches are being investigated, coupling of the power in to a fundamental mode waveguide and coupling the power into a free space quasi-optical beam.

7.7.3 Signal-to-Noise Limits

Another concern is whether the beatnote generated in the photomixer has an appropriately low level of AM-noise. Somewhere between 1 and 10 deg K per microwatt of RF power is the goal, although the effect of this type of noise would be mitigated by use of a balanced SIS mixer.

Figure 2 [SN]: http://www.tuc.nrao.edu/~ldaddari/pb7-7fig2.pdf Limitations of low-noise fiber transmission of a beatnote. Left Plot shows the Noise Temperature added to a Single Ended Mixer per microwatt of RF power. The right plot shows the maximum power that can be transmitted througha fiber at 1550nm due to Brillouin scattering.

Transmission of the two wavelengths over fiber, and generation of sufficient RF power with low noise levels is a topic discussed in a recent ALMA memo [17]. Many parameters, such as the type of mixer and number of mixer junctions, the type of RF coupling into the mixer, the total length of fiber, responsivity of the photoixer, ..etc are factors that need to be considered. However, there are some fundamental limits that need to be considered. These are illustrated in Fig. 2. First, the amount of noise (normalized to RF power) out of the photomixer decreases the more light you put in, until a limit is reached which is set by the laser relative-intensity-noise (RIN). A laser RIN of no higher than −160 dBc/Hz is thus recommended. Second, the amount of light that can be sent through the fiber is limited by Brillouin scattering by a power-distance law as shown in the figure. These two limits are in direct opposition in some case. For instance, for the furthest antennas, the power into the photomixer is limited to 2 mW, at which level the AM noise is above 50 K per microwatt. The result is either not enough RF power or too much noise. The use of optical amplifiers at the antenn will overcome the power limitation, but careful management of optical intensity noise will be required.

7.4.4 Optical Comb Generators

Optical comb generation is a technique that is being developed for optical frequency synthesis and appears to be an ideal method for translating our optically generated LO to an IF suitable for phase locking. In this method, a stabilized microwave source is used to phase modulate a laser in such a way that the optical carrier develops sidebands containing a comb of spectral lines at offset frequencies that are multiples of the microwave source frequency. Recent research has demonstrated that the comb can be made to extend to offsets as high as 4 THz from the carrier. If one of our lasers is modulated in this way, we can simply phase lock the other laser using whichever comb frequency falls conveniently close to it, i.e., with a difference frequency in the low frequency microwave region. The microwave difference frequency required for phase locking is obtained by combining the two optical signals in a photomixer. The conventional phase locked local oscillator in a mm-wave receiver also starts with a stabilized microwave oscillator, but this is applied at a high level to a diode harmonic mixer. The resulting diode conductance waveform contains high order harmonics, one of which beats with a sample of the LO to produce the desired microwave IF for phase locking. One difficulty with this approach is that each receiver waveguide band requires a separate harmonic mixer because the diode circuit must be matched to the waveguide being used. By comparison, the proposed scheme requires only a single photomixer operating in a relatively narrow range of optical frequencies to cover the desired LO frequencies (30-900 GHz). Thus, we can use a single comb generator for all bands in a single receiver, or potentially for all the receivers in the entire array. In addition, whereas the comb is actually at optical frequencies, none of the generated comb lines can appear in the receiver IF as interference.

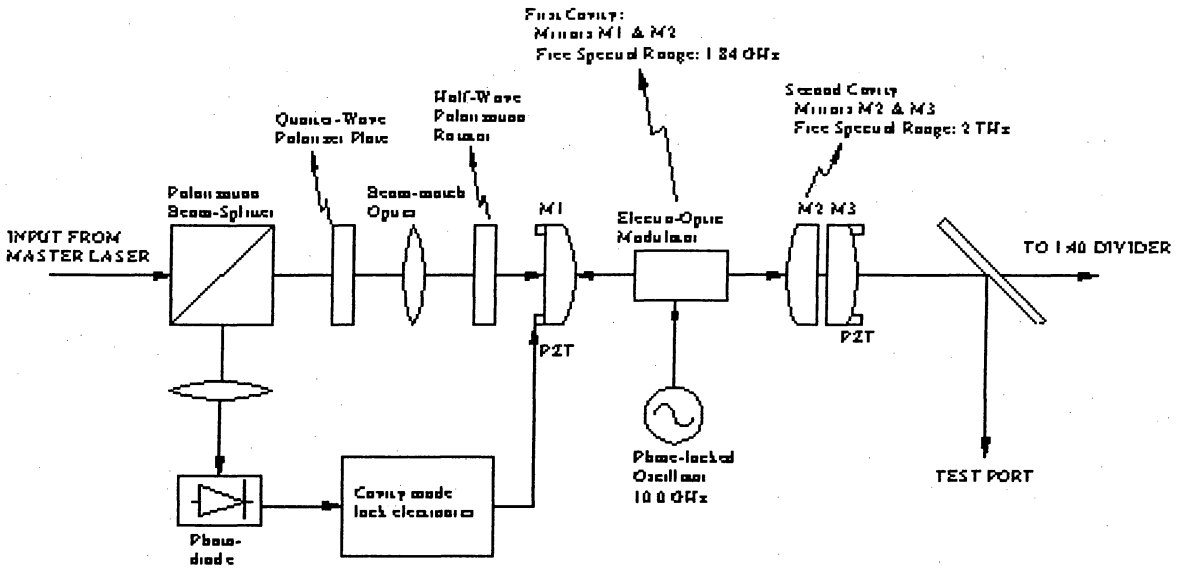


Figure 3 [freespacecomb]: Free-Space Optical Comb Generator [19].

Two methods of comb generation are possible: free-space and fiber. Both use the same principle, that an optical cavity allows multiple passes of light through a phase modulator. Fig. 3 shows the free-space optical comb generator, based on a technique developed for submillimeter phase locking [19]. The main cavity is formed by mirrors M1 and M2 an the phase modulator resides within the cavity. To the left of the cavity is an optoelectronic assembly for locking the cavity spacing exactly to the modulation frequency. The third mirror filters out unwanted comb lines. The result – a filtered optical comb – could be considered an optical frequency shifter as was depicted earlier in the Laser Synthesizer section. Fig. 4 shows the fiber implementation. A fiber loop forms the optical

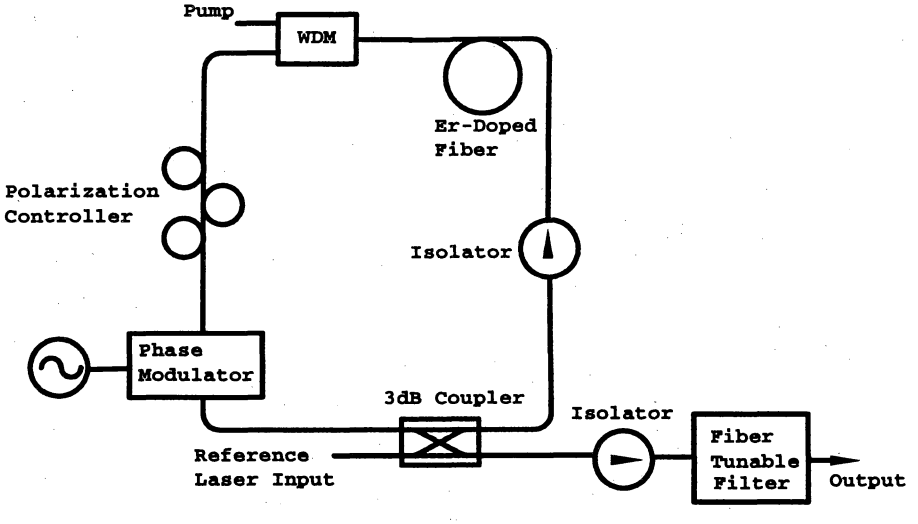


Figure 4 [fibercomb]: Fiber Comb Generator [21].

cavity. The fiber and its coupler form a low-Q cavity but this is overcome by having doped fiber within the loop to amplify the signal with each pass. The fiber modulator as before resides within the loop. A polarization controller is necessary to ensure that each loop round trip returns the same polarization. This technique is based on a paper written for DWDM development [21]. ALMA collaborators at University of Kent are investigating this apporach. It is not clear if this technique will generate a suitable low-noise frequency offset, due to the possibility of ASE noise added by the loop amplifier. However, the simplicity of implementing the comb generator in fiber is an attractive option.

REFERENCES

[1] L. D'Addario, "Timing and Synchronization." ALMA Memo 298, 2000-Mar.
http://www.alma.nrao.edu/memos/html-memos/alma298/memo298.pdf

[2] L. D'Addario, "Master Clock Implementation." Specification ALMA09001NX0002, 2000-10-24 (draft).
http://www.tuc.nrao.edu/~ldaddari/clockMasterSpec.txt

[3] L. D'Addario, "Fringe Tracking, Sideband Separation, and Phase Switching in the ALMA Telescope." ALMA Memo No. 287, 2000-Feb-15.
http://www.alma.nrao.edu/memos/html-memos/alma287/memo287.pdf

[4] W. Wild and J. Payne, "Specifications for the ALMA front end assembly," available at http://www.eso.org/projects/alma/develop/work-area.html, Aug. 31, 2000.

[5] E. Bryerton, D.L. Thacker, K.S. Saini, and R.F. Bradley, "Noise Measurements of YIG Tuned Oscillator Sources for the ALMA LO," ALMA Memo # 311.
http://www.alma.nrao.edu/memos.

[6] D.L. Thacker, E. Bryerton, K.S. Saini, and R.F. Bradley, "Phase Drift Measurements of YIG Tuned Oscillator Sources for the ALMA LO." ALMA Memo, in preparation..

[7] J. R. Tucker and M. J. Feldman, "Quantum detection at millimeter wavelengths," Rev. of Modern Phys., vol. 4, pp. 1055-1113, 1985.

[8] A. R. Kerr and S-K. Pan, private communication to E. Bryerton, July 2000.

[9] R. Brown, "ALMA Science Requirements," ALMA Project Book, Chapter 2. April 15, 2000 revision.

[10] V. Belitsky, "Local oscillator power requirements for ALMA SIS mixers." ALMA Memo No. 264, 1999-May.

[11] B. Shillue, "Millimeter-wave RF Power Measurements of a Commercial Photomixer" ALMA Memo No. 313, 2000-Jun.

[12] D. W. Porterfield, T. W. Crowe, R. F. Bradley, and N. R. Erickson, "A high-power fixed-tuned millimeter-wave balanced frequency doubler," *IEEE Trans. Microwave Theory Tech.*, vol. 47, pp. 419-425, Apr. 1999.

[13] L. Samoska, T. Gaier, A. Peralta, S. Weinreb, J. Bruston, I. Mehdi, Y. C. Chen, H. H. Liao, M. Nishimoto, R. Lai, H. Wang, and Y. C. Leong, "MMIC power amplifiers as local oscillator drivers for FIRST," in *UV, Optical, and IR space telescopes and instruments*, Proceedings of SPIE, vo. 4013 (2000), pp. 275-284.

[14] R. Bradley, K. Saini, E; Bryerton, and D. Thacker, "Frequency multipliers for ALMA." ALMA Memo, in preparation.

[15] ALMA memo #200, J.M.Payne, L.D'Addario, D.T.Emerson, A.R.Kerr, B.Shillue, "Photonic Local Oscillator for the Millimeter Array," Feb. 1998.

[16] ALMA memo #267, "Photonic Techniques for Use on the Atacama Large Millimeter Array," J. Payne, B. Shillue, A. Vaccari, June 1999.

[17] ALMA memo #319, "Photonic Local Oscillators for Radio Astronomy Signal-to-Noise Issues," B. Shillue, 08/00.

[18] ALMA memo #314, "Underground Temperature Fluctuations and Water Drainage at Chajnantor," Laura A. Snyder, Simon J. E., and Mark A. Holdaway, 6/00

[19] Jun Ye, Long-Sheng Ma, Timothy Day, and John L. Hall, "Highly Selective Terahertz Optical Frequency Comb Generator," *Optics Letters*, Vol. 22, No. 5, pp. 301-303, Mar. 1, 1997.

[20] L.N. Langley, M.D. Elkin, C. Edge, M.J. Wale, U. Gliese, X. Huang, and A.J. Seeds, "Packaged Semiconductor Laser Optical Phase-Locked Loop (OPLL) for Photonic Generation, Processing, and Transmission of Microwave Signals," *IEEE Trans MTT-S*, Vol. 47, No. 7, July 1999, pp. 1257-64

[21] Bennett, S.; Cai, B.; Burr, E.; Gough, O.; Seeds, A.J. "1.8-THz bandwidth, zero-frequency error, tunable optical comb generator for DWDM applications," *IEEE Photonics Technology Letters* , Volume: 11 Issue: 5 , May 1999, pp. 551 –553

[22] ALMA Systems PDR, Garching, 2000-02-28, http://www.alma.nrao.edu/~administration

[23] D'Addario, L.R. and M.J. Stennes,"Transmission of timing References to Sub-picosecond Precision over Optical Fiber," *Proc. SPIE*, vol 3357, pp. 691-701, 1998

[24] Zhang, C.Y. Yue, G. W. Schinn, W.R.L. Clements, and J.W.Y. Lit, "Stable Single-Mode Compound Ring Erbium-Doped Fiber Laser," *IEEE Journal of Lightwave Technology*, vol. 14, No. 1, Jan. 1996, pp. 104-109.

[25] [NTT] T. Nagatsuma, A. Hirata, Y. Royter, M. Shinagawa, T. Furuta, T. Ishibashi, and H. Ito, "A 120-GHz Photonic Transmitter," MWP2000, International Topical Meeting on Microwave Photonics, Oxford, England, Sept. 11-13, 2000

[26] [NTT2] H. Ito, T. Furuta, S. Kodoma, and T. Ishibashi, "InP/InGaAs uni-travelling carrier photodiode with 310 GHz bandwidth," *Electronics Letters*, 12 Oct 2000, Vol. 36, No. 21, pp. 1809-10

[27] [Jager] Stohr, A.; Heinzelmann, R.; Kaczmarek, C.; Jager, D., "Ultra-broadband Ka to W-band 1.55 micron travelling-wave photomixer," *Electronics Letters* , Volume: 36 Issue: 11 , 25 May 2000, pp. 970-972

[28] Blundell-Verghese, S.; Duerr, E.K.; McIntosh, K.A.; Duffy, S.M.; Calawa, S.D.; Tong, C.-Y.E.; Kimberk, R.; Blundell, R. , "A photomixer local oscillator for a 630-GHz heterodyne receiver, *IEEE Microwave and Guided Wave Letters* , vol 9 pp. 245-7, June 1999.

8 Systems engineering

Gie Han Tan

2000-12-07

8.1 Introduction

This chapter addresses the systems engineering efforts in the ALMA project. The accompanying tasks are under the responsibility of the ALMA Systems Engineering Group (ASEG).
Section 8.2 gives a concise summary of the concept of systems engineering and the areas which systems engineering will focus on in the ALMA project.
A major task for which the ASEG is responsible is the top-level system design of ALMA and this topic is summarized in section 8.3.
Sections 8.4 through 8.13 present more details on various important systems engineering management tasks and essential system wide requirements applicable to the ALMA project.

8.2 Survey of the systems engineering discipline

Complex systems like the Atacama Large Millimeter Array can only be created through the combined effort of many people. For this very reason, controlling the definition, design, production, integration, verification and validation of the instrument involves a major engineering effort as well as a technical managerial part. This is even strengthened by the fact that the people contributing to the project are geographically scattered among three, and when Japan joins four, different continents around the world.
Furthermore it is necessary that for all given life cycles of the instrument it be assured that the requirements of the astronomical users are also fulfilled.
Given this situation the need for a separate, systems engineering, group who takes care of these issues is probably obvious on first sight.

However, this is different from past other radio astronomical projects and therefore details about the existence of this ALMA Systems Engineering Group (ASEG) are discussed in this section.
The process of interaction between the ALMA Systems Engineering Group and other groups within the project is also described since this is crucial for performing their tasks in an effective way.

8.2.1 Raison d'être of ASEG

In "The Engineering Design of Systems: Models and Methods" various definitions of what systems engineering is about are summarized [1]. Of these definitions, probably the most concrete one applicable with the ALMA project in mind was given in MIL-STD 499A:

> Systems engineering is the application of scientific and engineering efforts to:
> 1. Transform an operational need into a description of system performance parameters and a system configuration through the use of an iterative process of definition, synthesis, analysis, design, test and evaluation.
> 2. Integrate related technical parameters and ensure compatibility of all related, functional and program interfaces in a manner that optimizes the total system definition and design.
> 3. Integrate reliability, maintainability, safety, survivability, human and other such factors into the total technical engineering effort to meet cost, schedule and technical performance objectives.

In previous radio astronomy projects so far, the tasks described by this definition were often distributed over the other participating disciplines. With the inherent risk that systems engineering issues would not be addressed properly resulting in less optimal technical solutions and/or the waste of resources.

Primary reasons to define a separate group responsible for systems engineering in the ALMA project are the following:

- The technical complexity of the ALMA instrument to be realized with limited resources and on a relatively short time scale need careful selection of the proper design technology for sub-systems:
  - A number of crucial technologies planned to be used in the design have never been demonstrated before at all or at least not at such a large scale. Trade-offs need to be made between science driven performance parameters like sensitivity and stability, and e.g. manufacturability and reliability.
  - Due to new technological capabilities sub-system functions can now be realized in more ways then in the past. Functions that were traditionally realized in hardware are now feasible in software and choices have to be made which approach is the most effective.
- ALMA will be a general-purpose instrument that should serve a versatile suite of different science requirements. A well-defined body must govern the flow-down of all these requirements to an appropriate system.
- The distributed nature of the project in terms of both geography as well as different, independent organizations needs a group that is responsible for achieving engineering commonality and interfacing in especially the design and construction phases of the project.

---

[1] Buede, D.M., "The Engineering Design of Systems: Models and Methods", pp. 9. Wiley & Sons Inc., 2000.

- The extent and complexity of the instrument requires that its reliability and maintainability be addressed already during the design & development phase in a complete and objective way to obtain a system having high availability for reasonable operating costs.

From the definition of systems engineering it becomes obvious that its tasks are related to probably every other activity in the project. Project members at all levels should be aware of this fact, the interaction between them and the systems engineering group is essential for success.

Finally, from the given definition it is also clear that systems engineering is not equivalent to system design, an often made misconception. But system design, the creative process of conceiving a solution for a certain requirement, is part of system engineering.

8.2.2 Working methodology and interaction with other groups

As depicted in figure 8.1, there will be a strong interaction between the ALMA project management, the project scientists and the ASEG. Within this ALMA management entity these three groups are responsible for the following areas:

- ALMA project management: resources and planning
- Project scientists: representation of astronomical users, defining science requirements
- ALMA Systems Engineering Group: engineering discipline

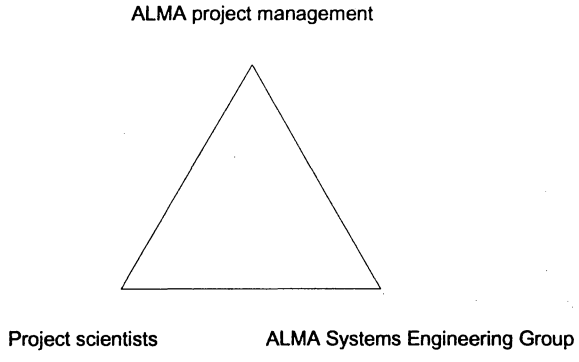


Figure 8.1, schematic representation of interaction within ALMA management team

The common main objective of the three groups is to optimally fulfill the scientific needs within the agreed constraints of budget and time.
The ALMA management is at the top of the triangle to represent its overall responsibility for the project.

The ASEG acts as the interface for engineering and technical issues between the project management and the technical groups contributing to the ALMA project.

Due to the relatively small size of the ASEG, involvement of specialist engineers from other disciplines is absolutely necessary to support the tasks of the ASEG. It is important that management and the specialist engineers are aware of this fact and that appropriate priority is given and time is allocated to these advisory tasks.

8.3 System design overview

Updated version of overview in MMA project book by Dick Thompson will be added here.

8.4 Product tree

The Product Tree (PT) is a top-down approach dividing the ALMA system systematically into successive levels of partial hardware and software based on the functions identified.

The products to be identified in the PT will as a minimum include:

- Items submitted to configuration control
- Items that are the subject of a project specification
- Items that are referred to in an interface control document

The primary aim of the product tree is to act as a tool in the systems engineering management of the project. It is a representation of which components the system consists off. Among others, it will improve the quality of keeping track of product specifications in a systematic way and the proper control of interfaces.

All items in the PT are partitioned on the basis of function, in some cases especially at the lower levels of the PT, an item might coincide with a physical assembly.

Since the PT is purely product oriented, activities that are not product oriented are not included in the PT and only appear in the Work Breakdown Structure. Typical examples are process-related tasks like project management and systems engineering.

- EMC, the IEC CSIPR EMC standards will be taken as the basis and adapted to the needs of the project. For example, additional requirements for emitted EM interference based on ITU-R RA 769-1 are foreseen due to the outstanding sensitivity of ALMA
- Safety
- Packaging, a modular approach is advocated that makes it possible to have Line Replaceable Units (LRU) that enable a/o efficient operation of the instrument
- Power distribution
- Environmental
- Preferred component lists, e.g. for embedded processor families

The information in this document will be to a large extent based on international standards from organizations like IEC and IEEE. Considerable use is made of the experience gathered in MIL-STD-2036A, VLT Electronic Design Specification and this is tailored to the specific needs of the ALMA system.

8.10 System availability and reliability

ALMA system availability is defined as the ability of the instrument to perform the required function under given conditions over a given time interval, assuming that the required external resources are provided.

So far, no quantitative requirement for availability has been determined for ALMA. It is planned to establish a baseline figure for ALMA on the basis of experience with existing radio astronomical instruments with similar technical and operational properties. Examples are the Very Large Array (VLA), Very Long Baseline Array (VLBA) and the Westerbork Synthesis Radio Telescope (WSRT).
This baseline availability figure will be used as a target for the ALMA system to be designed.

Various efforts are planned to assure that the required availability will be achieved. At least, the following tasks are foreseen:

- A uniform and objective method of reliability prediction based on MIL-STD-217F for electronic components and The Handbook of Reliability Prediction Procedures for Mechanical Equipment (NSWC-94/L07) is selected. These standards will be used to predict the reliability in terms of MTBF for every product in the ALMA system
- Failure Mode and Effects and Criticality Analysis (FMEA) will be carried out
- Fault Tree Analysis (FTA) will be performed
- The results of all three mentioned analysis methods will be fed back interactively to the design process to optimize the availability of the instrument

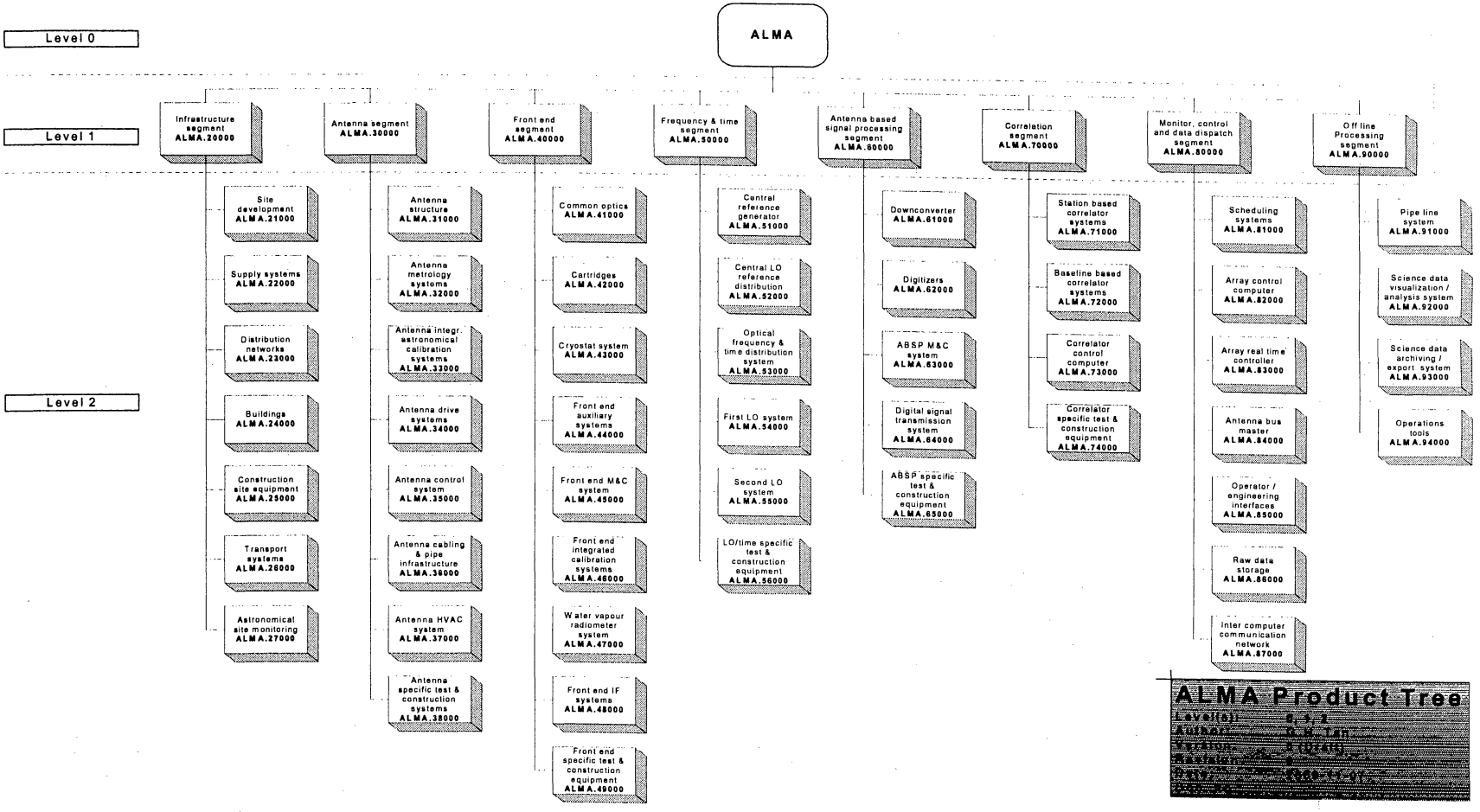


Figure 8.2, ALMA Product Tree / levels 0, 1 (approved by AEC) and 2 (preliminary)

Each Integrated Product Team will complete the PT at his level for the products under his responsibility. An exception is made for level 1 of the PT, which is the responsibility of the ALMA Systems Engineering Group.
The PT will be formally approved by the ALMA Executive Committee and will be under configuration control by the ALMA Configuration Control Board.

Figure 8.2 shows a preliminary version of the ALMA product tree. Level 1 of the PT is fixed and approved by the AEC, while level 2 needs to be completed in consultation with the responsible US Division Heads and European Team managers.

A full version of the ALMA product tree is available in a separate document "ALMA Product Tree" (in preparation).

8.5 Configuration control

What is under configuration and process description of configuration control to be included after consultation with AEC.

8.6 Quality management

Quality needs special care for a complex, hi-tech instrument like ALMA during its whole life cycle. Since this project book addresses phase 2 of the ALMA project, the quality management described in this chapter focuses on the development and production phases of the instrument.

It is anticipated that the required level of quality in the ALMA project must be obtained by using the following methods:

- Inspection of products
- Well described and controlled processes for, among others, Design & Development, construction and integration activities
- Use of standards
- Concurrent engineering

The traditional process of evaluation and inspection of products is definitely needed, but as the only measure not efficient to achieve the high level of quality as is needed in the ALMA project. Using this method only will for example result in the detection of defects, but does not identify the causes of these defects in an early stage. The repair of these defects will result in extra costs and delays in the project; a situation to be avoided by appropriate quality management measures.

A specific aspect of the ALMA project is the location where a major part of the instrument will be built, a site at more then 5000 m altitude having a hostile

environment for humans. Failures at this location can result in serious injuries or even death.
To avoid this problem, also here pro-active measures are necessary.

Based on industry experience, where it was also concluded that more inspection of products does not efficiently lead to better products, it is advocated to improve the definition and control of processes that play a role in the ALMA project. In industry this has been formalized under the well-known ISO 9000 standards.
According to the spirit of the ISO 9000 standards, a similar approach will be applied for the ALMA project. Several examples applicable to the ALMA project are:

- A well defined Design & Development process, including:
  - Clearly defined D&D phases
  - Proper review, based on well known terms, after every D&D phase
- Unambiguous and complete documentation, obtained through:
  - An ALMA document management plan
  - Use of templates and standard forms
  - The use of modern PDM tools to assure proper control of documentation

Improved process control will also be a major benefit to improve safety. Already in advance, when a process is analyzed and defined, possible risks are identified and solutions can be incorporated in the process to mitigate them to an acceptable, clearly defined level.

Where feasible the use of standards, either recognized general (e.g. ISO, IEEE) or project specific, should be used. The benefit of using standards is that these are well proven, widely known and accepted by a common group, thereby reducing the possibilities of making errors and minimizing the effort to get familiar with a certain topic.

Evaluation and inspection of products must already be addressed in the early design phase in the development of a product. Appropriate test plans must be established and the product design should be such that they can be easily and completely tested in the production phase.

From the description given above it will be clear that obtaining a proper level of quality is not only a systems engineering issue but can only be successful if there is commitment to it throughout the whole project organization.

8.7 Hardware design and production processes

A document providing a formal guideline for the hardware development and production process to be applied in the ALMA project is in preparation.
It has been prepared to standardize the hardware development and production process throughout the project and thereby assisting in:

- Creation of a common view on these processes
- Improvement of product quality
- Improvement of project control

Special attention is given to a methodology to track achieved performance during the development process against required product specifications.

It is recommended that this guideline is mandatory for the development and production of all hardware products that are under configuration control. Exceptions to this rule can be made after consultation with and approval by the ALMA Systems Engineering Group (ASEG).
It is encouraged to use this guideline also for the development and consequent production for other products that will be used in the ALMA system.

Due to the current status of the ALMA project a number of product development tasks have already reached the Preliminary Design Review milestone. In those situations the information given in the document with regard to objectives and deliverables at every phase can be used as a checklist if previous phases and milestones are completed according to this guideline.

Figure 8.3 shows a product development model, consisting of 7 phases with accompanying review milestones, which is applicable for hardware parts in the ALMA project. It is tailored to the needs of the project where high-tech products are developed and manufactured in large quantities.
Depending on the nature of the product, a/o technology maturity and production quantity, extra phases can be added or phases can be merged. For these specific situations where a different development model is proposed, the ASEG should be consulted and approval from this group is necessary to apply a modified product development model.
Note that the development model shown in figure 8.3 is represented as a classic linear model. The interaction between design, production and integration activities is not depicted explicitly, but should be taken into account.

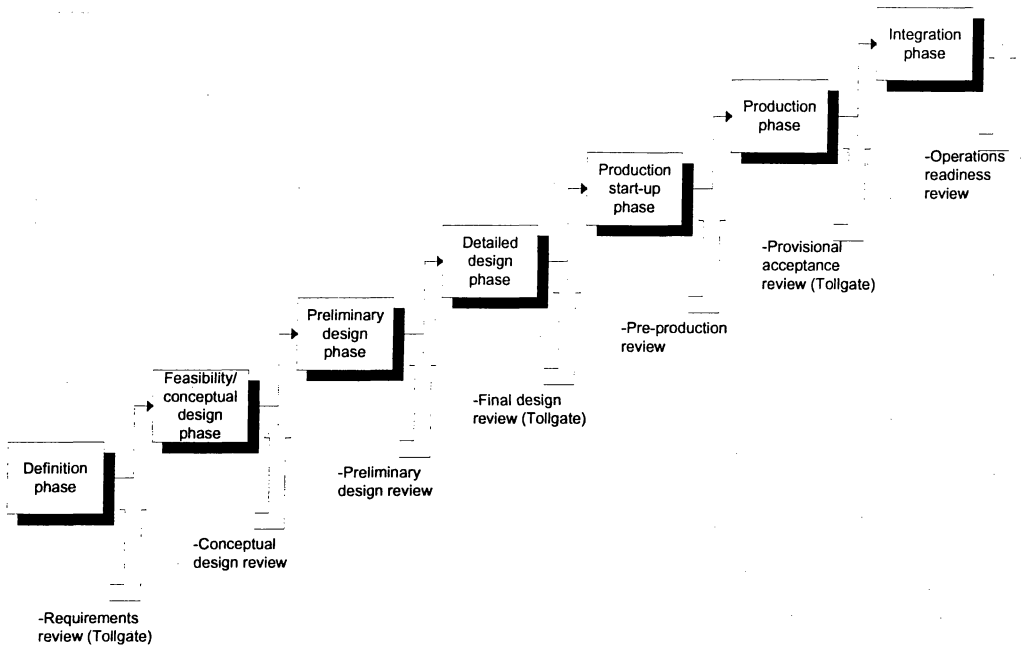


Figure 8.3, model of ALMA hardware development and production process

Note that some of the reviews in the ALMA development model have the status of tollgate. While a successful review milestone implies the formal completion of the previous phase, in addition at a tollgate explicit approval is given to commence with the next phase.

8.8 Software development process

Summary of "SE practices Software ...." Document by Chiozzi et al. To be added

8.9 Design standards

A set of common engineering rules that are in principle mandatory for the whole ALMA system will be defined. The purpose of these system wide requirements is:

- To assure integrity of sub-systems and complete ALMA system
- Effective design, construction and operation of the ALMA system
- To improve quality of the ALMA system

These system wide requirements will be described in a separate document and will address at least the following items:

8.11 Document management

It is recognized that documentation will play a vital role in the ALMA project. A/o documentation is the primary method to communicate information within the project organization and product design documentation is the only way, which defines the ALMA instrument and the sub-systems during all phases until final completion.

The ASEG will focus on and bear the responsibility for the management of engineering documents related to products that make up the ALMA instrument. However it is expected that the document management method will be general enough that other document types can be included without problem.

Important ingredients to the ALMA document management plan will be the following:

- Engineering document numbering system will be based on the ALMA Product Tree. This approach ensures a consistent and long lasting way of identifying product related documentation
- Archived documents, the originals, are only stored in electronic form. On paper, only copies of these original electronic documents are distributed when needed
- Submission, distribution and retrieval of documents is primarily done via Internet
- Different types of approval processes will be defined dependent on the type of document
- Mandatory standard file formats for general text documents will be introduced for the project. It is expected that for editable documents this will be Microsoft Office based file formats. For documents issued once only, storage in Adobe PDF format is advocated.
- For technical product documentation it is recognized that standardizing to a limited set of file formats is not feasible due to the large variety of CAD/CAE tools used across the project.
- Electronic distribution of document copies is primarily done in Adobe PDF format.
- The use of a modern Product Data Management (PDM) tool is essential in an efficient management of the project documentation. This tool must be able to handle the geographically distributed storage of documents into a single virtual database, assist in the approval and revision processes of documents, enable distribution and access of documents with proper security control

The complete ALMA document management plan will be described in a separate document.

8.12 Logistic and maintenance engineering

A maintenance philosophy for the operational phase of the instrument will be documented in a separate plan.
Already during the design of a product a maintenance requirements analysis must be carried out. This analysis will address at least:

- Mean Time Between Maintenance (MTBM) analysis
- Mean Time To Repair (MTTR) prediction
- Spare parts needs
- Necessary maintenance personnel skills and quantity
- Maintenance equipment

8.13 System safety

Based on international, as well as appropriate national standards, Chilean, European and U.S., a safety plan and system wide safety requirements will be established. Primary targets of the safety plan and system wide safety requirements is to preclude or limit hazards to personnel and equipment and to comply with applicable laws where needed during all life cycles of the instrument.
Safety plan and system wide safety requirements will be used as input for the design activity for all products.

ALMA Project Book, Chapter 9.

IF Processing and Signal Transmission

R. Sramek, W. Brundage,
J. Jackson, A. Baudry
2001-January-30

Revision History:

2000-December-12: First Completed Version

2001-January-30: Revised section 9.2 and 9.3; digitizer interface to the data-transmission system now 250 Mb/s; options for the Fast Demultiplexing Unit discussed. Minor text changes in section 9.1, 9.2.3, 9.3

Summary

The signal processing at the antenna from the IF (intermediate frequency) outputs of the Front-End to the digital inputs of the correlator is described in this chapter; the subsections are:

9.1) the Downconverter
9.2) the Digitizers
9.3) the Fiber Optic Data Transmission System.

The signals from the active front end at each antenna will pass through an IF selector switch within the front end package and then into the IF Downconverter. The Downconverter output signals will then go to Digitizers, then to the Data Transmission System, which will carry the digital signals to the central electronics building, then to the correlator.

9.1 The Downconverter

In each ALMA antenna there will be two Downconverter modules, one for each polarization, and the two inputs to each module will carry upper and lower side-band signals.

A block diagram of the Downconverter is shown in Figure 9.1.1 and the specifications are in Table 9.1.1 The input and output noise power spectral power distribution will be nominally flat over the passband as given in the specifications.

The Downconverter will take the wideband 4 - 12 GHz input signals received from the front end subsystem and produce four output signal channels each with a passband of 2 - 4 GHz suitable for bandpass sampling at by the digitizers, which are clocked at 4 GS/s

In addition to gain, frequency conversion and passband definition, the Downconverter module provides total power measurement of both the wideband input signal channels and the four octave band output channels. Also it provides switching which allows any output channel to tune either the upper (8 - 12 GHz) or lower (4 - 8 GHz) frequency portion of either input channel. The frequency tuning is provided by LO inputs in the 8 - 14 GHz range from four second LO synthesizers.

Table 9.1.1 Specifications for Downconverter

DOWNCONVERTER MODULE
SPECIFICATIONS for ALMA CONSTRUCTION

Document # ALMA06002NX0002

2000 November 15
W. D. Brundage

Reference: Block Diagram, Document # ALMA06002KX0002
*** indicates interfaces**

Number of modules	142 (2 x 64 antennas plus 14 spares)
***Inputs from front end**	
Number of inputs per module	Two: USB, LSB (upper and lower sidebands)
Frequency range, nominal	4-12 GHz or 4-8 GHz
Power level within any 2 GHz bandwidth when the antenna temperature is 290K	-40 +/-3 dBm, less loss of coax and connectors between front end outputs and module inputs (3m of phase stabilized Andrew FSJ1P-50A ¼ inch diameter, attenuation = 2.4 dB @ 12GHz)
Variation of power spectral density vs. frequency (flatness)	<+/-1.5 dB across the nominal frequency range
Headroom when the antenna temperature is 290K (see definition)	>20 dB
***Inputs from Second LO (LO2)**	
Number of inputs per module	Four (A, B, C, D), independently tunable
Frequency range	8.0-14.0 GHz nominal; frequency LO2 > frequency input
Power level	+13 +/-1 dBm
Power level of spurious frequencies	<-70 dBc, except <-40 dBc for 2nd harmonic
***Outputs to digitizers**	
Number of outputs per module	Four (A, B, C, D)
Frequency range	2 - 4 GHz nominal
Power level	-TBD +/-TBD dBm plus loss of coax and connectors between output and input to digitizer module
Headroom when the antenna temperature is 290K	>20 dB
LO2 spurious and leakage at outputs	<(power level -40 dB) for all combinations of frequencies of LO2-A, -B, -C, -D

Throughput from front end inputs to outputs to digitizers	
Input S_{11} reflection magnitude 4 – 12 GHz	<-20 dB (VSWR < 1.22) to minimize spectral ripples
Input noise figure 4 – 12 GHz	< 10 dB (2 610K); SP_{DC} < -164 dBm/Hz << SP_{FE} = -133 dBm/Hz
Image rejection	>20 dB
Filter, 4-12 GHz nominal bandpass for total power detection	passband <4.0 GHz and >12.0 GHz at -1 dB, max ripple +/-0.5 dB; stopband 3.5 GHz and 12.5 GHz at < -20 dB, 0-3.0 GHz and 13.0-18 GHz at < -40 dB
Filter, bandpass image reject	(may be revised after re-analysis of spurious mixer responses)
4-8 GHz nominal	passband <4.1 GHz and >8.4 GHz at -1 dB, max ripple +/-0.5 dB; stopband 4.0 GHz and 8.6 GHz at < -10dB, 0-3.0 GHz and 10-18 GHz at < -40 dB
8-12 GHz nominal	passband <7.6 GHz and >12.0 GHz at -1 dB, max ripple +/-0.5 dB; stopband 7.4 GHz and 12.4 GHz at < -10 dB, 0-6.0 GHz and 14-18 GHz at < -40 dB
Filter, outputs A, B, C, D (may be revised after re-analysis of mixer and digitizer spurious responses)	passband <2.1 GHz and >3.9 GHz at -1 dB, max ripple +/- 0.5 dB; stopband 0-2.0 GHz and 4.0-12 GHz at < -20 dB
Passband amplitude ripple	<1.0 dB peak-peak
Passband deviation from linear phase	<40 degree peak-peak
Gain stability	<0.1 dB peak-peak over 1 minute, <0.5 dB peak-peak over 60 minutes
Phase/delay stability	<10 degree peak-peak over 1 minute, <40 degree peak-peak over 60 minutes
Headroom[1] when the antenna temperature is 290K	>20 dB
Crosstalk (inverse of isolation) among any input and any unconnected output	>40 dB rejection
Attenuators in input path 4-12 GHz	
Steps	1 +/-0.3 dB
Range	>30 dB
Phase variation vs. attenuation	<20 degree peak-peak over attenuation range 0-20 dB
Deviation from linear phase vs. frequency 4-12GHz	<20 degree peak-peak over attenuation range 0-20 dB

Attenuators in output path 2 - 4 GHz	
Steps, nominal	0.25 +/-0.15 dB over attenuation range 0-20 dB
Range, nominal	>30 dB
Phase variation vs. attenuation	<10 degree peak-peak over attenuation range 0-20 dB
Deviation from linear phase vs. input frequency 4-12 GHz	<10 degree peak-peak over attenuation range 0-20 dB
Matching among all downconverters	
Amplitude vs. frequency	TBD
Phase vs. frequency	TBD
Total power detectors (TPD)	
Input path 4-12 GHz	
Number	two, one for each input channel
Response vs. input frequency at any LO2 frequency	< 2 dB peak-peak.
Output path 2 - 4 GHz	
Number	four, one for each output channel A, B, C, D
Response vs. input frequency at any LO2 frequency	< 1.5 dB peak-peak.
Linearity	<1 % deviation from square law over range -6 dB to +13 dB relative to antenna temperature = 290 K
Monotonic resolution of digitizer, minimum	16 bits for 13 dB headroom above antenna temperature = 290 K
*Readout	2 millisec integrations and dumps to MC-AMBTP card via serial or parallel interface
*Offset calibration	MC sets the input power to zero by either setting the preceding attenuator to >(20 dB + headroom) or by removing bias to the preceding amplifier
Stability of output relative to inputs from front end	<50 ppm in 1 second, <500 ppm in 60 seconds
*Interface to MC-AMBTP	dedicated total power data link to antenna bus master (ABM)
***MC control functions**	
Set levels of input total power detectors	1 byte for each of two attenuators
Set levels of output of each total power detector and input level of each output digitizer	1 byte for each of four attenuators
Set to zero all inputs to total power detectors	1 byte to remove bias to six amplifiers; or set all attenuators to maximum
Set all 3 matrix switches (select image filters for each output)	1 byte

***MC monitor functions**	
Total power detectors	3 bytes every 2 milliseconds for each of 6 detectors
Temperatures	2 bytes every 10 seconds for each of 8 locations
Supply voltages derived within module	2 bytes every 10 seconds for each of 8 voltages
***External power supply inputs**	+18 +/-0.5VDC @ <2.2A, -18 +/-0.5VDC @ <0.7A, +8 +/-0.3VDC @ <0.6A, +5 +/-0.1VDC @ <0.6A
Internal voltage regulators	
Output voltages @ amperes	+15 @ 2.2 (total of >1 regulator), -15 @ 0.6, +5 @ 0.6, -5 @ 0.1
Output regulation plus noise	0.01% peak-peak over time interval > 60 seconds
Timing generator	
*Inputs from Reference Receiver	25 MHz sine wave at 0 dBm; 20.833 Hz positive edge, 5V differential
Output for timing total power integration	500 Hz TTL pulses of >1 usec width synchronized to 20.833 Hz timing reference
Output for digitizer clock	TBD MHz to match digitizer; synchronized to 20.833 Hz timing reference
***Operational environment**	
Altitude	5000 meter (16,000feet)
Shock	Negligible
Vibration	TBD
Temperature of air flow past sides of module	Plenum temperature set 16 – 22 Celsius, variation < 2 C peak-peak
Air mass flow rate past sides of module	TBD
Specific heat of air flow	TBD
Packaging	
Module	3 to 6 width x 5U high x TBD depth standard module (ATNF) with extruded vertical heat fins on one side or both sides
*Multi-pin connector (power, MC-AMB, MC-TP)	One double density 100 pin D type [male]
*Coaxial connectors	12 OSP (M/A-COM) blind mating [male]

1. Define *headroom* as the dB ratio of *available power at 1% gain compression {P(-1%)}* to the *total system noise power {P}*. Typically, P(-1%) is 16 dB less than the available power at –1 dB gain compression and 26 dB less than the available power at third order intercept. -end-

Bill Brundage

2000 November 15

9.2 The Digitizers

9.2.1 Introduction, Top Level Specifications

The analog-to-digital converters, or digitizers, installed in the antennas provide the flexibility required for the fiber optic transmission of the IF. Digital conversion is of course indispensable to the correlator in order to derive the correlation function as a function of digital lags for spectroscopy. The digitizers are thus crucial and single-point-failure elements in the system. The ALMA system incorporates 3-bit digitizers thus improving the overall sensitivity compared to the classical 2-bit case.

The goal specifications are given in Table 9.2.1

Table 9.2.1

Input BW 2-4 GHz
Sample clock 4 GHz (250 ps)
Bit resolution 3 bits
Quantization levels 8
Aperture time ~ 50 ps
Jitter ~ a few ps
Small indecision region
Output demultiplexing factor 1/16
PLL Clock distribution 4 GHz, 250 MHz (125 MHz system clock)
Fine delay command
Low power consumption

A survey of tens of Web sites for commercial samplers or track/hold amplifiers with conversion rates above 1 Gsps show that the product required for ALMA does not exist off the shelf. Some commercial products go as high as 1-2 Gsps, one goes up to 4 Gsps but no digitizer has an input bandwidth up to 4 GHz . They are multi-application products with more bits than actually required for radio astronomy (and thus with high power consumption ~ 4 to 7 W).

9.2.2 Digitizer Overview

The block diagram of the ALMA digitizer is given in Figure 9.2.1. It includes several fundamental elements briefly described in the following Section. The input adapter amplifier, the comparators and associated latches and encoding are implemented in a single ASIC. The fast demultiplexing unit is separated from the ASIC to diminish any coupling of the digital output with the analog signal input. The ASIC and the demultiplexing unit form the digitizer proper.

The PLL box produces and distributes the sinusoidal 4 GHz sampling clock signal and the 250 MHz signal required by the demultiplexing unit. This is another separate unit which will be common to at least one digitizer pair. Fine delay setting may be obtained by controlling the sample phase in the PLL box. (Responsibility of the design of the PLL 4GHz box has not yet been attributed and is being discussed between both IRAM and NRAO.)

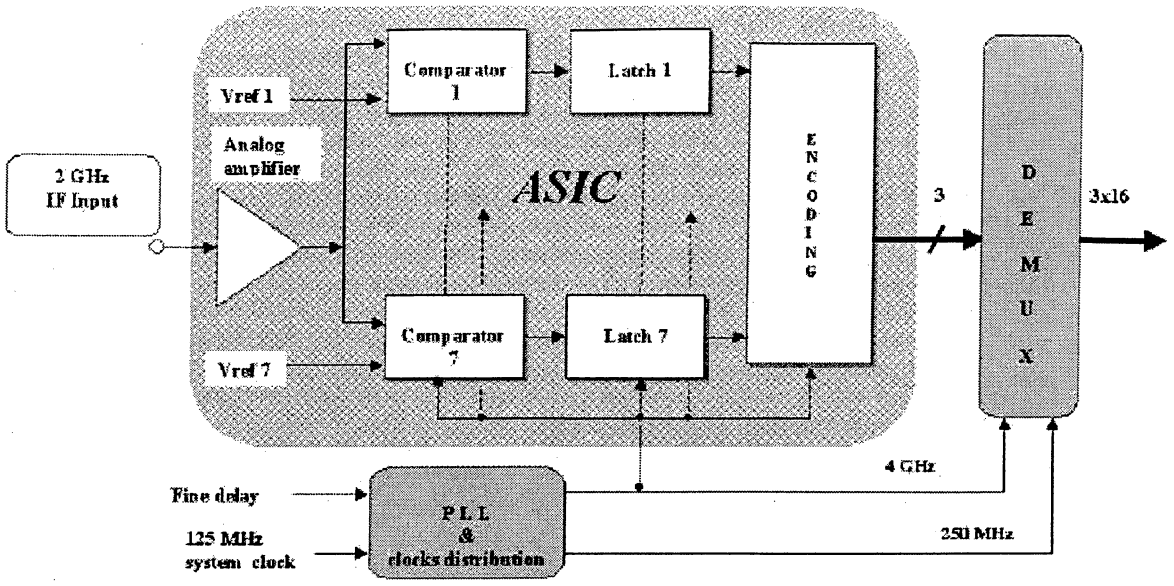


Figure 9.2.1 Digitizer Block Diagram

9.2.3 The Digitizers

Input Adapter Amplifier
The input analog signal is delivered from one of the four outputs (50 ohms impedance) of the IF down-converter module in the range 2 to 4 GHz. It is random with Gaussian statistics. The response of this amplifier is flat within +/- 0.5 dB over 2 GHz bandwidth and linear up to about +15 dB above the r.m.s. input signal voltage. The voltage supply required for the adopted ASIC technology is +/-1.25 V.

The digitizer input level is controlled in the IF downconverter with +/- 0.25 dB attenuators placed in the 2-4 GHz output paths of each IF downconverter. This allows us to minimize platforming effects and to keep the quantization thresholds constant and at their optimum level for maximum quantizing efficiency.

If the dynamical tests of the digitizer prototype would not be good enough (see Digitizer Test Bench below) further downconversion to the 0-2 GHz band just before the input amplifier stage could become necessary. This would be made with 4 GHz mixing from the PLL box.

Comparators and Quantization Thresholds
The sampling function is performed in the comparators which include two latches operated in a master-slave configuration and clocked at 4 GHz. The 4 GHz clock signal is equally distributed to 7 comparators. It is shaped internally in a dedicated amplifier driven by the external 4 GHz sinusoidal signal. The seven thresholds comprise a zero reference voltage and are set around +/- 0.5 'sig' , 1 'sig' and 1.5 'sig' where 'sig' is the r.m.s. voltage at the common input of the 7 comparators; these levels are kept constant and their exact value is tuned with an accurate division voltage chain so as to minimize the quantization losses. First simulations of SiGe digitizers indicate that the sampler indecision region is small and at the level of 1% of the smallest comparison threshold.

Encoding
The digitizer encoding is not yet finally adopted. It is not dictated by the correlator specifications because the digital FIR filters have look-up tables to translate between the digitizer code and the 4-level correlator chip code. However, the encoding should minimize the power consumption and should permit easy identification of the digitized signal sign (Walsh demodulation). We plan to deliver SCFL differential logic levels.

Adopted Technology
At the moment, the adopted technology is based on high speed SiGe bipolar transistors from ST-Microelectronics. In order to check the technology performances, the layout of an experimental ASIC comprising an input amplifier, two comparator-latches, two output buffers and clock distribution has been prepared with design tools and verification software from ST-Microelectronics. Simulations are encouraging and a first layout has been sent to the foundry. The technology will evolve from BiCMOS6G to BiCMOS7G all designs being made with 2.5 V DC supply voltage.

Packaging
The encapsuled sampler ASIC (as delivered from the industry) will be placed on a printed board and in a shielded enclosure with coaxial connectors. Each three bit output is sent through a coaxial connector to the input of each single bit demultiplexing board (if option a) described below in the Interface to the Data Transmission System is adopted), and each board is placed in a separate shielded enclosure. Shielding and coaxial connectors will minimize cross-talks and digital feedback to the sampler input.

Digitizer Test Bench
Thorough tests of the digitizer ASIC prototypes will be undertaken on a dedicated test bench. This bench comprises a broad band input noise source with low / bandpass filters (0-2, 2-4, 0-4 GHz). The 3-bit output data from the sampler ASIC are sent to a Fast Demultiplexing Unit (FDU) with 2 logic layers and synchronizers. The FDU data are then processed in a simplified (16 lags) low frequency correlator based on FPGAs. The demultiplexing stage of the test bench is identified as a 'risky' area in the design. Figure 9.2.2 shows the basic principles of the FDU.

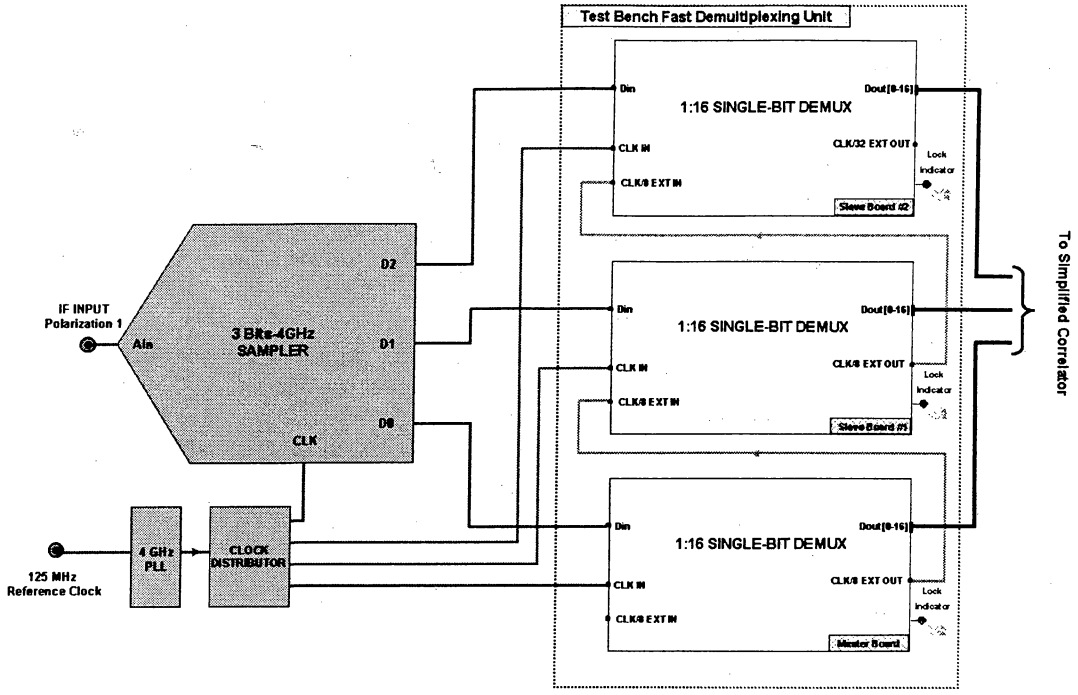


Figure 9.2.2 Test Bench Fast Demultiplexing Unit (FDU) Block Diagram

The FDU is made up of three single bit demultiplexing boards. Each board comprises a 1:16 demultiplexer and a synchronizer allowing multi-bit demultiplexing operation. The 1:16 demultiplexer consists of several logic layers the first one using commercial high speed GaAs ICs. The synchronizer can be seen as a PLL stage detecting whether the (equivalent) 1:16 logic counters of two chained demultiplexers operate in the same phase state or not.

Interface to the Fiber Optic Data Transmission System
Designing a series-reproducible version of the test bench FDU poses a significant problem because there are no commercial demultiplexers phasing 3 bit signals at high rates. We envisage three main possibilities: a) synchronize the three single bit commercial demultiplexers as performed in the test bench; b) use commercial high speed gate arrays, if available; c) develop a new ASIC. While option b) needs to be investigated carefully, we do not know at the moment whether it will be reliably possible with option a) to go to the production stage for hundreds of units.

The demultiplexed output is interfaced to the Virtual Parallel Bus (VPB) transmitter. It delivers 16 times 3 bits or 48 lines at 250 Mbps consistent with the 12 Gbps output data flow from each sampler ASIC, and consistent with the input of the VPB digital serializer and optical combiner needed for the IF/FO downlink.

We propose to place a digitizer pair (twice 2 GHz IF from two polarizations) and their 6 associated demultiplexing boards (option a) above) within another bigger shield to deliver 96 lines to the digital block of the VPB transmitter. Common shielding with the digital and optical VPB blocks should also be investigated provided that there is no RFI damage.

Alain Baudry
revised, 19 January 2001

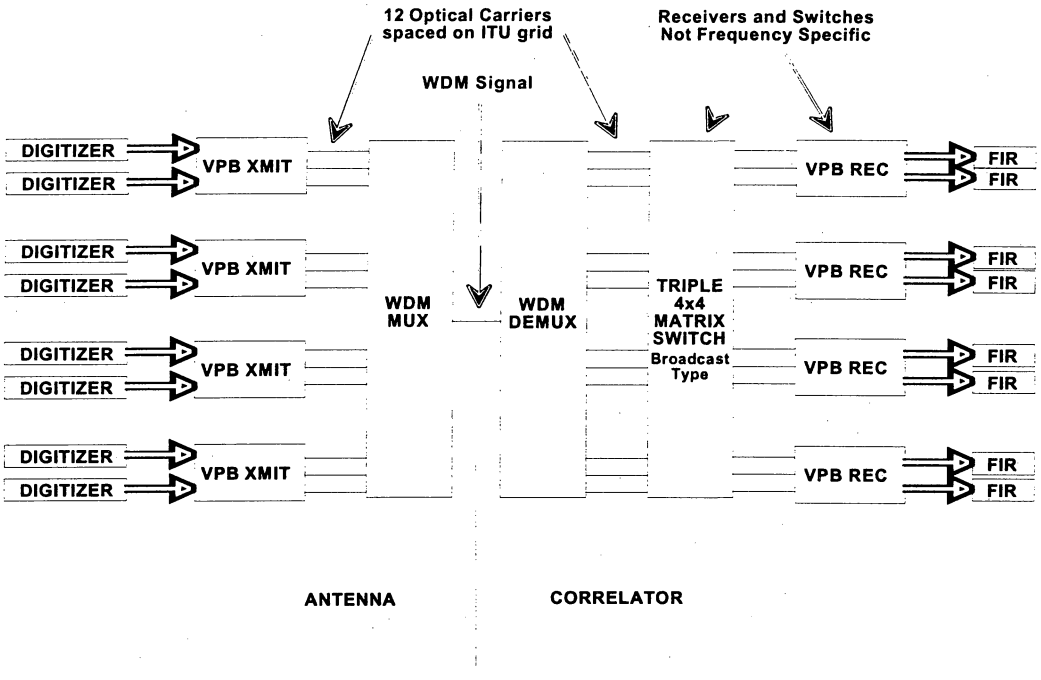
9.3 The Digital Fiber Optic Data Transmission System

The current hardware design of a Virtual Parallel Bus (VPB) system for the serialization, synchronization and transmission of the digital IF data from the ALMA interferometer antennas to the correlator is presented here. The transmission protocol will be presented in a future ALMA memo entitled "Digital Transmission System Signaling Protocol" by Robert W. Freund. The fiber optic IF link for the Atacama Large Millimeter Array (ALMA) and the ALMA Test interferometer utilize the same design.

The Virtual Parallel Bus system accepts 96 bits of parallel 250 Mbps data from a pair of digitizers in the antenna and transmits it over a fiber optic cable to the Correlator. The system includes a digital multiplexer, an optical transmitter composed of a laser and electro-absorptive modulator, a wavelength division multiplexer, single-mode fiber, optical de-multiplexer, an optical receiver, and finally a digital de-multiplexer.

Twelve links will be needed for each antenna in the ALMA array in order to transmit 3-bit data from eight digitizers to the Central Electronics Building. A full twelve-carrier system is shown in Figure 9.3.1.

Figure 9.3.1 Intermediate Frequency Fiber Optic System Block Diagram



9.3.1 Transmitter

A block diagram of a single IF data transmission channel for the Digital Fiber Optic IF Transmitter is shown in Figure 9.3.2. The current design contains all of the multiplexing and electrical to optical conversion on a doublewide AT style module with connections to the digitizers via a backplane board or high-density cable(s). The optical to electrical conversion hardware includes the laser with an integrated Electro-absorption modulator, and laser temperature and bias current control hardware. In addition, there is a provision for an optical wavelength locker and the associated feedback control circuit. Optical power monitoring is handled through the interface to the Laser Driver module. The card is controlled by a Monitor and Control (M&C) interface.

The following list specifies the components in transmitter module.

Power Supply: The board contains on-board regulators to produce the various voltages required by the VPB from the 48-Volt DC input voltage. The voltage requirements are +1.8VDC, +2.5VDC, +3.3VDC, +5.0VDC, -5.2VDC, and +8.0VDC. Other voltages may be required and can be produced with inexpensive 48VDC input switching regulators.

On-Board Control System: The on-board control for a single VPB transmit module consists of three devices, an AMBSI2 (formerly known as ZASI) module and two PIC 16C74 microcontrollers. The first 16C74 microcontroller controls internal operation of the three XILINX Virtex–E FPGA's. The second PIC 16C74 microcontroller provides all necessary monitor and control functions required by the three laser power controllers and the three laser temperature control modules. Both 16C74 microcontrollers communicate with the AMBSI2 via an SPI bus connection. The AMBSI2 module interfaces to the CAN based ALMA Monitor and Control System.

Clock distribution: This block uses the 125 MHz sine-wave system clock for the generation of a 625 MHz clock for the 10Gbps Multiplexer IC, and a 312.5 MHz clock for the FPGA. The board also requires a 20.833 Hz signal at LVDS levels,

625 Mbps Serializer/Protocol Engine: A Xilinx Virtex-E FPGA. This device will contain all of the logic for multiplexing the 250 MHz input data up to the 625 MHz rate required by the 10Gbps Multiplexer IC. It also provides all framing and synchronization functions as described in the Signaling Protocol document.

10 Gbps Serializer: This consists of a Giga GD16555/85 10Gbps multiplexer IC and associated electronics. The multiplexer converts the 625 MHz 16-bit parallel data into a 10 Gbps serial data stream for transmission over the fiber optic link. This circuit is essentially as shown in the data sheets, application notes and evaluation board design from Giga.

Driver Amplifier: Amplifies differential CML output of 10Gbps Multiplexer IC to sufficient levels to drive the optical Electro-Absorption Modulator (EAM).

Optical/RF Board: This includes the integrated laser / EAM package plus the laser power and temperature controllers.

Wavelength Locker: This is an optional item and is not shown because it is not necessary for the basic operation of the system. It may be necessary to use a wavelength locker and feedback control to keep the operational frequency of the laser from drifting out of the passband of the DWDM components and to minimize time-of-flight variations due to frequency dependence of the index of refraction in the fiber. It is

estimated that lasers currently available may drift out of the optical passband of the DWDM components in 2 to 4 years time. However, new components are becoming available in the market that advertise ultra-stable frequency operation, maintaining operation in a 100 GHz bandwidth over as much as 20 years, which would obviate the need for wavelength lockers. Wavelength lockers are not an attractive option because of their extremely high cost. Wavelength lockers currently cost over $1,000 each and require additional circuit complexity to implement.

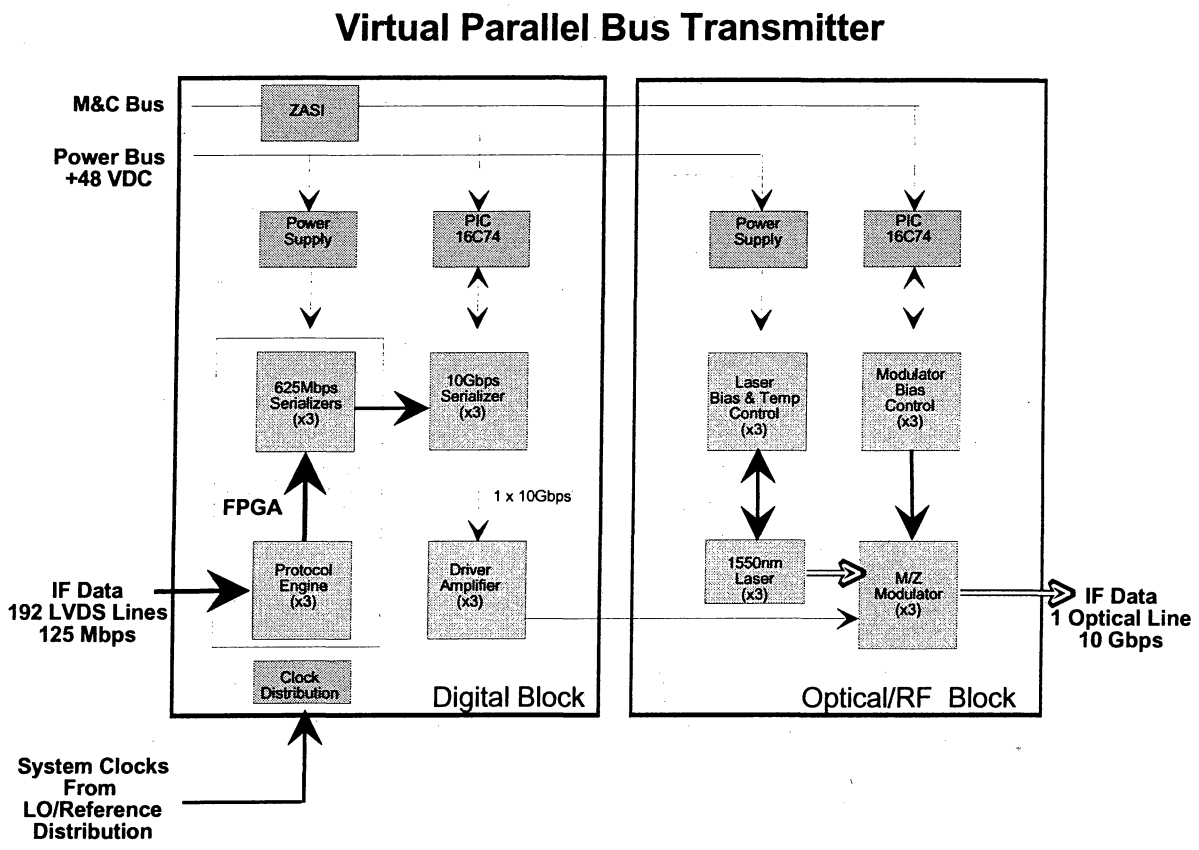


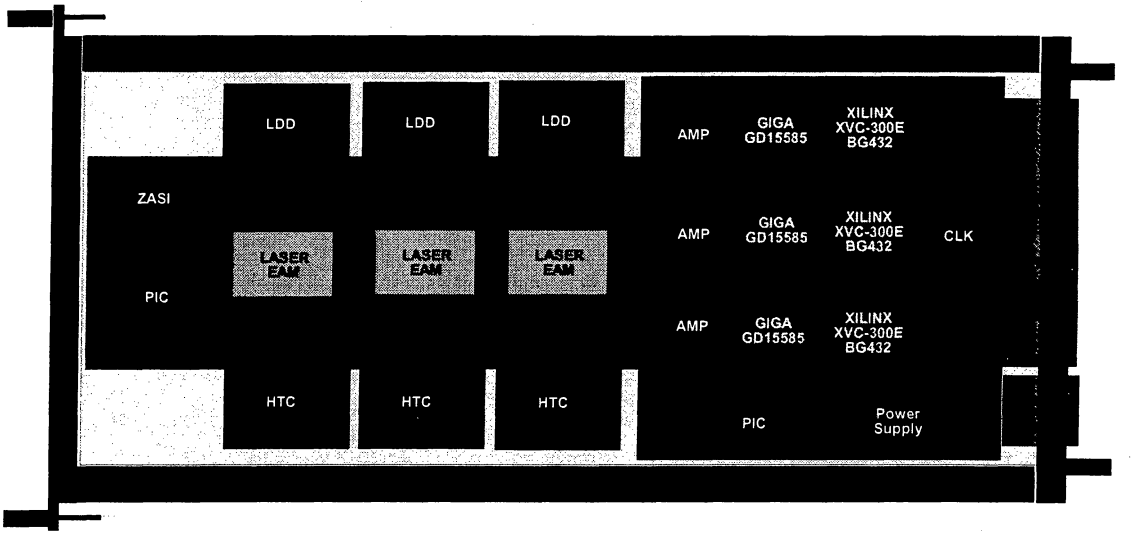
Figure 9.3.2 Transmitter Block Diagram

Packaging: The Transmitter will be mounted in a doublewide RFI shielded Australia Telescope (AT) style module. Data will be fed into the modules on high-density connectors through a backplane or high density shielded cables. The choice of connection method will be driven primarily by RFI considerations The transmitters will have a minimum of two printed circuit boards, one carrying the digital electronics and M&C interface, and the other holding the Optical components and associated electronics. The separate Optical board simplifies sparing since each laser must be tuned to an operational frequency in order to match the transmission passbands of the DWDM components while the digital boards are consistent throughout all VPB modules. Figure 9.3.3 shows a conceptual drawing of the VPB transmitter module. A third

mezzanine board may be used for the 10Gbps multiplexer IC to simplify replacement of that device due to a rapidly changing market.

The output of this transmitter module is three optical carriers that will be combined with the carriers from the remaining three VPB modules using WDM techniques for transmission to the VPB receivers in the correlator rack over a single fiber. Connections to the optical fibers from each module are made through Diamond E-2000 blind-mate backplane connectors. The use of blind-mate backplane connectors reduces maintenance costs by enabling easy replacement of the modules, and significantly reduces the chances for damage during replacement. The fibers are all routed to a separate DWDM module.

Virtual Parallel Bus Module - Right Side Interior View



- Left Side Panel is Extruded Heat Sink
- Heat Generating Components are mounted on the Heatsink side of the PCB's and directly coupled to the Heatsink.

Figure 9.3.3 Conceptual diagram of a two channel Transmitter Module.

9.3.2 Receiver:

Figure 9.3.4 shows the functional block diagram of a receiver. The optical signal is received by a *pin* photodiode and the photo-generated current is used to recreate the 10Gbps digital signal, and to determine the average received optical power. The relatively weak digital signal is passed on to an amplifier that generates the voltage levels needed by the demultiplexer, which converts the data into 16 parallel streams at 625 Mbps. The FPGA does the final demultiplexing to 64 streams at 125 Mbps.

PD & Amps: This includes a *pin* photodiode, amplification stage, and an optional electrical filter. The output is a 10 Gbps data stream that is fed to the 10 Gbps Demultiplexer IC. The electrical filter is a low cost item that is typically used for SONET applications in the telecom industry.

10 Gbps demultiplexer : The first demultiplexer stage is a Giga GD16544 or GD16584 10Gbps demultiplexer IC and the associated electronics. It converts the 10 Gbps serial data stream from the fiber optic link to 625 MHz 16-bit parallel data. This circuit is essentially the same as shown in the data sheets, application notes and evaluation board design provided by Giga.

Receive Shift Register/Protocol Engine: A Xilinx Virtex-E FPGA converts the 16 bit 625 MHz data from the demultiplexer IC into 80 bit 125 MHz data for presentation to the Correlator. The FPGA implements all synchronization, timing and data integrity functions as described in the Signaling Protocol Document.

On-Board Control System: The On-board control for a single VPB receive module consists of three devices, a AMBSI2 module and two PIC 16C74 microcontrollers. The first 16C74 microcontroller controls internal operation of the three XILINX Virtex–E FPGA's. The second PIC 16C74 microcontroller provides all necessary monitor and control functions required by the three optical receivers. Both 16C74 microcontrollers communicate with the AMBSI2 via an SPI bus connection. The AMBSI2 module interfaces to the CAN based ALMA Monitor and Control System. A list of interface commands has been generated, but needs revision to reflect the M&C protocol. Further modification to this list will also be necessary as the design evolves. This information will be detailed in the appropriate Interface Control Document.

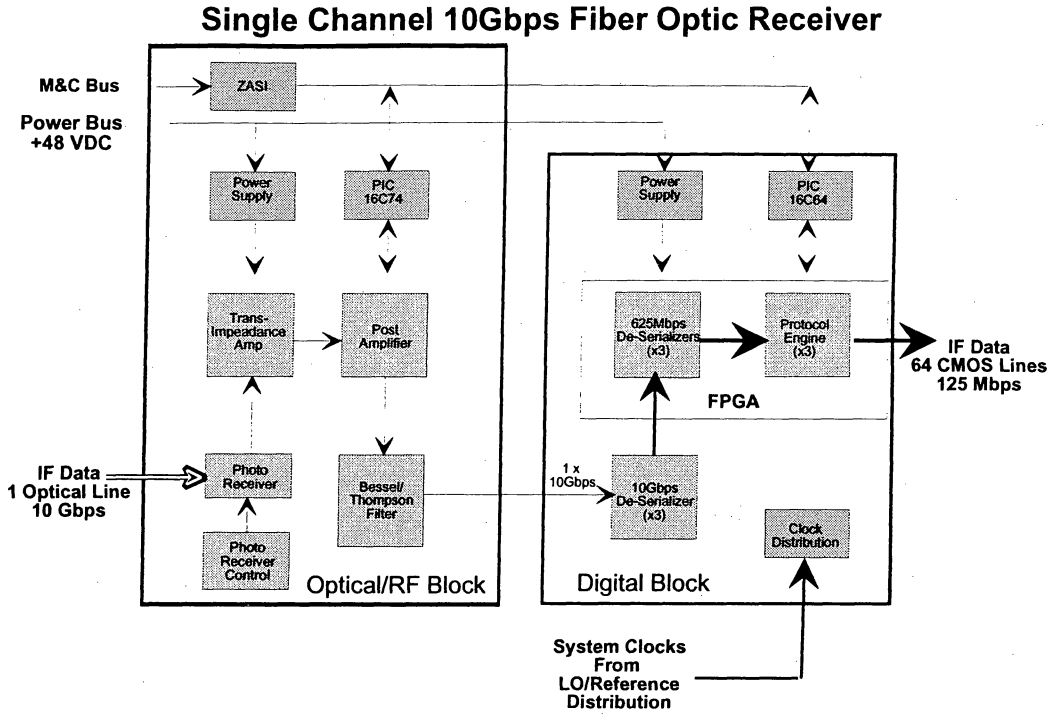


Figure 9.3.4 Receiver Block Diagram

Packaging: The VPB Receiver is packaged in a 6U x 280mm size Euro Card of the same general design as the cards in the ALMA Correlator. These boards will be installed in the correlator rack with the FIR filter cards. The optical portion of the module is on a separate card to aid in sparing and future modification issues. Heat generating components are cooled with individual heatsinks bonded to the components. Figure 9.3.5 shows the VPB Main Boar d and Figure 9.3.6 shows the Optical Board.

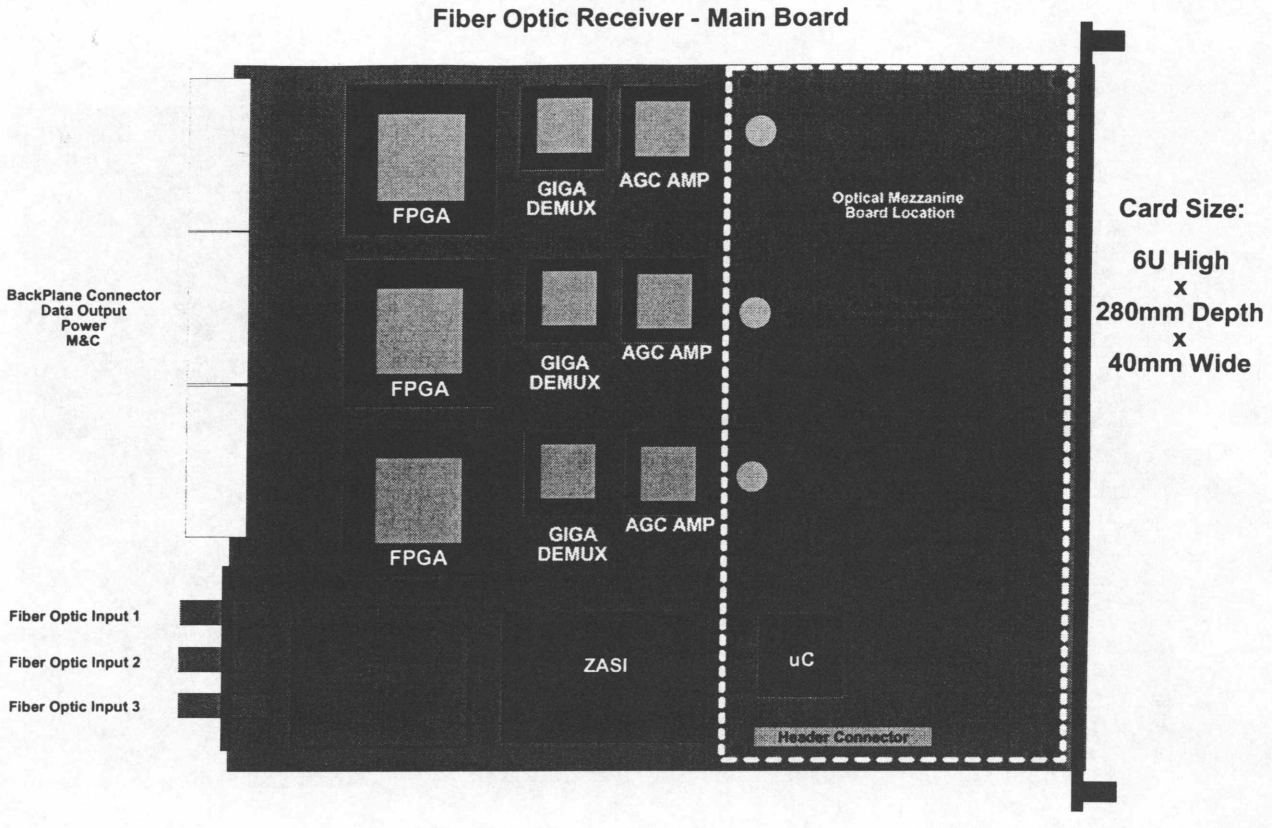


Figure 9.3.5 VPB Receiver Main Printed Circuit Board

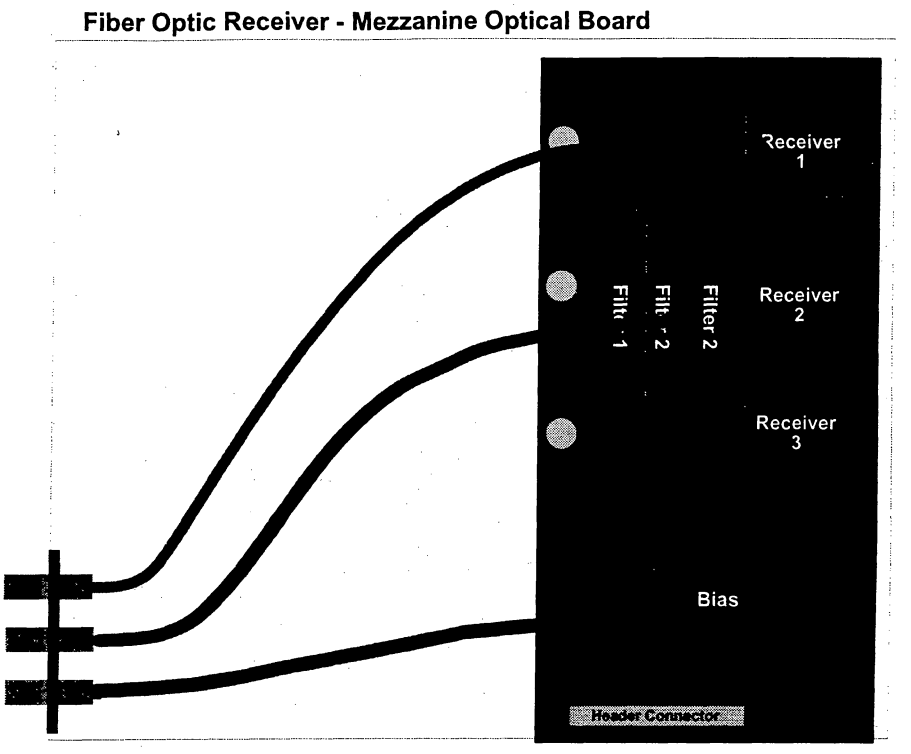
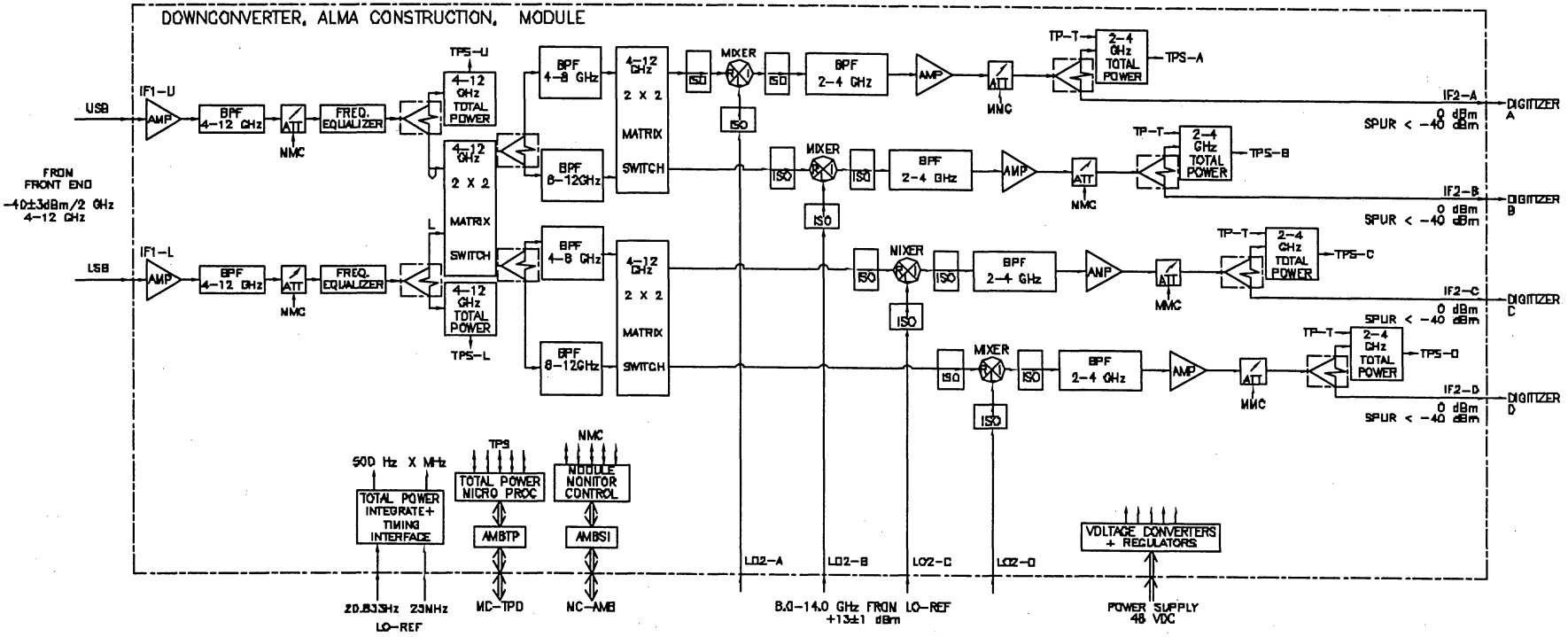


Figure 9.3.6 VPB Receiver Optical Mezzanine Board

Jim Jackson
2001 January 22

Figure 9.1.1 Block Diagram of Downconverter

REV	DATE	DRAWN BY	APPRV'D BY	DESCRIPTION



ACAD : 06002KX0002

UNLESS OTHERWISE SPECIFIED DIMENSIONS ARE IN INCHES. DIMS. ARE IN [MILLIMETERS]

TWO PLACE DECIMAL (.XX) ±	.01
THREE PLACE DECIMAL (.XXX) ±	.005
FOUR PLACE DECIMAL (.XXXX) ±	.0005

MATERIAL :

FINISH :

NEXT ASSEMBLY | DWG. TYPE

ALMA

ANTENNA SIGNAL PROCESSING (ASP) DOWNCONVERTER (DC)

ALMA CONSTRUCTION BLOCK DIAGRAM

NATIONAL RADIO ASTRONOMY OBSERVATORY
SOCORRO, NEW MEXICO 87801

DRAWN BY M. SULLIVAN	DATE 2000/09/29
DESIGNED BY B. BRUNDAGE	DATE 2000/09/29
APPROVED BY B. BRUNDAGE	DATE 2000/10/02

SHEET NUMBER 1 OF 1 | DRAWING NUMBER ALMA06002KX0002 | REV. – | SCALE NONE

ALMA Project Book, Chapter 10

ALMA CORRELATOR

John Webber
Ray Escoffier
Chuck Broadwell
Joe Greenberg
Alain Baudry
Last revised 2001-02-07

Revision History:

1998-09-18: Added chapter number to section numbers. Placed specifications in table format. Added milestone summary.

1999-04-09: Revised milestone dates and made date format conform to adopted standard. Revised tables and some text to reflect adoption of digital FIR filter. Changed text to reflect architectural change in delay line implementation. Revised block diagram.

2000-02-04: Changed maximum number of antennas to 64. Changed block diagram to correspond to current thinking about system architecture.

2000-03-31: Revised from MMA to ALMA baseline correlator. Incorporated changes resulting from correlator and systems PDRs.

2000-04-12: Clarified path of digitizer development and meaning of subarrays.

2000-12-05: Minor revisions to a few numbers. Addition of section describing possibilities for a future correlator.

2002-02-07: Minor revisions to text, mostly editorial. FX correlator reference added.

Summary

This section describes the ALMA correlator. The design described here is for a lag correlator with a system clock rate of 125 MHz. The goals of Phase 1 are to produce paper designs and some simulations of all major correlator elements, including the correlator chip, and to fabricate and test prototype hardware. The goals of Phase 2 are to produce a prototype minimally populated correlator, deliver such a prototype for use in the test interferometer, and deliver the complete correlator to the ALMA site.

Table 10.1 ALMA Correlator Specifications

Item	Specification
Number of antennas	64
Number of baseband inputs per antenna	8
Maximum sampling rate per baseband input	4 GHz
Digitizing format	3 bit, 8 level
Correlation format	2 bit, 4 level
Maximum baseline delay range	30 km
Hardware cross- correlators per baseline	1024 lags + 1024 leads
Autocorrelators per antenna	1024
Product pairs possible for polarization	HH, VV, HV, VH (for orthogonal H and V)

Although the specification is for 64 antennas, the design may be changed to accommodate a larger or smaller number of antennas with some impact on schedule. However, once the production units are begun, no change in the maximum number of antennas can be made without substantial redesign.

Table 10.2 Principal milestones for ALMA correlator

Preliminary Design Review	2000- 01- 20
Prototype correlator Critical Design Review	2001- 07- 31
Deliver prototype correlator to VLA site	2003- 05- 30
Deliver first quadrant to Chajnantor site	2004- 06- 18
Deliver last quadrant to Chajnantor site	2006- 10- 06

10.1 System Block Diagram

The system architecture described has been chosen as the best tradeoff to produce high reliability, robust operating margins, a minimum number of integrated circuits, and a minimum number of cable interconnects (see ALMA Memo 166). The performance of this architecture permits high versatility in correlator operation (see ALMA Memo 194). The adoption of a digital FIR filter eliminates many potential sources of systematic error (see ALMA Memo 204 and ALMA Memo 248).

The correlator system envisioned for ALMA includes digital filters, mode selection, a delay line and data format conversion stage, cross- and auto- correlators, long term accumulation, and initial digital computer processing.

A simplified functional block diagram for the ALMA correlator is given in Figure 10.1. This diagram presents a fairly conventional lag correlator except for the presence of the packetization stage.

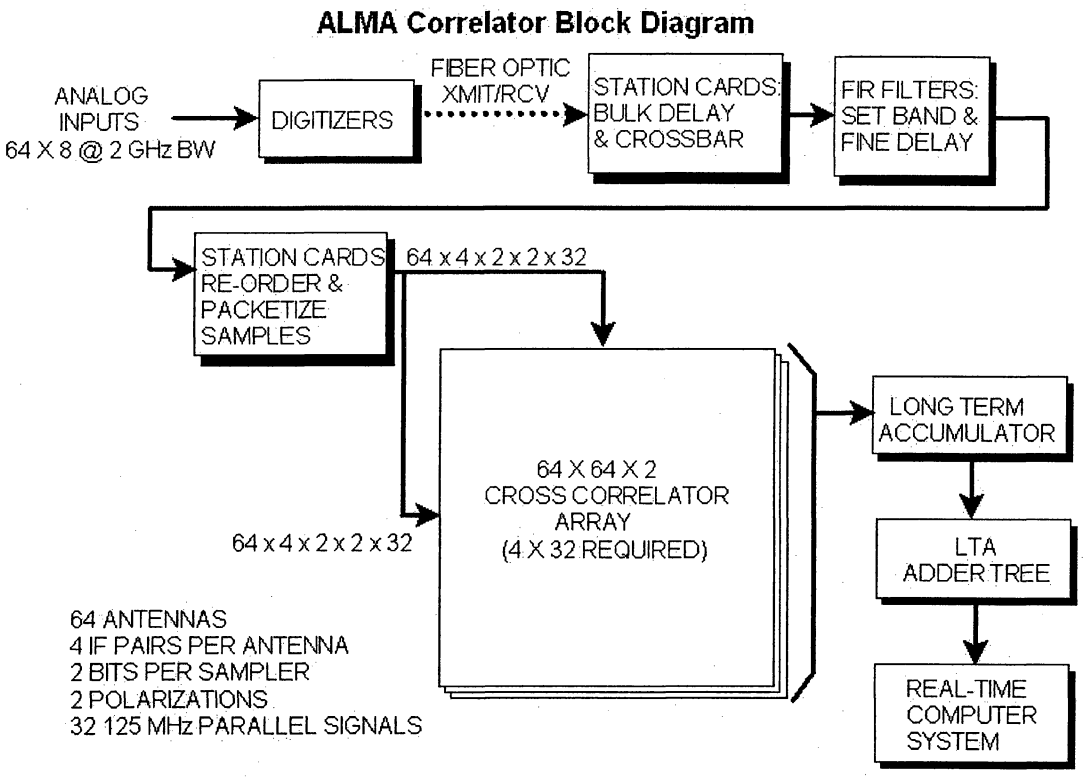


Figure 10.1: Simplified correlator functional block diagram. The digitizers, fiber optic IF transmission system, and real-time computer system are not part of the correlator but are shown for clarity. Computers which control the correlator are omitted.

The analog outputs of the baseband system drive digitizer inputs where 3- bit, 8- level digitization is performed at 4 GSamp/second. For details, see section 9.2. The samples are transmitted over fiber optic cables at a rate of 96 Gbit/s from each antenna to the correlator.

When less than 2 GHz bandwidth is desired, the samples are used as the input to the digital filter. The use of 3- bit quantization at the FIR filter input results in a small (~4%) loss of SNR compared to perfect analog filtering; the output re- quantization to 2 bits provides suitable input to the correlator. State counters are provided on both the input and output stages to permit setting analog levels to the digitizers correctly and for use in total power calibration.

Logic in the station cards routes outputs from the digital filters into the station cards which perform the packetization function. When fewer than 8 digitizers per antenna are being used, this stage will assure high system efficiency by replicating active digitizer outputs into unused memory areas and hence into otherwise unused correlators where additional lags can be generated. In this way, maximum performance will be obtained for the observational mode desired.

The digital filter stage will also do the sample decimation for observations in which sample rates less than 4 GS/second are needed. A 32- sample delay is required just before the digital filter in order to perform the finest resolution delay adjustment.

Adjusting the signals to the appropriate timing by means of a bulk delay is provided on the station cards which precede the FIR filters, in very efficient high density RAMs. For a 30 KM delay range, 524,288 RAM bits per digitizer output bit are required.

The packetization block seen in figure 10.1 will take the 32 parallel outputs of each digitizer and, using RAMs, both generate lags and re- sort the samples. In this block, the 32 parallel outputs of a high speed digitizer would be converted from each carrying every 32nd sample to each carrying short (about 1 msec) bursts of contiguous samples. If the N- wide parallel (2- bit) output of a high speed digitizer (each output carrying every Nth sample) were to drive the correlators using a conventional architecture, an N- by- N matrix of correlators would be required to insure every sample is correlated with every other sample. For N = 32, this would mean a matrix of 1024 small correlators to correlate the output of every baseband input of every baseline.

By using the format conversion scheme, the 32- wide parallel output from a high speed digitizer will be transformed into 32 parallel signals each carrying 1 millisecond data packets of contiguous samples that need drive only an N- by- 1 array of correlators. This simplification in the correlator circuit requirements is obtained at the cost of an inefficiency of about 0.2% which results because the end bits in adjacent 1 msec time segments of samples will not get correlated with each other.

(Note that the conversion from a conventional N- by- N architecture to an N- by- 1 architecture does not improve the spectral resolution performance of the correlator. The performance is set by the number of hardware correlators in the system. The conversion does, however, greatly simplifying the system wiring in that all N- by- N signals from two

antennas do not have to be wired to closely spaced electronics, thus simplifying the wiring matrix driving the cross correlators as well as reducing the number of I/O pins required by logic cards and integrated circuits.)

An additional benefit of the format conversion strategy is that it allows the system the same advantage as a recirculating correlator: when the bandwidth being processed is reduced by a factor of 2, the number of lags the system is capable of generating goes up by a factor of 2. This results in a factor of 4 increase in frequency resolution for a factor of 2 decrease in bandwidth.

Still another advantage of the format conversion (by far the most important in the ALMA correlator) is that it allows a minimum cable interconnect complex between the station electronics and the correlators. It also eliminates any requirement to interconnect correlator arrays in low bandwidth modes. Since the number of data interfaces between these two stages in the ALMA correlator surpasses that of any other astronomical correlator system by a factor of almost 100, this aspect of the system architecture is most important.

The cross correlator matrix of figure 10.1 is used to correlate the digitizer outputs of every antenna with those of every other antenna. At the intersection of any antenna X and another antenna Y in this matrix, there will be a correlator chip. This correlator will compute lag products for the XY baseline while the antenna Y and antenna X intersection of the matrix computes the baseline lead products. Auto correlation products for each antenna are obtained from correlators on the matrix diagonal.

In order to minimize further the station electronics to cross- multiplier cable interconnect, a very compact cross correlator matrix is essential. The design for the ALMA correlator places an entire 64 X 64 cross correlator matrix on 4 adjacent printed circuit cards, constituting a correlator *plane*. Each plane handles a 1/32 slice of the 4 GHz decimated digital data stream, at a 125 MHz data rate.

The distribution of signals from the station electronics to interface boards ("paddleboards") on the rear of the correlator planes assures that no signal drives more than one load. Two versions of the paddleboard will be produced: the first version will distribute the signals such that all IF signals from up to 32 antennas can be fully processed using only one quadrant of the correlator; the second version will distribute the signals such that all IF signals from the full 64 antennas can be processed using the entire correlator. This allows for interim operation of up to half of the array with one quadrant of the correlator while the remainder of the correlator (and antennas!) is built.

One disadvantage of this architecture is that once the number of antennas for the array has been set, future expansion of the correlator beyond this number is not practical.

The custom lag correlator chip has a dual 4- by- 4 array of correlators, each handling 2 polarizations. The chip can be programmed via a microprocessor supplied program word for its position in the matrix, which will select one of three correlator configurations:

1. four short correlators to compute the lags of all 4 polarization products (HH, VV, HV, and VH).
2. two longer correlators to compute just the lags for the two polarization components (HH and VV).
3. a single long correlator to compute lags for only one of the two baseband

inputs.

The estimated size of this custom correlator chip is approximately 2,000,000 gates. It will run on a 125 MHz clock. The chip package will an industry standard surface mount package with 240 pins.

Each individual lag of the correlator chip consists of a 2- bit, 4- level, times 2- bit, 4- level, multiplier whose output is integrated in a 25- bit accumulator. Each accumulator has secondary storage for the 16 most significant bits. Each chip contains 256 * 16 = 4096 of these lags.

There are eight Xilinx FPGAs on each correlator card for reading the accumulated results from the correlator chips and transmitting the results to the Long Term Accumulator Cards.

Each correlator card also has five Xilinx FPGAs to perform the Analog Sum function, which is required for forming a single beam from a set of participating antennas for Very Long Baseline Interferometry.

For observations in which fewer than 8 baseband inputs are being used, more lags can be produced by dedicating more than one correlator array to process the outputs of active baseband inputs. In this case, cards in the data format conversion stage will be used to form a virtual connection, the effect of which is to link two or more correlator arrays in series. The delayed input to the correlator chips that are to compute the higher level lags will be displaced in time the appropriate number of bits by offset RAM addressing in the data format conversion cards.

The long term accumulation block seen in figure 10.1 integrates the correlator outputs for the desired duration. The correlator chips will produce a total of 1,048,576 lag results in one plane. The long term accumulation block must provide double buffered integration storage for every result (since in some modes, every plane has distinct sets of lag results) in a total of four separate bins. This is a requirement of 1,073,741,842 storage locations, spread over 64 long term accumulator cards, or 16,777,216 results per card.

The adoption of a digital FIR filter has a potential system- wide consequence: it makes more attractive the baseline plan of performing the digitization at the antenna and transmitting the data to the correlator over a digital rather an analog fiber optic link. This is due to the fact that, with analog filters, sampling at the antenna implies placing the analog filters at the antenna, with resulting stringent specifications on filter temperature stability which could be difficult to meet. The advantage of digitizing at the antenna is that the limited SNR and gain instability of an analog fiber optic link are eliminated. The disadvantages are possible shielding difficulties for the sampling clock and the (at present) high cost of digital transmission compared to the cost of two 8 GHz wide analog channels.

10.2 Performance

This section gives performance parameters for some typical operating modes of the ALMA correlator. The ALMA correlator will be programmable on a baseband by baseband basis and, hence, some baseband inputs may be processed in one mode

while other baseband inputs are processed in other modes.

Bandwidths per baseband input range from a maximum of 2 GHz down in factor of 2 steps to 31.25 MHz. For 8 baseband inputs per antenna, this yields a maximum bandwidth per antenna of 16 GHz.

Sub- arrays will also be possible using the ALMA correlator. The maximum number of correlator sub- arrays for ALMA is limited to 16 by the adoption of automatic transfer of results from the Long Term Accumulator to the VME computers which will receive the data.

There are 8 digitizers per antenna. The baseband inputs driving the digitizers will consist of 4 dual polarization pairs; for each pair, 4 polarization cross- products may be computed. Each digitizer is assumed to digitize at 4 GHz and hence to be driven by RF signals at most 2 GHz in bandwidth. The maximum bandwidth processed is thus 16 GHz split into 2 GHz pieces. Note that the analog baseband constraints of the planned ALMA baseband processing system will impose limits as well.

The smallest division of lags in the projected correlator chip is 64 lags. Because of the architecture, this will produce 64 lead and 64 lag channels and hence 64 spectral points per product. This smallest correlator division means that in the full- up configuration, all baseband inputs active at maximum bandwidth and all 4 polarization products being computed, 64 spectral points will be produced for every baseline, every spectrum. This gives a frequency resolution per spectral channel of 31.25 MHz.

Given the full- up performance as defined above, the number of lags that the correlator can produce for a given experiment results from the following considerations:

1. If polarization cross- products are not required, a factor of 2 more lags (finer resolution) can be obtained. The particular configuration can be selected on a baseband pair by baseband pair basis.

2. If fewer than 8 baseband inputs are required, lags go up as 1 over the fraction of baseband inputs used (1/2 the baseband inputs, 2 times the lags).

3. If a lower bandwidth than 2 GHz per baseband input is required, lags go up as 1 over the fraction of maximum bandwidth (1/4 the maximum bandwidth, 4 times the lags) until a factor of 32 is reached. After that, the number of lags stays constant. The particular configuration can be selected on a baseband by baseband basis.

Note that item 3 implies the characteristic described above that for each reduction by a factor of 2 in bandwidth, an increase of a factor of 4 in resolution is obtained (up to the factor of 32 limit after which the resolution improves by only 2 for each factor of 2 reduction in bandwidth).

Table 10.3 below illustrates some of the possible modes. The first four columns relate to the correlator proper. The columns relating to velocity range and resolution assume 90% of the analog bandwidth will be usable. (See ALMA Memo 194 for additional illustration of the ALMA correlator performance.)

Table 10.3 Selected correlator modes

# of Digitizers	Bandwidth/ Digitizer	Cross-pol Products?	Channels/ Product	At 230 GHz, in velocity space:	
				Range	Resolution km/s
8	2 GHz	Yes	64	9391	40.8
8	2 GHz	No	128	18783	20.4
8	1 GHz	No	256	9391	5.1
8	500 MHz	Yes	256	2348	2.5
8	250 MHz	No	1024	2348	0.32
4	2 GHz	Yes	128	4696	20.4
4	1 GHz	No	512	4696	2.5
4	500 MHz	Yes	512	1174	1.3
4	250 MHz	No	2048	1174	0.16
2	2 GHz	Yes	256	2348	10.2
2	1 GHz	No	1024	2348	1.3
2	500 MHz	Yes	1024	587	0.64
2	250 MHz	No	4096	587	0.08

Two natural time intervals associated with the correlator are 1 msec and 16 msec. These are the two short term integration cycles available in the correlator chips. The 1 msec short term integration cycle is available only on the array diagonal (auto-correlation results only). The 16 msec cycle is available both on and off the array diagonal (auto- and cross- correlation results). The Long Term Accumulator (LTA) provides longer term accumulation for the 16 msec results, in integer multiples of 16 msec, up to approximately 65 seconds maximum. The LTA does not provide longer term accumulation for the 1 msec results. It does provide buffers for 16 consecutive sets of 1 msec results, so access to the results is on 16 msec boundaries instead of 1 msec boundaries.

The function of the adder tree block seen in figure 10.1 varies with correlator mode. At

full bandwidth (2 GHz), the lag results from all 32 planes must be summed together in the adder tree, while at minimum bandwidth, distinct sets of lags are produced in each plane and must be passed through the adder tree block. Intermediate bandwidths require intermediate sets of planes to be summed.

In cross- correlation mode, using 16 msec integrations at minimum bandwidth, a 32 GByte/sec output rate would be required, if all results were transmitted out of the correlator. The correlator output capacity is specified as 1 GByte/sec (64 MByte/sec on each of 16 streams). Alternatives of longer integration times, restricted numbers of lags, or fewer baselines are provided, allowing lower output rates.

The current functional specification of the LTA is given in ALMA Memo 294.

10.3 Size and Power Requirement Estimate

Table 10.4 Preliminary ALMA correlator module & printed circuit card requirements

Item	# required	Size	Power
FIR filter card	512	6U euro card	80 w
Station card	512	6U euro card	20 w
Correlator card	512	9U euro card	160 w
Control card	160	6U euro card	40 w
Long term accumulator	64	9U euro card	60 w
TOTALS	1760		143,360 w

It is estimated that the station- dependent part of the system (digitizer, filter, mode, and memory) will require 1/4 a rack per antenna, or 16 racks for 64 antennas. The remainder of the system, proportional to the number of antennas squared (correlator, control, and accumulator) will occupy 16 racks for 64 antennas. The grand total of racks is therefore about 32.

The power estimates given in Table 10.4 above are based on the experience gained in the development of the GBT spectrometer. The biggest unknown at this time is the dissipation to be expected in the custom correlator chip, 32,768 of which will be required in the system. The GBT correlator chip dissipates about 5 watts with a clock rate of 125 MHz. Such a high chip dissipation in the ALMA correlator would mean both high system power requirements and lower reliability because of the difficulty in removing the heat from the system at the high altitude site.

By using low voltage chip technology it is predicted that the custom correlator chip described in this document can be built with about a 1.5 watt power requirement, 2 watt maximum. The chip represents about a factor 2 increase in the level of integration when compared to the GBT correlator chip (twice the number of transistors). By using a more

modern process, with finer component features and low voltage technology, a smaller chip with lower power requirements should be possible. The smaller silicon size should also mean a higher yield in the manufacturing process.

10.4 Second Generation Future Correlator

The ALMA Baseline Correlator described above will provide a correlator which meets the baseline science requirements. However, it is clear that advances in semiconductor technology and correlator architecture will eventually make that correlator obsolete, and that planning for a future correlator is necessary.

10.4.1 Introduction and Correlator Nomenclature

While the NRAO team designs the ALMA Baseline Correlator with delivery of the first quadrant scheduled by mid-2004, European teams in the ALMA project are working on a preliminary design study and sub-systems prototyping for a second generation Future Correlator. The goal is to offer higher correlation efficiency, more spectral channels and perhaps more flexibility than with the Baseline Correlator while benefiting from the technical advances one can anticipate before completion of the full ALMA Correlator. At the same time, the Japanese are working on an FX architecture design (c.f. ALMA Memo 342 at http://www.alma.nrao.edu/memos/html-memos/abstracts/abs342.html.) The current nomenclature is then, Baseline Correlator, Future Correlator and FX Correlator for the baseline ALMA project and the European and Japanese team projects, respectively.

With Japan joining the ALMA project the Enhanced ALMA will require an Enhanced Correlator or a change to the baseline correlator design to accommodate any additional 12-m antennas beyond the 64 served by the current baseline correlator design. The 10 to 16 smaller antennas of the ALMA Compact Array may be served along with the 12-m antennas by a single correlator, or by a separate, dedicated Compact Array correlator. This Enhanced Correlator should be defined by the joint efforts of the NRAO-European-Japanese correlator teams.

We concentrate here on the main features of the European Future Correlator. Europe proposes to prototype and build a digital hybrid correlator whose architecture is somewhat intermediate between the XF and FX architecture designs and incorporates features common to both designs. The basic concept was first discussed in a European kick-off meeting held in April 2000. Further presentation was made in September 2000 during the ASAC meeting and during an informal meeting held with the Japanese team. (See minutes of the latter meeting in http://www.eso.org/projects/alma/committees/backwg/minutes10sept_correl.doc.)

The European Enhanced Correlator concept and the WIDAR correlator concept proposed by the Canadian group for the VLA are closely related.

10.4.2 Future Correlator Overview

One key idea behind the digital hybrid correlator concept is that for a given spectral resolution the cross-correlation requirements diminish with the number of sub-bands sub-dividing each digitized 2 GHz input band. The current concept considers 16 partially overlapping sub-bands in order to match a correlator clock frequency of order 250 to 300 MHz. In this frequency-division

demultiplexing scheme there are several important issues under investigation: flexible interconnection of correlator cells, digital filtering and requantization, digital total power calibration for sub-band concatenation and FFT, etc. A simple block diagram is shown in Fig. 10.2. 3- or 4- bit correlation is a major goal of the Future Correlator study. The advantage is significant because it is beneficial to all types of observations compared to 2-bit correlation. Going from full 2-bit to full 3-bit operation diminishes the quantization losses and is equivalent to adding about 10% more collecting area.

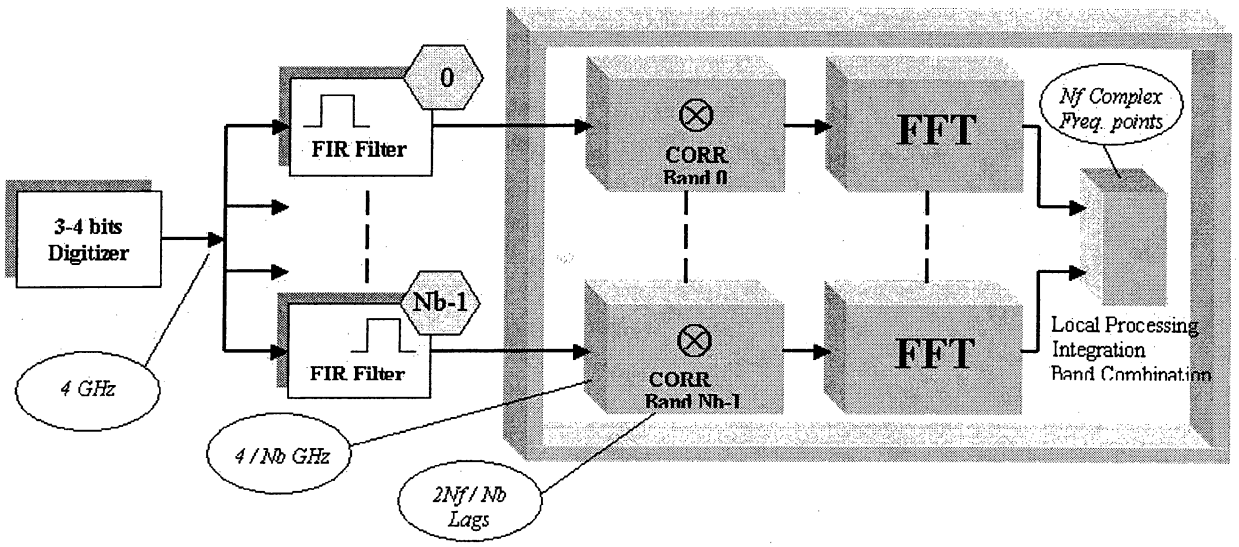


Figure 2. Future Correlator simplified block diagram

The number of sub-bands per 2 GHz total band is an a priori free parameter in the digital hybrid correlator design. This parameter has an impact on both the total cost (more FIRs means lower clock rates) and power consumption (depending not only on the adopted correlator chip design but also on the performances of the FIR filter ASIC design).

In contrast with the Baseline Correlator in which the FIR filters are used to narrow the input bandwidth, in the Future Correlator FIRs are used for both band narrowing and frequency demultiplexing of the 2 GHz input band. It is interesting to note that FIR filtering is similar to the FFT stage in the FX correlator design. In the Future Correlator the total number of spectral channels is driven by several issues and the number of lags per baseline associated with each 2 GHz sub-band is not yet adopted. The maximum spectral resolution depends on flexible allocation of lag resources using cross-bar switches to interconnect the correlator cells. The present goals are 8000 spectral channels across 8 GHz bandwidth and about 5 kHz maximum

resolution.

An initial document describing the signal routing and the hybrid system architecture based on generic correlator modules, cell partitioning and interconnections of an NxN array of cells fed by N correlator input signals has been prepared by A. Bos for discussion in the European team. Other on-going tasks include fast backplane technology studies, signal processing implementation platforms and VHDL simulations, and design of FIR filter ASICs operating above 100 MHz. In addition, there is good progress in the construction of an end-to-end model to simulate the Future Correlator signal flow and the effects of requantizing after the FIRs. An operational mock-up comprising both hardware and software sub-systems will be prepared in 2001 to test the Future Correlator concept and identify the most critical components/sub-systems. Results will be given in the ALMA Phase 1 Design Study document.

ALMA Construction Project Book, Chapter 11

HOLOGRAPHY

Antonio Perfetto & Darrel Emerson
Last revised 2000-Sep-28

Revision History
2000-09-28: First draft version

Summary

Holography, using a terrestrial transmitter, will be used as the prime measurement tool for antenna adjustment. The system has already been developed for the Test Interferometer antenna measurements at the VLA site in New Mexico. Essentially the same measurement hardware will be used for the initial antenna adjustments as antennas are assembled at the Operations Center in San Petro, prior to being transported to the high site at Chajnantor.

Measurement System

The holography hardware and software used for antenna evaluation at the VLA site will also be used in Chile. This is described in detail in the Chapter 5 (Holography) of the ALMA Test Interferometer Project Book. The holography transmitter is mounted on a tower, some 50 meters high, and about 300 meters from the antenna. The corrections, to be applied in the reduction software, for the transmitter being in the near field of the transmitter are very important. Dual frequencies, of ~80 GHz and ~104 GHz are used. The dual frequency measurements give extra information that will confirm many aspects of the near field correction, and also of other corrections applied to allow for various diffraction artifacts. Please see the above reference for details.

Operational Considerations

The antenna manufacturer will be responsible for the initial antenna adjustment, to a precision of around 100 microns rms surface error. The specification for the final antenna surface precision is 25 microns, with a goal of 20 microns. The holography measurement system has a specified error of 10 microns rms, with a goal of 5 microns.

In practice, the antennas at the San Pedro center will be handed over to NRAO with a 100 micron surface precision, and then holography using the transmitter on the 50-meter tower will be used to measure the surface to a precision of ~10 microns. The measurement requires that the antenna control system and data acquistion system will support the holography observing mode. For each antenna, it is anticipated that two or three iterations of measurement, adjustment, and measurement will be required in order to reach the necessary antenna surface precision. It is assumed that transportation to the high Chajnantor site can be made in a way not detrimental to the surface precision. In the current operational plan, there is no intention to perform further single-dish holography at the high site using the dedicated holography receiver and transmitter. Interferometric holography, using celestial sources, will be carried out at the

high site, using standard ALMA receivers, as soon as the system in place there permits.

---

12 Computing

Revision History:

2000-09-29: Complete rewrite.
2000-10-06: Correct some typos and phrasing.
2000-11-22: Update data rate requirements, ZASI description.
2001-01-22: Reflect new AMBSI names, replace ATM with Gb Ethernet.

12.1 Introduction

B. Glendenning (NRAO)
Last Revised: 2000-11-22

This chapter describes the operational software required for ALMA. This includes real-time and near-real-time software to monitor and control hardware devices, software to schedule the array, software to format the data suitably for post-processing, software to archive and restore the data, software to perform fundamental calibrations (*e.g.*, pointing) required to operate the array, commissioning software (*e.g.*, holography), and software to implement a near-real-time imagine pipeline. It does not generally include post-processing software, firmware that is "inside" the device (excepting the correlator), or engineering test software that is not needed during operations.

ALMA Computing, principal requirements

Sustained data rate, science data	6 MB/s (Average) 60 MB/s (Peak sustained)
Image pipeline	First-look images produced automatically for standard observing.
Dynamic scheduling	Nearly automatic scheduling of the array, accounting for current weather and other conditions, to optimize the scientific throughput of the array.
Archiving	Networked archive of all ALMA raw science data and associated calibration data and derived data products.

12.2 Science Software Requirements

R. Lucas (IRAM)
Last Revised: 2000-11-22

The scientific requirements for ALMA software have been described in ALMA memo 293 rev.1), from which we reproduce here the introductory remarks. The reader is encouraged to read this report for details. A new version of this report is being prepared, which will include more formal requirements and execution scenarios (Use Cases).

The operation of ALMA will have to deal with a larger variety of projects than previous instruments: on one hand at long wavelengths (1-3 mm) due to the high sensitivity and quality of the site, and a long experience with millimeter-wave interferometry, we can predict with reasonable certainty the observing modes that will be used, the relevant observing strategies to schedule the instrument, and the data reduction techniques. On the other hand at the highest frequencies (~300μm) no array has been operational yet; we plan to rely on techniques such as radiometric phase correction, fast phase switching and phase transfer between frequency bands, that have been demonstrated, but not applied with the operational scale that we foresee for ALMA. We thus will have to combine in the software a high level of automation, needed to deal with the large information rate that will be available, with a high level of flexibility at all levels to be able to develop and implement new observing methods and reduction procedures. For simple projects, the astronomer with little or no experience of radio techniques should be able to use the instrument and obtain good quality results; however, experts should easily be able to perform experiments we do not even foresee today.

The expert user/developer will need to be able to send direct commands to the instrument through a simple, easily editable command language. Atomic commands in a script language will directly send orders to the basic software elements controlling the hardware: antenna motion, instrument setup, or transmitting parameters to the data processing (pipeline). The script language will support loops, structured conditional tests, parameterized procedures, global variables and arrays ... These scripts, once fully developed and tested, will evolve into the basic observing procedures of the instrument.

The general user will need more user-friendly graphic interfaces to many components of the system. They will propose several templates, corresponding to the available observing modes, and provide a simple way to pass astronomy parameters to the basic observing process and to the corresponding data reduction procedures of the pipeline. Input parameters will preferably be expressed in terms of astronomical quantities, which will be translated into technical parameters by sophisticated configuration tools.

Proposal submission will be in two phases, the first before proposal evaluation, the second to provide information needed for the actual scheduling and observation.

We believe that dynamic scheduling is an essential feature of the instrument and should be installed from the very beginning of its operational life. Though the site is undoubtedly one of the best for submillimeter observations, it will only be usable at the highest frequencies for a fraction of the time; to improve the total efficiency we must be able to make the best use of all weather conditions, by selecting in quasi-real time the project most suited to the current weather and to the state of the array. This means we should always be able to observe a given project in appropriate weather conditions. This philosophy can be extended to the point where a given project can change its own observing parameters according to variations in observing conditions (such as atmospheric phase rms).

The whole real-time system will be under control of a telescope operator, through a specially designed interface. This must provide an overview of what observation is occurring, the state of the instrument, and observing conditions on the site, and should enable the operator to react to any unexpected event. A general monitoring interface must be also accessible through the network.

The instrument should produce images, aiming to be final for most projects, even when projects are spread over several sessions and configurations, and/or include short/zero spacings. For this purpose an on-line pipeline is required. It will include calibration of the array itself, to reduce measurements of baseline, delay offsets, and determine pointing models during specific sessions. During standard observing sessions reference pointing and focusing measurements will have to be

reduced, with fast loopback of results to the observing process; the phase fluctuations on the phase calibrators must be evaluated, with a feedback to both the real time process and the scheduler. Calibration will be applied on-line and maps/datacubes will be produced according to data processing parameters input by the observer. Single-dish observing sessions will also be reduced on-line. The pipeline must be able to reduce on line the quasi totality of the data, which is expected to be produced at an average rate of 6MB/s, with a peak rate of 60MB/s for some observing sessions.

For most projects the data pipeline will produce results in a form suitable for quality evaluation, and astronomical processing, hopefully leading to fast publication. Uncalibrated UV data will be archived together with the calibrations curves and the resulting images. The archive should enable fast access to the observing parameters and full reprocessing of the data set with improved processing algorithms.

A general requirement is that the various parts of the system should be developed in a highly consistent way, from the very beginning of design; they may however be installed progressively, provided the critical elements are implemented first.

12.3 Architectural Overview

12.3.1 Physical

B.E. Glendenning (NRAO), G. Raffi (ESO)
Last Changed: 2001-01-22

As shown in Fig.1, the ALMA computers and networks are organized around a few geographical areas:

- Antenna computers, which are repeated and essentially the same at the 64 antennas
- Central computers at the ALMA site, with central control and coordination functions linked to the whole array.
- Operation Support Facility (OSF) computers to offer user and maintenance interfaces to the whole array from San Pedro.
- Data centers in Europe and USA with archiving and remote access functions
- Other computers accessing ALMA via Internet. Although these should be seen as external to the ALMA project, their presence and functions might have to be supported via certain categories of software and so they have to be considered.

It is important to note that operation at OSF is considered as local operation, i.e. all normal and maintenance operations are available at the OSF with full functionality and performance. The only reason to perform maintenance at the ALMA site has then to be the need for physical access, replacement or movement of parts, but not any limitation in the capabilities offered at the OSF. These should include video and audio links to the central computers at the ALMA sites and possibly to the antennas.

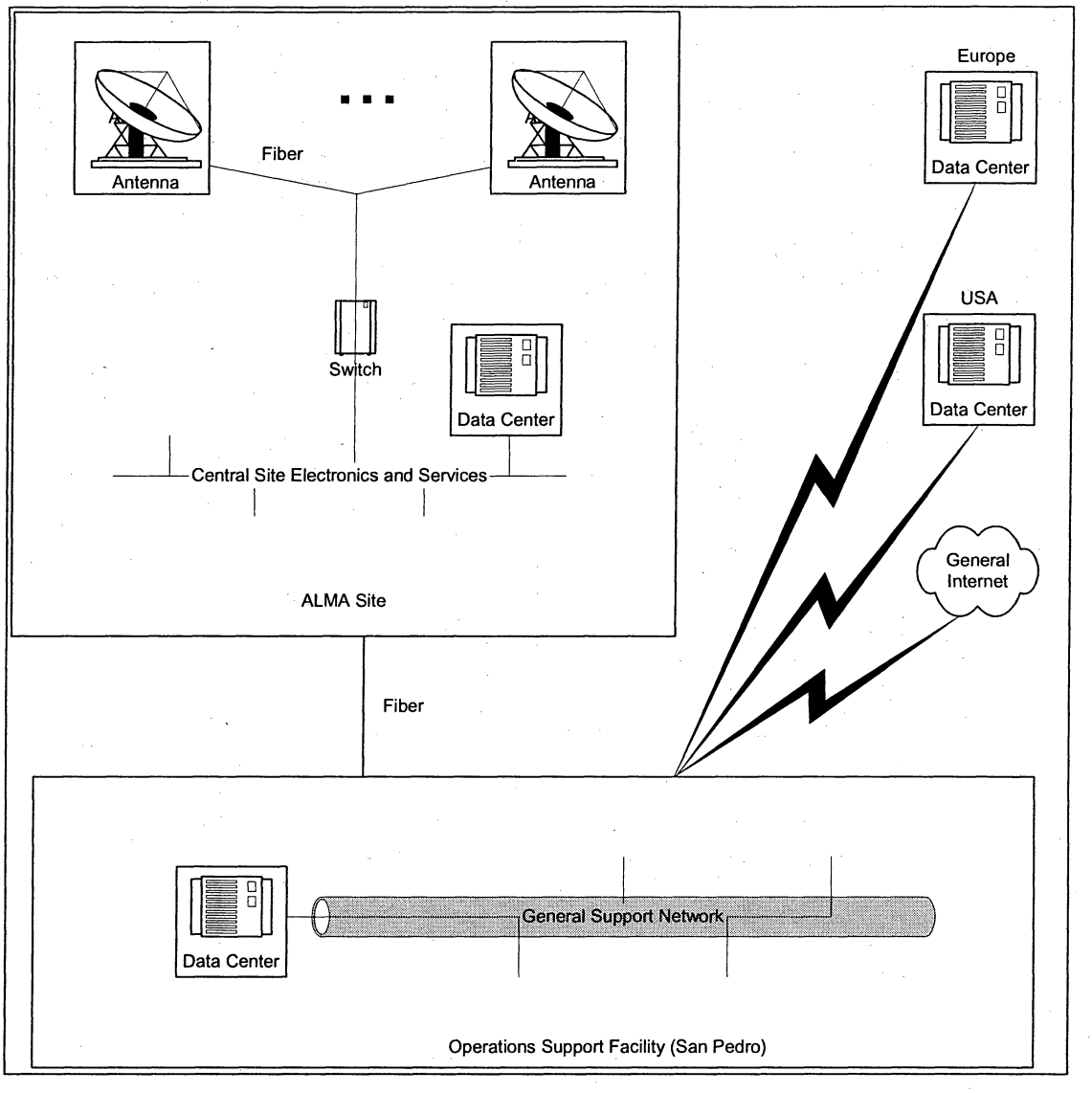


Figure 1 - Overview of ALMA Computer Systems and Networks

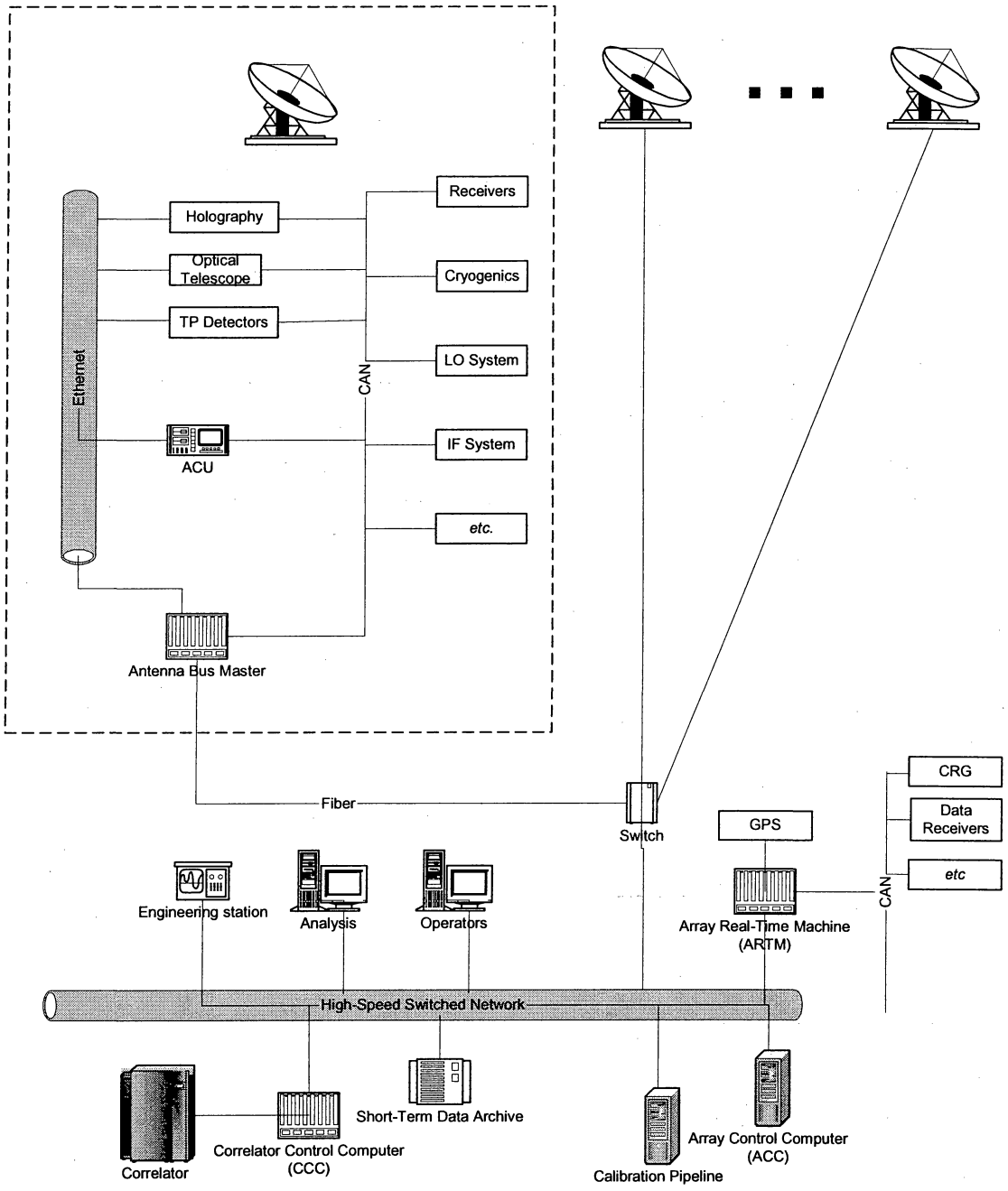


Figure 2 - Layout of computers at the site

At each antenna there is an Antenna Bus Master (ABM) – a VME Power PC based VxWorks computer. Its principle role is to provide real-time control of the devices at the antenna based upon infrequent time-tagged commands from the center. It serves also as a router for the antenna Ethernet segment.

All devices with computer interfaces are attached to a Controller Area Network (CAN) bus, through which they are controlled and monitored. More details about the properties of these interfaces are described under Monitoring and Control. A particularly important device is the

Antenna Control Unit (ACU) that is provided by the antenna vendor. It is implemented as a VME/PPC/VxWorks computer. While like any other device it is controlled over the CAN, it also has an Ethernet interface for software maintenance and the setting of static parameters that do not need to be changed in normal operations.

It is assumed that data transmission from the total power detectors and holography backend is done over Ethernet.

The optical telescope (for the antennas where this will be mounted) is also commanded via the CAN bus. Its video output will be digitized and transmitted over Ethernet and via the ABM to the central computers.

Each ABM is connected to the central systems via a point-to-point Gigabit Ethernet network that terminates at a switch. The switch is in turn connected to a high-speed switched network on which all central ALMA computer systems required to operate the array are attached.

At the center there are two real-time computers. The Array Real-Time Machine (ARTM) plays the role of the ABM, providing local real-time control of its attached CAN devices. The Correlator Control Computer (CCC) is the other central real-time computer. It provides the interface for the correlator, and provides detailed control of the correlator hardware. Both the ARTM and CCC are VME/PPC/VxWorks based systems.

The coordination function is implemented via the Array Control Computer (ACC), which is a high-end workstation running the Linux operating system. It is responsible for controlling all hardware in the array (indirectly through the ABM, ARTM, and CCC computers) under the command of a high-level observing script. It also runs various ancillary software such as model (*e.g.,* delay) servers, data formatting, *etc.* If necessary for performance reasons the ACC functions could readily be split into multiple computers.

There will be a few general purpose Linux systems on the switched network for operator access, astronomical data analysis, software development, and the like. Computers with Labview will be used for engineering and maintenance purposes.

Data from a meteorological data station will also be available at the OSF.

12.3.2 Logical

G. Harris (NRAO)
Last Changed: 2000-09-28

The ALMA Control Software System has a number of concepts, among them are session, sequence and data products. The Conceptual Diagram portrays some of these important notions.

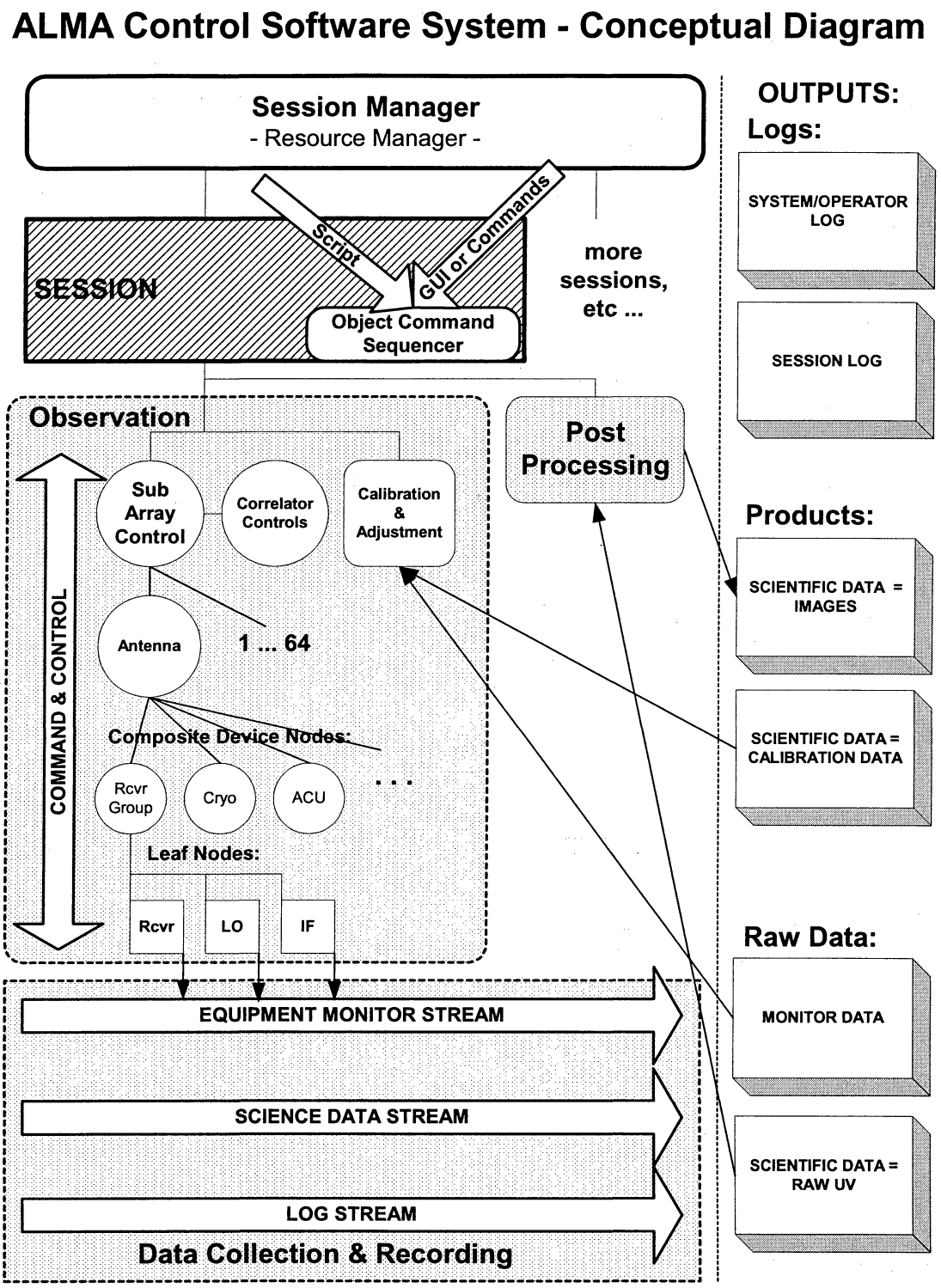


Figure 3 - ALMA Control Software System - Conceptual Diagram

First of all, there is a session manager or executive, described below, which creates observer work sessions and allocates resources such as antennas, correlators and post processors to them.

This observer session will perform observations and perform other processing. It is directed by a programmatic script created in advance by the observer, or entered directly in interactive mode through a GUI or commands. These commands will expand through the action of the software system into a sequence of device commands.

While operating an observation, for example, control commands will be issued to a software structure as shown in the shaded box labeled *Observation*. This structure operates the sub-array of antennas and correlator as a unit, issuing device operations in a timed sequence. Then later, the same session might drive a *Post Processing* operation to process the raw data placed in the archive by the *Observation* activity and make it into finished images.

The *Observation* box shows some of its major activities by component. First, the commands flow down the structure of sub-array and antennas in a hierarchical manner. The further hierarchical structure of device groups and devices, shown in the composite and leaf device nodes, leads us to name this structure a control tree, as shown in later drawings. Each tree has groups of devices, such the receiver group with receivers, filters, LOs and IFs. Ultimately, the software tree reaches down to leaf nodes connecting to hardware.

As the observer operates the instrument, the observer's session will create *Outputs* as shown in the diagram. There will be logs, both of individual sessions, and of the system as a whole. There will be raw output from an observation plus processed images and calibrations. Some of these activities may be in parallel, allowing feedback into observation parameters.

The system also needs to operate independently as observer sessions come and go, keeping logs and equipment status, performing maintenance and calibration and other activities. Supporting this is *Data Collection and Recording*, which operates at the system level. This is shown by another shaded box in the diagram. Data collection from the equipment is not dependent on a control session, nor vice versa.

The collection process streams information from the antennas to central storage and makes it available for the system as well as the observer sessions when they are operating. There is equipment information: temperatures, frequencies, etc.. There is science data processed by the correlator. There is also log information, constantly from the system and operator, and periodically from observer work sessions. There is a method, described below, to subscribe to the information published by the data collection process.

Now moving from the general concepts of session, sequence and data products, look at the details of session operation as shown in the figure for Logical Architecture.

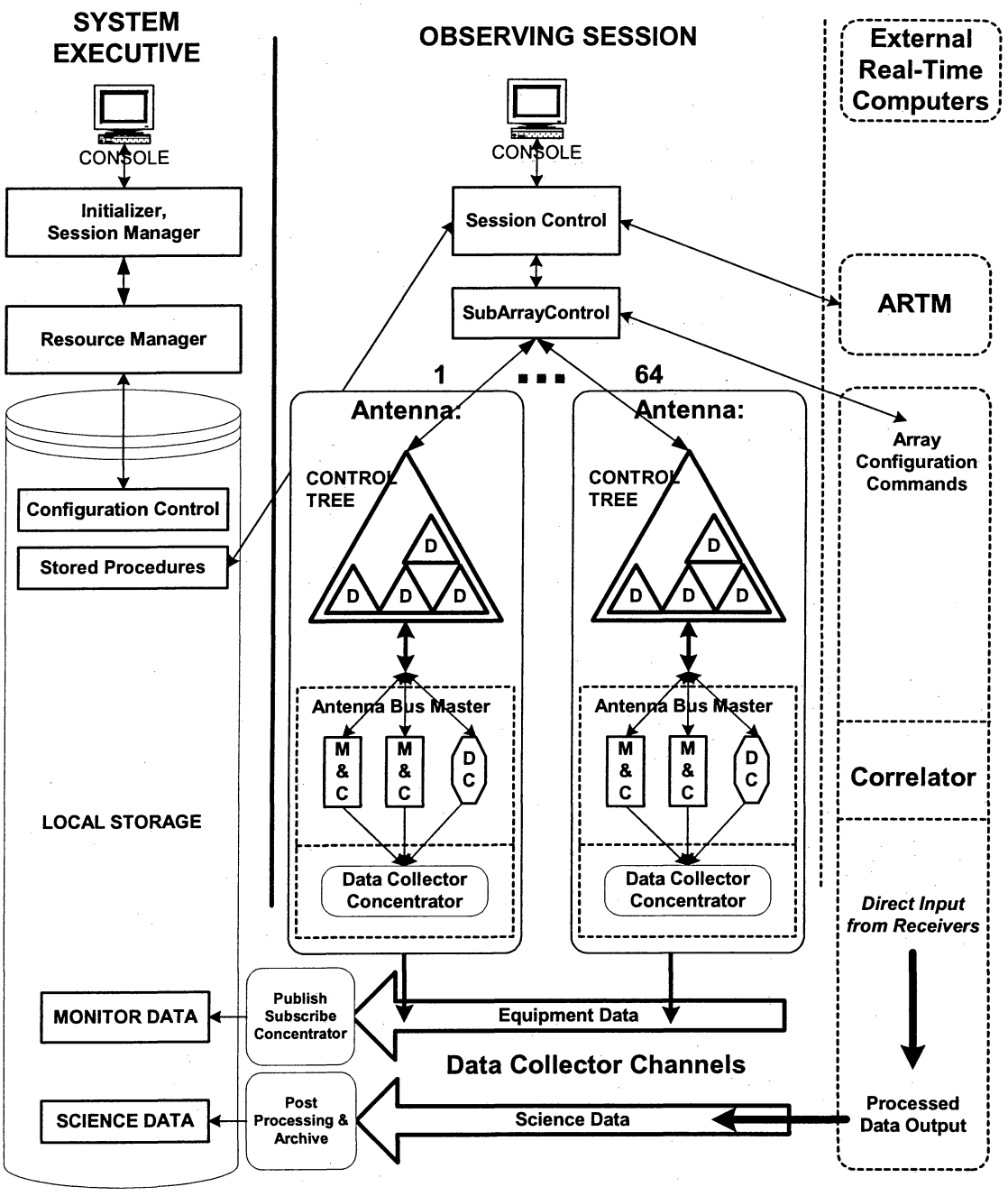


Figure 4 - ALMA Control Software System Logical Architecture

First of all, ALMA has a supervisory program called the system executive. The executive initializes the instrument, creates processes for observers called work or observing sessions, allocates resources, and launches services which collect and process data. It also supports operator activities.

The action is similar to initializing an operating system:

- During startup, a resource manager locates and initializes the devices, sometimes using information from a configuration control database.
- The system processes are started and necessary communications connections established, making devices available to processes in the system. This is especially important for services such as object component naming, which supports various parts communicating with each other.
- The initial allocation of these resources is performed, to a resource pool or to an observer.
- Then observer processes are started to use the resources and perform work. These are typically called sessions.

There may be many observer sessions, operating simultaneously. There may also be additional operator work sessions, as the executive will also provide for some maintenance and configuration activity.

Using the logical architecture diagram, we can see how this will operate. The operator will start the executive after operating system startup.

The executive will first start the resource manager. The resource manager will be responsible for using the configuration database to initialize the antennas. Each antenna will be completely initialized, including the control tree and Antenna Bus Master [ABM], and allocated as a unit to a work session. This process will also register all the objects [devices] in the antenna in the object naming system.

Once these devices can be identified, the executive can also start the data collection process in each antenna's ABM computer. The ABM collector concentrates the equipment information from about a thousand points per antenna. There is also a central collector, which gathers the data from each antenna. [There will be 64 antennas.] This is called a publish/subscribe mechanism. Once the data is collected into the central Array Control Computer [ACC], other processes may subscribe from this central publisher.

With all antennas initialized and the data collector operating, the executive can start work sessions for the observers and give them appropriate resources depending on the desired observing mode. When a work session terminates, the executive will clean up the resources and possibly re-initialize them or place them in a known state for another session. Operator sessions or direct executive commands may be used for various standard activities, parking or stowing, shutdown, special calibrations etc.

[Now look in the drawing at the section labeled SESSION.]

Using the work or observing session, the observer enters commands or retrieves stored commands on the system if in interactive mode. If in automatic mode, prepared observing scripts are executed in the order presented by the system scheduler. The observing commands are propagated into the SubArrayControl and then to each antenna through a control tree, a software construct representing the devices on the antenna.

The method by which this is done is called CORBA [Common Object Request Broker Architecture], an industry standard. Use of distributed object oriented technology is quite extensive in this instrument. Every possible piece of the software is object based, from the configuration data, to the scripts and their functions, to the devices in the control tree and even to the hardware device control points on the antennas. This facilitates the creation, management and execution of the control software.

Some devices are simple and others are treated as a group. The letter D inside the control tree triangle represents a Device, perhaps a composite hierarchy of receiver parts such as LO, filters,

etc. The hierarchy of devices, represented by the triangle, expands commands as they propagate through the control tree. Sometimes a simple command by the observer results in many commands to the bottom level devices.

Some software devices are in the central computers, in what may be called the logical time part of the system. Other devices are in external computers such as the correlator, the Array Real Time Machine [ARTM - the time server] or even out on the antennas, next to their associated physical devices in what we call the ABM [Antenna Bus Master]. **These external computers are represented by dotted lines and operate in real time.** The dotted line between the control tree of an antenna and its ABM out on the antenna represents the transition from logical time to real time.

Once commands reach the ABM, they go to appropriate Monitor and Control points [M&C]. Monitor points read the value of a single hardware access point and Control points can set the value of a single hardware access point.

Sometimes a behavior is needed in real time which physical devices do not have. Rather than trying to perform this more complex behavior remotely from the control tree, a software construct called a device controller [DC] is used. For example, a nutator sequence may be given as a single command from the command tree resulting in many actions by a device controller in the ABM, all synchronized in time.

Another feature of this architecture is the decoupling of the control system from the data recording. The control system is not dependent on the data monitoring and recording system. Or vice versa. Equipment may be monitored even when not being used for observations. Just as the antennas are configured and initialized before giving them to a work session, so also is the data collecting system turned on and initialized by the system executive.

The correlator is accessed on two levels. First it is configured on an array basis and commanded as shown in the drawing by the SubArrayControl at the subarray level As data coming from the receivers is sent to the correlator, it is processed and the correlator output goes into the science data.

Another independent computer with support electronics is the Array Real Time Machine [ARTM]. It coordinates with external clock sources, internal frequency generators and other equipment to provide a timing standard to the array.

When operating in interferometric mode, a single session drives many antennas constituting a sub-array. The SubArrayControl distributes the current command from session control to the antennas, correlator, etc.. Generally all antennas receive the same command, but each can also be individually addressed. Any per antenna differences in processing a command are handled by each antenna's control tree and ABM.

Other activities not shown in this diagram for simplicity are:

- Display sessions using the information from the central data publisher,
- Observing proposal preparation,
- Observation Scheduling and Dispatch,
- Image Processing Pipeline, and the
- Feedback from monitor data, and Image Processing to observations
- Archiving and distribution of data.

12.4 Data Flow Software

J. Schwarz (ESO)
Last Changed: 2000-11-22

The Data Flow Software will encompass the end-to-end processing of an ALMA Observation, from the preparation of a proposal by a prospective observer, through to production and archiving of cleaned images/data cubes from the array. Support will thus be provided to:

- The proposer/observer;
- The proposal reviewers and those responsible for granting ALMA observing time;
- The array operations staff;
- The archival researcher.

The Data Flow Software will provide:

- Tools for the preparation of observing proposals; these tools will allow the proposer to specify his/her scientific requirements insofar as possible. The tools themselves will then be responsible for translating these requirements into array-level configurations and observing scripts.
- A dynamic scheduler to maximize the observing efficiency of ALMA in service mode. Observations will be scheduled based on assigned scientific priority, as well as appropriate match to the existing observing conditions. Some (user-specified) observing parameters may be adjusted to take account of changes in these conditions.
- A near-real-time pipeline, capable of feeding back results to the observing process, and of producing images that should be of final quality in the majority of cases.
- An archive to store *uv*-data, images/data cubes, calibration data, observing programs (including proposals, scripts, and observation parameters).

12.4.1 Observing Proposal & Programme Preparation

In order to ensure that ALMA is accessible to the largest possible segment of the astronomical research community, the SSR has decided that prospective observers should be able to input their scientific desiderata (*e.g.*, angular resolution, spectral resolution, field of view, and signal to noise ratio) and that the software tools provided by the Observatory should evaluate these for consistency and feasibility, and derive the necessary Array parameters--from configuration through correlator setup and selection of calibration sources to total observing time.

At the same time, there will be a mode that allows experts to bypass this feature of the tool entirely, specifying in detail the low-level observational parameters to be used. The SSR envisages that a scripting language will enable direct control of the Array. Moreover, modification of accepted programmes is to be allowed up to the time that a Schedule Block begins execution.

How Observatory Operations will coordinate these two types of programmes is TBD; in the case of VLT Service Observing, only a limited set of templates are available to the Observer, and low-level control of the telescope by, for instance, a user's *ad hoc* script, is not allowed. Observation Blocks, the basic units of VLT observing programmes, are vetted individually by the operations staff. Modification of accepted Phase II Programmes by the observer is only allowed under special circumstances.

12.4.2 Observation Scheduling

Although interactive observing will be supported by ALMA, it is expected that the Array will usually be operated in Service Mode. As in the case of the VLT, the argument here is that ALMA

is so expensive an instrument that every effort should be made to maximize its observing efficiency.

In spite of the existence in the observing community of sophisticated software scheduling tools, scheduling remains as much an art as a science. At the VLT, where the Spike inference engine developed for the Hubble Space Telescope was adopted for long- and medium-term scheduling, the final decision about which Observation Block is next to be executed is determined manually by the Operations Staff. Spike is used primarily as a visualization tool at the VLT. It is quite complex, and as it is written in Lisp, requires that comparative rarity, a Lisp programmer, to maintain it. A commercial version, written in C/C++, is also available, and its purchase might alleviate the maintenance problem. Until this is done at the VLT, automatic scheduling is likely to remain a low priority. The VLT Operations Team has not found the manual scheduling procedure to be an excessive load on their manpower resources so far.

On the other hand, the radio astronomy environment may be more hospitable to dynamic scheduling. Wright (1999) described his dynamic scheduler for BIMA, which selected observations on the basis of weather conditions, as a technical success. The algorithm for ALMA, which will include the possibility of modifying some observational parameters (*e.g.*, dwell time on phase calibrators) in response to changes in atmospheric conditions, will be more sophisticated, however. Such decisions will depend upon near-real-time feedback from one of the data reduction pipelines (see below).

Although scheduling "what comes next" may be relatively simple, VLT experience demonstrates that what seems like tonight's optimal solution may cause problems (*e.g.*, a schedule "hole") weeks or months later. A scheduling tool capable of generating possible timelines over longer periods than just one night is likely to be needed. Moreover, no schedule can be optimized over a several-month period unless the awarding of telescope time takes into account the distribution of hour angle ranges and observing conditions that will be available.

12.4.3 Observation Execution & Interface to the Array Control Software

The interface to the Array Control Software is not yet defined and will be discussed in a future revision of this document.

12.4.4 Near-real-time Pipeline Processing and Feedback to Operations

There is a requirement to feed back the results of *e.g.*, pointing and phase calibrations to Observatory Operations in near real-time. Some form of quick-look imaging will also be needed, though definitive image production will probably need to wait for the results of calibrations that may not be available simultaneously with the science data.

Design and performance of the pipeline will be discussed in a future revision of this document.

12.4.5 Archiving and Final Data Reduction

The ALMA Archive must serve two purposes:

1. To support observatory operations and to provide the observer with the data he/she wants and needs in a timely fashion.
2. To support archival research by users *other* than the original observer.

Whether this will require two archives, a "short-term" and a "long-term" one remains to be seen. The SSR has proposed allowing each observer to decide which data he/she wants archived, subject to a limit determined by the correlator data rate (over the long term this is estimated to be 6 Mbytes/sec). If both corrected and uncorrected *uv*-data are to be retained, the effective rate goes

to 6 Mbytes/sec, implying that of order 180 Terabytes must be archived each year. In any case, an archive based on individual observers' decisions is likely to be very inhomogeneous, and thus may not be suitable as the official ALMA long-term archive.

In addition to *uv*-data, the archive will include images produced by the automatic pipeline reduction system, as well as complete information concerning the conditions under which the observations were performed and the data reduced. On the assumption that ALMA will have adequate computing resources for the task, it will be possible for the archive researcher to request reprocessing of the archived *uv*-data with improved calibration data and tailored algorithms.

12.4.6 Acknowledgments

Dave Silva, Bruno Leibundgut, Michele Peron and Peter Quinn of ESO's Data Management and Operations Division have been generous of their time in discussing the VLT Data Flow experience with me.

12.4.7 References

R. Lucas, B. Clark, J. Mangum, P.Schilke, S. Scott, F. Viallefond, M. Wright, "ALMA Software Science Requirements: Version 1.0 Report", ALMA Memo #293 (Revised), 2000.

M.C.H. Wright, "Dynamic Scheduling", ALMA Memo #282, 1999.

12.5 Control Software

12.5.1 Overview

B. Glendenning (NRAO)
Last Updated: 2001-01-22

This section summarizes the design for the ALMA Array control software. At present, it is very closely modeled on the test interferometer design. As we gain experience with the test interferometer, the designs will separate (for example, the sub-array facilities for the ALMA array will be very much more elaborate).

In general, we would describe this design as conventional in the sense that designs like it have been implemented for many recent telescope control systems. There are no drivers that we know of that would require us to come up with a radically new design.

Fundamentally, the role of the control software is to take high-level observing commands from the scheduling system and to turn these high-level commands into detailed control of the array, finally producing science data products and monitor data archives. In brief, the scope of the control software is "scheduler to data."

Almost all devices will be attached to a CAN bus operating in a master/slave (polled) fashion. The bus will operate at 1Mbps and is capable of at least 2000 polled operations per second (up to 8 bytes of data per transaction). Devices on the CAN bus will be responsible for implementing a simple in-house protocol to map CAN message IDs to internal device addresses. A few devices will have other connections, in particular Ethernet.

The CAN buses are attached to (mastered by) a Power PC (PPC) VME based computer running the VxWorks real-time operating system. There is one ALMA PPC system at each antenna, and two at the center. One central system is used to master the central CAN devices. The other system is embedded within the correlator. Centrally there is a general-purpose computer running Linux

that is the overall master for the system, which is known as the Array Control Computer (ACC). There are also some ancillary systems in the center for, *e.g.,* operators to sit at.

The antenna-based systems are connected to the central systems via point-to-point Gigabit Ethernet network connections. Commercial off the shelf (COTS) solutions for Gigabit Ethernet are available that support 40km fiber runs and quality of service (QoS) guarantees. The central computers are connected with Gigabit or 100Mb (TBD) switched Ethernet on a network segregated from the rest of the networks site networks.

Logically, the software is partitioned so that control flows in a master-slave fashion from a central executive who controls high-level ("composite") software devices that in turn control their constituent parts. The lowest level software devices are referred to as device controllers, and represent a proxy for the actual hardware – that is, they communicate with the hardware. Data – monitor and backend – is collected from the devices by a collecting process in the real-time computer attached to the hardware, and buffered up for distribution via a publish/subscribe mechanism to consumers of the data, which include the processes that format and archive the data. The software is distributed amongst computers so that only the device controllers and software directly concerned with low-level device activities are on the local real-time computer. All higher-level software entities are concentrated on the ACC.

Engineering access to devices that are installed on the test interferometer will be implemented via access to the device controller interface or to the I/O routines directly from engineering workstations.

The ALMA Time System will establish synchronized switching cycles and mode changes, and provide time-stamping for the resulting measurements. This must be done across the entire instrument including the central building and the geographically dispersed antennas. Additionally, the Time System must be accurately related to external measures of time to correctly determine the position of astronomical objects of interest. The fundamental time system of the interferometer is TAI time maintained in a central master clock. Although TAI is the internal system used by the software, other time systems such as UTC and local sidereal will be accepted from user input routines. Most devices do not have precise timing requirements. For those that do, the control software must arrange to have monitor and control commands sent to the device in precisely defined windows within the pervasive 48ms timing period. This is accomplished by sending time-tagged commands from the center sufficiently in advance of when they are required to account for the non-determinism in the network and general-purpose ACC. The commands are then staged in the ABM until they are required at the hardware. The slave clocks are given the array time of a particular timing event, and they maintain time thereafter by counting timing events.

There are several characteristic ("looping") timescales of interest to the control software:

- 2ms: This is the shortest timescale at which any device will require interaction (the total power detectors). Timescales faster than this are always handled by hardware.
- 48ms: This is the period of the pervasive timing event sent to all hardware with precise timing requirements.
- 16ms: Fastest correlator dump time. (Autocorrelations achieve 1ms dump times but are "bundled" in 16ms increments).
- 1s: This is the fastest timescale for observational changes, e.g., source change or change in correlator setup.
- >1s: Most devices will be monitored or controlled at rates slower than 1Hz, often much slower (300s).

Some model values are required for some devices – for example, the correlator will need a delay model. These calculations will be encapsulated in "model servers" running centrally. The central control software will call these servers and send a parameterized result (e.g., polynomial as a function of time) to the device that needs them.

12.5.2 Device (M&C) Interface

M. Brooks (NRAO)
Last Changed: 2001-01-22

12.5.2.1 CAN Interface

The CAN bus has been selected as the primary interface to distributed hardware devices within the ALMA project. CAN is an ISO standard multi-drop communications medium used extensively in automotive and industrial applications. A higher layer master/slave protocol has been defined (ALMA Computing Memo #7) which governs the behavior of slave nodes at the devices. This bus is hereafter referred to as the ALMA Monitor and Control Bus (AMB).

12.5.2.1.1 Bus Master Nodes

It is required that on any single AMB there be only a single bus master node. Bus master nodes have been coded and tested for the following hardware and software combinations:

- O/S: VxWorks. SBC: MVME2700, MVME1603, MVME2604. CAN Interface: Tews Datentechnik TP816 Dual and Single Port PMC CAN modules (available in the US from SBS Greenspring)
- O/S: Windows NT. Any PC platform with PCI. CAN Interface: National Instruments Dual Port PCI CAN board.

The following hardware and software combinations are planned for support:

- O/S: Linux. Any PC platform with PCI. CAN interface: Janz CAN-PCI/K2O

It is the responsibility of the ALMA Computing Group to design, implement and test all code developed for these bus master platforms.

12.5.2.1.1.1 Bus Slave Nodes

There will be two standard interfaces (SI) to the AMB, the AMBSI1 and the AMBSI2. All designers of hardware interfacing to the AMB should use one of these two; new slave node implementations may be considered where sufficient justification can be provided. All slave node implementations must comply with the bus specification detailed in ALMA Computing Memo #7, which requires that no slave node initiate transactions on the bus unless polled by a bus master.

12.5.2.1.1.2 AMBSI1

The purpose of the AMBSI1 is to provide a highly flexible interface board with on-board CAN and device-side I/O consisting of parallel, serial and bit-wise ports. The board would be delivered to hardware designers with a standard firmware package supporting basic CAN access to the I/O ports and external bus. The micro-controller on the AMBSI1 will have sufficient spare processing capacity to run additional user-specific code, such as providing local displays. The ALMA Computing Group would be responsible for developing such code in conjunction with the

hardware developer and for integration testing the product to ensure that the user code does not interfere with the CAN slave code. All AMBSI1 boards should be similar in hardware, and identical in size and connector arrangement. Code may be loaded into on-board flash memory by means of the CAN bus or the local RS232 port.

The CAN bus servicing code is written to be entirely interrupt-driven. It runs only in response to interrupts from the CAN controller, and those interrupts are set to the highest possible priority (in the Infineon C167, this is interrupt level 15, group level 3). When not servicing an interrupt, the processor executes an idle loop. User-written code may include a function to execute from the main idle loop, or it may simply consist of the callbacks for CAN message reception interrupts.

12.5.2.1.1.3 AMBSI2

This device would perform the function of a CAN bus to Serial Peripheral Interface (SPI) converter. The AMBSI2 is designed for use in systems which either have their own microprocessors or need a very minimum amount of I/O to the CAN bus. It is currently proposed that the resulting subsystem would require +5 Volts at 200mA max and occupy a small daughter PCB of approximately 1.4 by 1.8 in. The PCB would be either mounted to the main PB Board by several SIP headers that also serve as the I/O connectors for the subsystem.

The board would have a unique serial number (programmed into the microprocessor or using the same Dallas part as the AMBSI1, TBD). There would be about 16 bytes of RAM which could be accessed by either the CAN bus or the SPI port. An ATN output would indicate when RAM contents had been altered by the CAN bus. The unit will have its own oscillator and reset circuitry, completely independent of the host microprocessor. All of the software in the subsystem would be developed by and be the responsibility of the software group.

In addition to providing the CAN interface, the unit would also provide the interface for both the RESET and TIMING pulses which come in through the AMB DB-9 connector or module wiring as RS-485 and be available as TTL signals to the application circuitry.

12.5.3 Correlator

J. Pisano (NRAO)
Last Revised: 2000-11-22

12.5.3.1 Summary

This section describes the ALMA correlator control computer hardware and software. This computer system is meant to control the baseline ALMA correlator and process its lags. The real-time computer system consists of two parts. The control computer, a VME PowerPC computer running VxWorks, configures and monitors the ALMA correlator via a CAN interface. The data processing computer, a Beowulf cluster of COTS PCs running Linux with customized device drivers, interfaces to the ALMA correlator LTAs via a FPDP interface and converts raw lag results to spectra.

12.5.3.2 Correlator Computer Hardware Overview

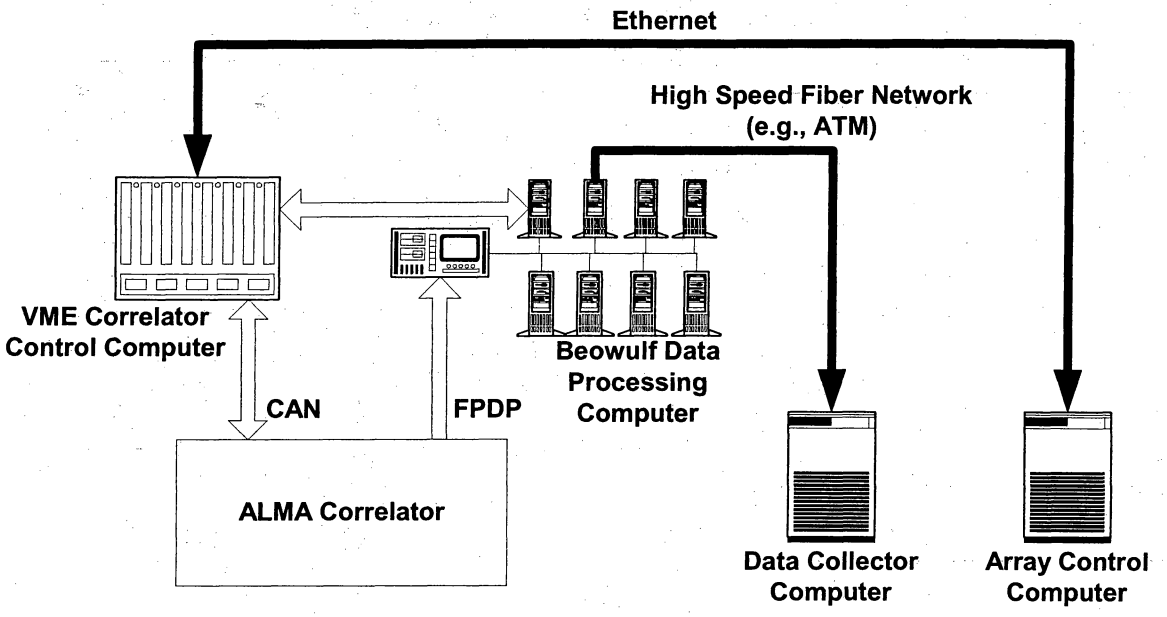


Figure 5

Figure 5 shows a block diagram of the correlator computer relative to other systems to which it interfaces. The Array Control Computer (ACC) acts as a master to the correlator control computer issuing commands to the latter. The Data Collector Computer accepts well-defined spectral data sets, applies further processing (flagging bad data, etc.), writes these processed results to a distribution format and forwards them to an archive.

The Correlator Control Computer (CCC) is a Motorola MV-2700 PowerPC computer which runs the VxWorks real-time operating system. Its primary function is to translate commands from the ACC via an Ethernet connection into a format understood by the ALMA correlator using the CAN protocol. The CCC serves as a CAN master to the slave nodes in the correlator hardware. The kinds of commands that the ACC issues include correlator configuration commands, monitor requests and correlator integration control, i.e., starting and stopping of integrations.

The Correlator Data Processor computer (CDP) accepts raw lags from the ALMA correlator hardware via a 32-bit parallel bus (Front Panel Data Port - FPDP) from the LTAs, converts them to spectral results and transports them to the Data Collector computer via a high-speed link. The CDP must accept lags and perform FFTs at extremely high rates. It has been estimated that the data transfer rates can range from 0.5 MB/second to over 3GB/s and that processing rates can range from 25 MFLOPS (million floating point instructions per second) to 25 GFLOPS.

Currently a straw man design of a Beowulf (see http://www.beowulf.org/ for more information) cluster includes the following items:

- Rack-mounted computer(s) to support 32 FPDPs interface directly to each compute node.
- 16 - 32 Dual processor Pentium III (or equivalent) nodes
- Gigabit Fiber and router(s), e.g., MyraNet, Giganet, InfiniBand to interconnect each node (the Beowulf network) in a multi-drop configuration. A possible option for this is the "Flat Neighborhood Network" topology – see http://aggregate.org/FNN/ for more information.

- High-speed network interface to the Data Collector computer in order to support the burst data rates from the correlator to the Data Collector computer of 60 MB/sec.
- One or more dedicated Beowulf nodes that act as bridges between the Beowulf network and the network.

There exists a link between the CCC and the CDP computer systems that will most likely be Ethernet. This provides a communication link between the ACC and the CDP allowing the CCC to send configuration or timing information to the CDP.

12.5.3.3 Correlator Computer Software

The CCC software will be written in C++ as a multi-tasking application utilizing the VxWorks RTOS as there are many time-critical functions in controlling the ALMA correlator.

The CCC software will have the following properties:

- It will utilize the ALMA Common Software (ACS) libraries for basic software services that include peer-to-peer communications, logging of errors and messages, time services, events and alarm systems all of which incorporate the CORBA architecture.
- It will be aware of the 48 ms array time Timing Events in order to synchronize its operation with other hardware and software components in the array.
- It will be capable of controlling the ALMA correlator on 16-ms time boundaries that are the basic time ticks of the correlator hardware.
- It will be able to support the ALMA correlator's specification to handle up to 16 sub-arrays each having a specific set of antennas, bandwidths, correlation modes and dump times – see ALMA Memo 294, *The ALMA Correlator Long Term Accumulator* for details.
- It will be multi-threaded allowing high priority, time critical operations to proceed with out interruption from lower priority, non-critical operations to proceed without interruption.
- It will utilize the CAN protocol to communicate commands to and to retrieve monitor data from the ALMA correlator.
- It will follow the general characteristics of a "Device Controller" which is a standard ALMA control computing architecture defining how computer-controlled devices operate.

12.5.3.4 Data Processor Computer Software

The CDP software will be written in C++ for Linux. The software architecture will take advantage of the parallelization of the hardware. Each "data spigot" of the LTA will be connected to a Beowulf compute node to handle set of lags for a given baseline. FFTs will then be run on these lags and sent to the Data Collector computer.

The CDP software will have the following properties:

- It will be able to support all of the correlator modes available to the one quadrant of the ALMA baseline correlator.
- It will be capable of supporting the minimum correlator dump times of 16 ms for cross correlations, 16 1-ms auto correlations every 16 ms.
- It will be capable of delivering an output data rate of 6MB/s with a burst rate of 60 MB/sec.

- It will use the FFTW library from MIT (see http://www.fftw.org for details) to compute the FFTs. This is an extremely fast FFT engine especially if there are many FFT computations done on similar data sets. Also there is a version of FFTW which can be run on a parallel computer architecture using MPI, although at this time we do not think that this feature will be necessary.
- It will allow the optional processing capabilities of:
  - Van Vleck correction
  - Hanning Windowing
  - Spectral Decimation
  - Spectral Averaging
  - Apply coarse and fine geometric delays for cross-correlations
  - Sum all channels together to provide a single value.
- It will utilize the standard Message Passing Interface (MPI) – see http://www.mpi-forum.org/ for more information – utilized by many Beowulf systems. It is envisioned that there will be little peer-to-peer communication as each compute node will have all of the lags for a given baseline and will not need to transmit lags between compute nodes. It is likely that the master node will need to communicate information to the compute nodes including configuration and status information and possibly timing information.
- It will have a set of software components that extract lags from the ALMA correlator via the FPDPs and distribute the lags to the appropriate compute nodes. As these data rates are extremely high, this will most likely be a real-time embedded computer with access to the Beowulf network.

12.6 Telescope Calibration

R. Lucas (IRAM)
Last Changed: 2000-09-28

This section outlines the software needs of the operations to be performed in order to determine the calibration parameters that will be needed to successfully execute and reduce standard astronomical observations. These operations will be under the responsibility of the operators and staff astronomers. They may require all the antennas, but in many cases, only a sub array of antennas will be used. Most will actually be performed by observing standard astronomical sources, so many/most of the general science requirements apply; in particular the data will be processed through the pipeline, and the results fed back to the observing processes; the raw data will be archived as science data. These calibration operations will be implemented in the same way as the standard observing modes. A complete list of those operations cannot be made at this time, since the general calibration plan will be subject to evolution during the next years and even the full life of the array.

Among those operations we may cite:

- Pointing calibration sessions:
  A set of pointing calibrations are performed on several pointing calibrators all across the sky. This needs to be done after some antennas have been moved, but also for periodical, systematic checks. Occasionally one should measure the relative pointing of the different frequency receivers. Pipeline data reduction will be the same as for the standard pointing measurements during project observations; then a least square determination of the pointing model parameters is performed for each antenna involved.

- Baseline calibration sessions:
  A set of cross-correlation scans on several calibrators all across the sky are performed, for instance some antennas have been moved (but some antennas which have not been moved have to be included in the sub-array). The data reduction is performed by a least square fit of baseline offsets to the observed phases.
- Delays calibrations:
  They are needed for most interferometric observations. One observes a strong point source calibrator), in the observable sky, for a very short time to measure the relative delays to the antennas by fitting a straight line to the frequency dependence of the observed phases.
- Beam shape calibration
  This involves holography measurements on cosmic sources, to monitor the beam shape and the focusing of the antennas, for instance as a function of elevation.

A later revision to this section may also include the processing of array-wide atmospheric transparency monitoring (FTS?).

12.7 Post-Processing Software

B. Glendenning (NRAO)
Last Changed: 2000-09-28

The AIPS++ package (http://www.aips2.nrao.edu) will be developed as required to cope with general ALMA data processing needs. The ALMA data products will be written in a FITS-based format so that other packages may be used for special-processing or user preference. This section will be expanded in a later revision.

12.8 Common Software

G. Chiozzi (ESO)
Last Changed: 2000-09-28

12.8.1 Overview

The ALMA Common Software (ACS) is located in between the ALMA application software and other basic commercial or open source software on top of the operating systems. It provides basic software services common to the various applications (like antenna control, correlator software, data pipelining).

ACS is designed to offer a clear path for the implementation of applications, with the goal of obtaining implicit conformity to design standards. In a distributed environment like the one of ALMA, the application software will then become more uniform and therefore more maintainable.

The main users of ACS will be the developers of ALMA applications. The generic tools and GUIs provided by ACS to access logs, Configuration Database, active objects and other components of the system will be also used by operators and maintenance staff to perform routine maintenance operations.

It is intended to develop ACS incrementally via periodic releases. Automatic regression tests are planned in order to achieve that ACS becomes a reliable and robust platform.

12.8.2 Technologies

The choice of the technologies used in the ALMA Common Software, and as a consequence on the whole ALMA Software is based on the following initial and explicit decisions:

- Select state of the art but consolidated and widely accepted technologies.
- Adopt an Object Oriented architecture
- Share software rather than re-invent it

The ALMA Common Software is designed based on the experience of the ALMA partners in their previous projects, but is implemented using the technologies and architectural concepts that have found wider industrial acceptance in the last few years.

At the very core is the decision to use CORBA[0][0]. The reasons for using CORBA are in short: Object Orientation, support for distributed systems, platform independence, it is a communication standard, it provides a variety of services.

The ALMA software will have to be as much as possible independent from the operating system and will actually run on multiple platforms (Linux and other flavors of UNIX and VxWorks). We have therefore decided to select the Adaptive Communication Environment (ACE)[0] as the basic multi-platform software. This package provides portable operating system interface services and implements a wide set of classes specifically designed for the implementation of distributed real-time systems.

ACE is also the core of The ACE ORB (TAO)[0], a high performance real-time CORBA implementation.

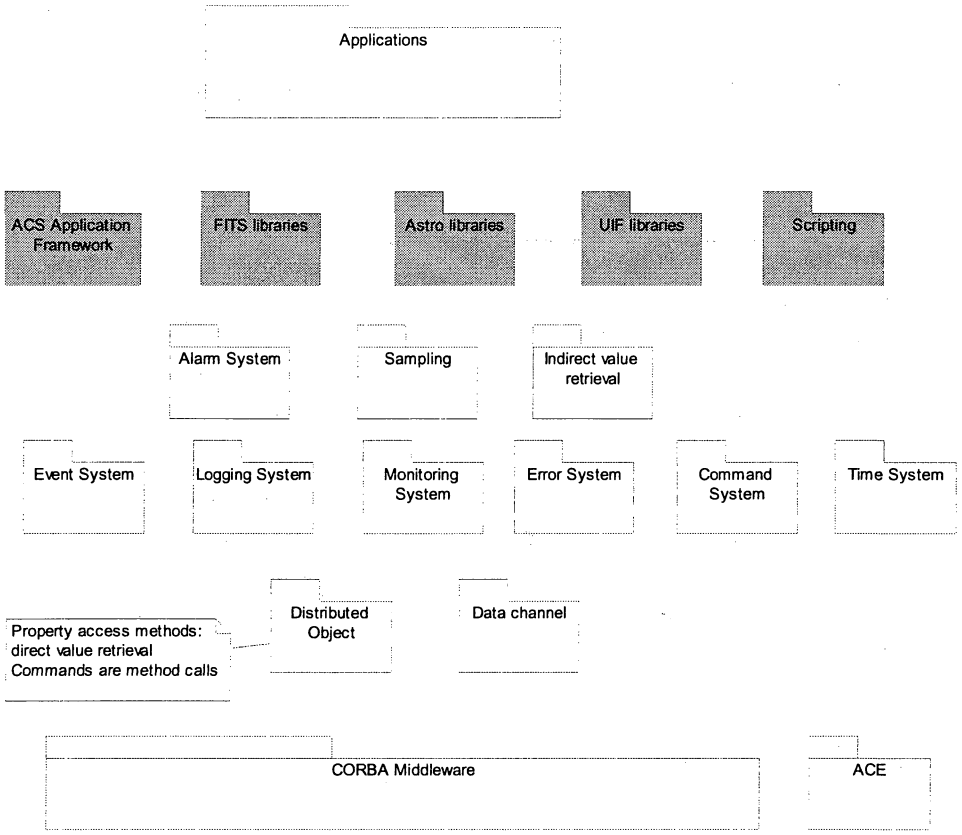
The Object Model for ACS is based on the concept of Distributed Object[0], that identifies three entities:

1. **Distributed Object** - Instances of classes identified at design level in the ALMA system, are implemented as Distributed Objects. In particular, at control system level, Distributed Object is the base class used for the representation of any physical (a temperature sensor, a motor) or logical device in the control system.
2. **Property** - Each Distributed Object is characterized by dynamic values, called Properties, that are monitored and controlled (status, position, velocity, electric current).
3. **Characteristic** - Static data associated with a Distributed Object or with a Property, including meta-data such as description, type and dimensions, and other data such as *units*, *range* or *resolution* are called Characteristics.

The choice of CORBA for the implementation of Distributed Objects and of all services that are part of ACS makes it possible to have every software operation available in a transparent way both locally and at the Control center in San Pedro. This applies also to all data, logs and alarms.

12.8.3 Services

The following UML Package Diagram shows the main services provided by ACS, represented by the packages that implement them.



Packages have been grouped in layers in order to limit as much as possible the relations between the layers (not shown in the diagram): package are allowed to use services provided by other packages on the lower layers and on the same layer, but not on higher layers.

At the very bottom:

- **Distributed Object**: This package provides the IDL interfaces for the Distributed Object class, for the Property classes to access values of all basic data types (like integer and floating point numbers) and for Characteristics. It provides also the basic server-side CORBA implementation for these same classes. Server-side applications can use directly these classes or implement sub-classes for specific I/O devices. Client-side applications use the IDL interfaces to access all Distributed Objects, not being concerned by the specific implementation details of the server.
- **Data Channel**: implementation of a data pipe to transfer efficiently continuous flows of data. It is used by many of the higher level services.

The second layer provides essential services that are necessary for the development of any application:

- **Event System**: implementation of data retrieval by event, monitors and periodic timers.
- **Logging System**: API for logging of data, actions and events. Transport of logs from the producer to the central archive. Tools for browsing logs.

- **Value retrieval System**: Value retrieval and archiving of system and engineering values on a fixed periodic basis or on the occurrence of specific changes in the value.
- **Error System**: API for handling and logging run-time errors, tools for defining error conditions, tools for browsing and analyzing run-time errors.
- **Command System**: Tools for the definition of commands, API for run-time command syntax checking, API and tools for dynamic command invocation.
- **Time System**: Time and synchronization services.

The third layer provides services that are not strictly necessary for the development of prototypes and test applications or that are meant to allow optimization of the performances of the system:

- **Alarm System**: API and tools for configuration of hierarchical alarm conditions based on Fault Tree analysis, API for requesting notification of alarms at the application level, tools for displaying and handling the list of active alarms.
- **Sampling**: low-level engine and high-level tools for fast data sampling (virtual oscilloscope).
- **Indirect value retrieval**: low-level engine and configuration tools for data mirroring and caching.

The fourth and last layer provides high level APIs and tools. More will be added in the future. The main goals for these packages is to offer a clear path for the implementation of applications, with the goal of obtaining implicit conformity to design standards and maintainable software.

12.8.4 References

CORBA - Object Management Group home page (http://www.omg.org/)

Advanced CORBA Programming with C++, M.Henning S.Vinoski, Addison-Wesley, 1999

ACE, Adaptive Communication Environment
(http://www.cs.wustl.edu/~schmidt/ACE.html)

TAO, the ACE ORB (http://www.cs.wustl.edu/~schmidt/TAO.html)

Implementing Distributed Controlled Objects with CORBA -
M.Plesko, PCs and Particle Accelerator Control Workshop, DESY, Hamburg, 1996
(See http://kgb.ijs.si/KGB/articles.htm for this and other related papers).

12.9 Software Practices

B. Glendenning (NRAO)

Last Changed: 2000-09-28

ALMA has adopted a software process based on the Rational Unified Process (RUP). It is using the Unified Modeling Language (UML) as the software design language. It will use the CORBA Interface Definition Language (IDL) as the basis for software/software Interface Control Documents. It uses both configuration control and version control. These items and others will be spelled out in greater detail in a later release of this document.

ALMA Project Book, Chapter 13.1

Imaging Requirements for the ALMA

M.A. Holdaway
Last modified 1998-Jul-13

H.A. Wootten
Last revised 1998-Nov-10
Revised by Al Wootten
Last changed 2000-Apr-18

Revision History:
1998-11-10:Format modified to Project Book Standard

2000-04-18:Format modified to ALMA

Summary

Nonlinear mosaicing--the development of both techniques and algorithms--poses the most pressing set of imaging problems for the Atacama Large Millimeter Array.

Table 13.1 ALMA Imaging Requirements.

Simulation	SDE/AIPS++ and other software packages
Mosaicing	AIPS++ capability of mosmem routine in MIRIAD

Timescale:

- There are no principal goals to be achieved by the end of the ALMA Design and Development (D&D) Phase.
- The imaging capability will be needed by the time the first few antennas are available in Chile, probably by the end of 2005.
- However, some capability of supporting holographic measurements during the test phase at the VLA site will be needed by 2001 Oct 1.

The primary imaging requirement for the ALMA which is not well developed for use at the VLA, IRAM, OVRO, BIMA and other interferometric imaging institutions is the development of nonlinear mosaicing techniques. These are currently fairly primitive within AIPS. Development is currently taking place using the MIRIAD package and BIMA data.

13.1 Simulation Capability

Full simulation capability: given a source structure, particular hour angle tracks (i.e., observing strategy), phase stability, opacity, (add other errors as required), what will the image sensitivity be like? How should one calibrate? What will the image quality be like? Simulation capability should be a tool to aid in the proposal process, and available to the astronomer checking the imaging as well.

Visibility weights should reflect the current noise level (i.e., reflect both changes in T_{sys} and opacity).

13.2 Mosaicing

Mosaicing combines multi field interferometric and total power data. The routines may be based on either maximum entropy (MEM) or CLEAN techniques; mosaics may be constructed using both linear and non-linear algorithms. At present no interferometer works well in total power modes; usually data from an antenna external to the array which is larger in size than the individual array elements provides total power data. When these data are combined, vexing decisions regarding matching the calibration of the flux scales of the two data sets must be faced. Even then, the options for mosaicing within existing software programs are restricted.

To provide the total power data, a set of antennas will synchronously sweep the heavens, providing data without stopping at fiducial sampling points. This is the On The Fly (OTF) observing mode. Data produced in the OTF mode will be combined with interferometer data to produce the dataset to be imaged. Currently, there is no seamless procedure to follow in combination of these data. Better OTF algorithms may be required, as errors in the total power imaging may limit the overall quality of mosaiced interferometer plus total power images.

Each of the methods of image restoration has advantages. CLEAN treats point sources quite well, as it approximates the image by a set of delta functions in one implementation. However, this does not often result in a pleasing portrayal of the millimeter/submillimeter sky, which consists of extended interconnected weak structures. MEM represents these structures well but has not produced results which are a quantitative as CLEAN. A combination of the two algorithms may result in more effective imaging techniques.

Mosaicing--the seamless integration of images taken at adjacent sky positions--is becoming more frequently used as interferometers gain sensitivity at higher frequencies, where their beamsizes are small compared to the scale of heavenly structure. Since enabling this technique within some software packages has required retrofits, the manipulation of multiple pointing data sets can be difficult. We expect this observing mode to be the norm at the ALMA; there must be simple methods of manipulating complex multifield data sets.

For bright or rapidly changing objects, it may be desirable to do interferometry On The Fly. This can challenge the throughput rates of the data system; it has not been attempted on any interferometer yet. The integration times will be set by the minimum allowed by the correlator at the maximum dump rate for the desired number of channels. So, if the correlator is dumping with 0.1 s rates, the ALMA will be slewing at about 3-4 beams per second.

When a fairly bright point source lies within an area to be imaged, small pointing excursions may be tracked over time by self-calibration. For determined antenna pointing offsets which have been determined, then a mean array offset may be measured in this way. This can, in turn, be applied to mosaic or other data to improve the imaging (Holdaway 1993 ALMA Memo 95). Algorithms for application of pointing offsets to mosaicing data sets must be developed and tested.

Over time, each antenna will deform differently, owing to different solar heating, subtle structural differences or other factors. Since each antenna then departs from the idealized model primary beam, the quality of mosaic images will be affected. This problem becomes most severe at the highest frequency. It

may be possible to develop self-calibration algorithms which solve for the varying antenna-to-antenna voltage patterns arising from these deformations. A final image would then be consistent with both the measured data and the deparltures of the individual antenna voltage patterns from the ideal. Such an algorithm would require a computationally intense routine, possibly unusably so.

References

Holdaway, M. 1993 ALMA Memo 95.

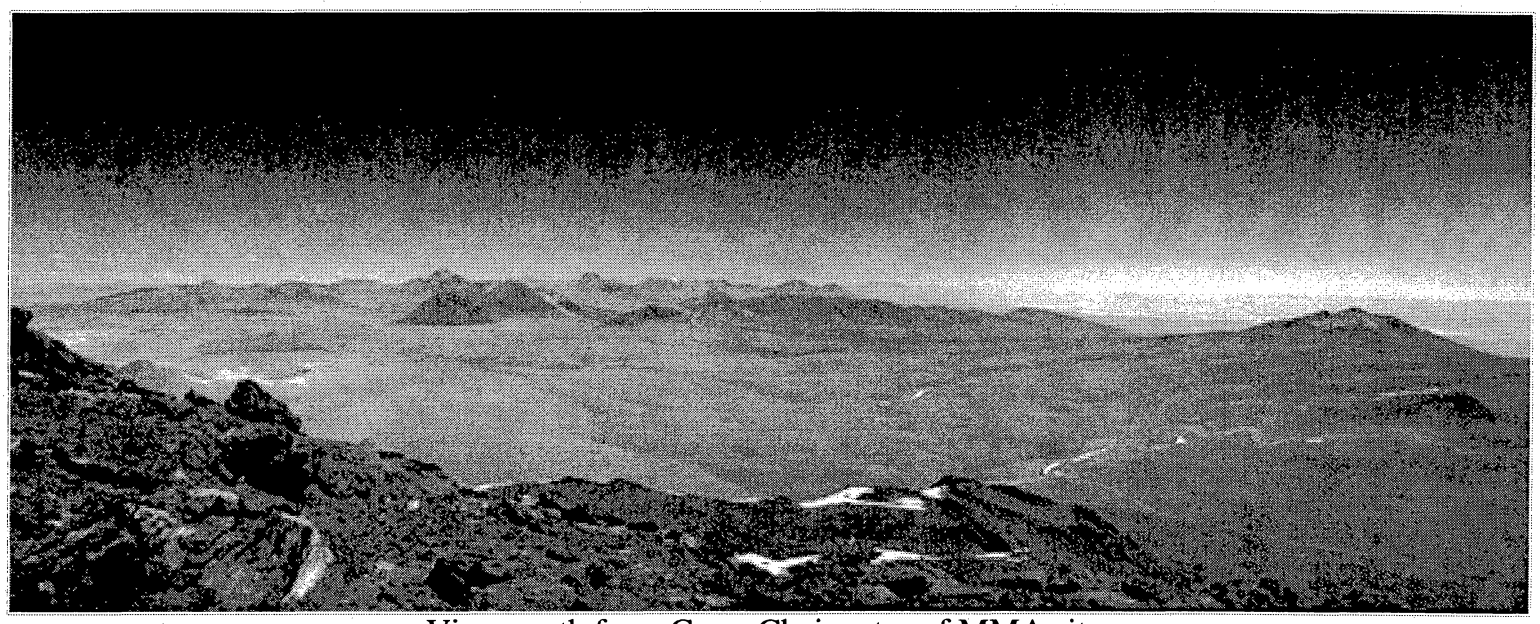
mma@nrao.edu

ALMA Project Book, Chapter 14

SITE CHARACTERIZATION AND MONITORING

Simon Radford

Last revised 2000 September 6



View south from Cerro Chajnantor of MMA site.
Photo: S. Radford, 1994 November.

Revision History:
2000 September 6: Updated, removed WBS numbers.
1999 February 4: Updated URLs.
1998 October 16: Reorganized to match WBS, added section numbers.
1998 July 15: Original version.

14.1 Goals

During the project's prehistory, NRAO conducted extensive measurements to characterize several candidate sites for the Millimeter Array. These studies culminated in the recommendation of an array site on the high (5000 m) plateau southwest of Cerro Chajnantor, Chile, about 40 km east of the village of San Pedro de Atacama. The goals of further site characterization and monitoring are:

- to identify and quantify site conditions and their influence the instrument design or operations concepts,
- to provide a historical record of site conditions to guide priorities for instrument development and operation,
- to maintain a continuous presence on the site through development and construction to the start of operations, and
- to maintain contact and coordinate efforts with other groups working on or near the site.

14.2 Areas of interest

At millimeter and submillimeter wavelengths, pressure broadened molecular spectral lines make the atmosphere a natural limitation to the sensitivity and resolution of astronomical observations. Tropospheric water vapor is the principal culprit. The translucent atmosphere both decreases the signal, by attenuating incoming radiation, and increases the noise, by radiating thermally. Furthermore, inhomogeneities in the water vapor distribution cause variations in the electrical path length through the atmosphere. These variations result in phase errors that degrade the sensitivity and resolution of images made with both interferometers and filled aperture telescopes. The site characterization effort addresses these areas:

- Radiometric properties of the atmosphere
  - stability
  - transparency
- Physical structure of the atmosphere
  - meteorology
  - stratification
  - turbulence
- Physical characteristics of the site
  - topology
  - geology

14.3 Site Infrastructure

The ALMA operations base on Chajnantor are two 20 foot (6 m) long ocean shipping containers. These provides shelter for personnel and physical support for the instruments. In 1995, NRAO installed the first container and in 1998 June ESO installed a second container 15 m north of the NRAO equipment. In 1998 October, a third container was installed 1 km west of the ALMA containers as a launch base for radiosondes.

14.3.1 Safety program

Site inspections (every six months? year?); inventory, inspection, and, if necessary, repair or replacement of safety supplies and equipment; identification, and, if necessary, remediation of safety hazards; training (first aid, high altitude illness, oxygen therapy, fire safety, industrial safety). Note NRAO safety rules.

14.3.2 Solar power system

All three containers have arrays of solar panels and battery banks to supply electrical power. The system on the NRAO container can supply about 500 W continuously (24 VDC and 110 VAC 60 Hz), with sufficient reserve capacity to weather a storm of a few days. With current instrumentation, this system operates near capacity. The ESO container has a slightly smaller system (24 VDC and 220 VAC 50 Hz), also near capacity. A wind turbine has been installed on two occasions to augment the NRAO system, but it broke quickly. System maintenance includes a periodic (yearly) check and refill of battery water.

14.3.3 Communications

Voice and low-speed (<= 9600 baud) data are transmitted by cellular and satellite (Inmarsat A) telephones. In 2000 October, an Inmarsat M4 satellite terminal will be deployed to provide voice and ISDN (64 kbaud) data communications. The Inmarsat A terminal then will be decommissioned. Handheld radios will be used for communications on and around the site.

14.3.4 Transportation

Four wheel drive vehicles are required to access to Chajnantor, especially during inclement weather.

14.3.5 High resolution digital elevation model

1n 1996, contour maps and digital elevation models were prepared from aerial photographs (Maps of the Chajnantor Zone). These cover two 8 × 8 km regions of the Chajnantor and Pampa la Bola areas at 5-10 m resolution. In 1999, these maps were extended to the entire science reserve (18 × 19 km). They will be used for hydrodynamic studies of airflow over the site, for planning the array configurations, and for planning civil works.

14.3.6 Computers and network

All NRAO and ESO instruments are controlled by PCs running Windows 95. They are interlinked with ethernet, which extends to the LSA container. The PC clocks are synchronized to a GPS receiver that provides an absolute time reference good to about 1 s. The GPS receiver was used to determine the position of the container (Memos 261 and 312)>

14.3.7 Auxiliary instruments

- A surveillance camera, installed on 1997 June 15, takes pictures of the southwest horizon every two hours. Data are retrieved about once a month and the images posted.
- A subsurface temperature probe was operated 1997 June - October and 1998 March - May. Data are analyzed in Memo 314.
- A seismometer was installed in 1995. Data are analyzed by Chilean group (K. Bataille). Firmware was updated in 2000 July to accomodate GPS date rollover.

14.3.8 Physiology studies

John West, MD (UCSD) is investigating strategies for improving worker comfort and performance at high altitude. These include enhancing the oxygen concentration of the air in working and living quarters (Memos 191 and 302).

14.4 Atmospheric stability

Inhomogeneities in the distribution of water vapor cause variations in the electrical path length through the atmosphere. The resulting phase errors degrade the sensitivity and resolution of observations with both

interferometers and filled aperture telescopes.

14.4.1 12 GHz interferometer

The site test interferometers directly measure the tropospheric phase stability. They observe unmodulated 11.5 GHz beacons broadcast from geostationary satellites and measure the phase difference between the signals received by two 1.8 m diameter antennas 300 m apart. Because the atmosphere is non-dispersive away from line centers, the results can be scaled to millimeter and submillimeter wavelengths.

Four instruments have been constructed by NRAO's Tucson office. The first was operated near the VLBA antenna (3720 m) on Mauna Kea, Hawaii, from 1994 September to 1996 June, then installed at the VLA in in 1997 May. The second has been operating on Chajnantor (5000 m) near San Pedro de Atacama, Chile, since 1995 May. A third was built for the LSA project. ESO installed it at Pajonales in 1997 April and moved it to Chajnantor in 1998 June. A fourth instrument, with a 100 m baseline, was installed at Green Bank in 2000 March.

The design and operation of these instruments are described in Site Test Interferometer (Radford, Reiland, & Shillue 1996, PASP 108, 441). From the phase time series, we obtain the r. m. s. path fluctuations on a 300 m baseline, the power law exponent of the phase structure function, and the velocity at which the turbulent water vapor moves over the array. Memo 129 describes the site test interferometer data reduction in detail, and Memo 130 illustrates the agreement between two different methods of deriving the mean velocity of the turbulent water vapor flow in the atmosphere.

In 1998 June, the ESO interferometer was set up alongside the NRAO interferometer. They share essentially the same baseline, but observe different satellites about 5° apart on the sky. Lag correlation of the data from the two interferometers will indicate the height of the turbulent layer (see Memo 196).

The interferometers operate autonomously. Status reports are received daily and data are retrieved about once a month. The data are analyzed in Tucson and monthly summaries are posted. Current activity includes operation and maintenance, including sporadic repair as required, data retrieval, and data analysis.

14.5 Atmospheric transparency

Pressure broadened molecular spectral lines, principally of tropospheric water vapor, make the atmosphere semi-opaque at millimeter and submillimeter wavelengths. The translucent atmosphere radiates thermally, which increases the system noise, and attenuates incoming radiation, which decreases the signal.

14.5.1 225 GHz tipper [11.1.3.1]

The 225 GHz tipping radiometer is the benchmark instrument for site characterization. It measures the atmospheric transparency every 10 minutes and the stability of atmospheric emission every fifth hour. Operation is automatic. Daily and monthly data summaries are posted. The data are made available to interested parties in machine readable form. Current activity includes operation and maintenance, including sporadic repair as required, data retrieval, and data analysis.

14.5.1 Submm tipper

A tipping photometer was been developed in collaboration with Carnegie Mellon University to directly measure the atmospheric transparency at 350 μm wavelength. This instrument is based on an ambient temperature, pyroelectric detector. The spectral response is defined by a resonant metal mesh. A compound parabolic (Winston) cone and offset parabolic scanning mirror together define the 6° beam on the sky. The detector is internally calibrated with two temperature controlled loads and views the sky through a woven Gore-tex window. Identical instruments have been deployed on Chajnantor (1997 October), at the CSO on Mauna Kea (1997 December), and at the South Pole (1998 January). An incomplete fourth unit was supplied to the University of New South Wales in 1999 July for modification prior to remote Antarctic deployment. In 2000 June, a fifth instrument equipped with a filter wheel and 200, 260, 350, and 1300 μm filters was deployed at Chajnantor. In 2000 October, this will be redeployed to 5700 m on Sairecabur to investigate the dependence of transparency with altitude in the area of Chajnantor.

The instruments operate autonomously. Status reports are received daily and data are retrieved about once a month. The data from these instruments are being analyzed with the aim of making an unbiased comparison of the three sites. Current activity includes operation and maintenance, including sporadic repair as required, data retrieval, and data analysis. Further work includes cross calibration between the submm tipper and other instruments, namely the 225 GHz tippers, SCUBA, CSO, and AST/RO,

14.5.2 Fourier Transform Spectrometer

To measure the atmospheric emission spectrum at Chajnantor, the Smithsonian Observatory has deployed a Fourier transform (polarizing Martin-Pupplet) spectrometer. This cryogenic instrument covers 350 - 3000 GHz with 3 GHz resolution and a 3° beam. The instrument recorded data for most of the 1998 winter season. NRAO provides the base for field operations.

14.6 Physical structure of atmosphere

The vertical profiles of atmospheric water vapor and turbulence may affect the success of radiometric phase calibration schemes.

14.6.1 Radiosonde campaign

Radiosondes carried by weather balloons provide *in situ* measurements of pressure, temperature, humidity, and wind speed and direction over the launch site. From these data we learn about the stratification of the water vapor over Chajnantor and about shear layers that may generate turbulence. A surplus radiotheodolite was acquired, upgraded by the manufacturer, tested in Tucson, and deployed at Chajnantor. Beginning in 1998 October, balloon flights have been made whenever appropriate personnel are at the site. This campaign is a collaboration between NRAO, Cornell, ESO, and SAO. The balloons are launched from a container placed 1 km west of the main site.

14.6.2 Hydrodynamic models

Calculations of airflow over Chajnantor, with emphasis on turbulence generated by local topography. Collaboration with NOAO.

14.6.3 Sodar

Acoustic sounding, or sodar, senses thermal turbulence in the lower atmosphere. Engineering tests of an ESO sodar unit were made in 1999 November. We are evaluating our interest we have in pursuing further measurements.

14.6.3 Weather stations

Additional weather stations will be deployed to measure the variation of meteorological parameters over the site.

14.7 Technical planning with collaborators and neighbors

Several groups are carrying out site characterization studies or astronomical experiments nearby. NRAO encourages these groups and takes interest in their results. As needed, NRAO and the other groups coordinate activities.

14.7.1 ESO

In 1998 June, ESO redeployed its site characterization equipment to Chajnantor. The ESO equipment, located 15 m north of the NRAO equipment, includes:

- Several weather stations. These are currently deployed adjacent to the containers, but will be deployed across the site in the last quarter of 1998.
- A 12 GHz interferometer. This is set up alongside the NRAO interferometer, sharing essentially the same baseline, but observing different satellites about 5° apart on the sky. Lag correlation of the data from the two interferometers will indicate the height of the turbulent layer (see MMA Memo 196).
- Dual three channel 183 GHz radiometers. These (instruments, designed and constructed by MRAO, OSO, and ESO, measure the H_2O line shape. They are installed at the ends of the LSA interferometer and look in the same direction as the interferometer. Variations in the line shape will then be compared to the phase fluctuations measured with the interferometers. See Memo 271.
- A single channel 22 GHz radiometer (deployment uncertain).

ESO provides field support for the ALMA site characterization program.

14.7.2 NRO/LMSA

At Pampa la Bola, about 7 km northeast of the MMA equipment, the LMSA project has installed:

- Weather stations (Memo 322),
- Dual 220 GHz tipping radiometers,
- A 12 GHz interferometer, and
- A Fourier Transform Spectrometer (temporary deployments).

14.7.3 Cornell

The CAT project is making optical seeing (DIMM) measurements. Campaigns in 1998 May, July, October, etc.

14.7.3 MAT

Observations of fluctuations in the Cosmic Background Radiation by a Princeton/Pennsylvania group. Campaigns in 1997 and 1998.

14.7.3 CBI

Observations of fluctuations in the Cosmic Background Radiation by a Caltech group. Deployment occurred in late 1999, with observations throughout 2000.

14.8 Site Characterization Reviews

Scientific reviews of site characterization data obtained by all groups.

14.8.1 USNC/URSI meeting

At the USNC/URSI National Radio Science Meetingin 1999 January in Boulder, there was be a session on Atmospheric Transmission at Millimeter and Submillimeter Wavelengths. Results from the NRAO site characterization program will be presented.

14.8.2 Mid-term Review

2000 March 22

14.8.3 Final Review

2001 March

References

ALMA Site Studies at NRAO ALMA Site Studies at ESO

MMA site NRAO safety rules

NRAO 1998 May, Recommended Site for the Millimeter Array [also ps]

Delgado, G., Otárola, A., Belitsky, V., & Urbain, D., The Determination of Precipitable Water Vapour at Llano de Chajnantor from Observations of the 183 GHz Water Line, MMA Memo 271

Gerard, A. B., McElroy, M. K., Taylor, M. J.,Grant, I., Powell, F. L., Holverda, S., Sentse, N., &West, J. B., 2000, Six Percent Oxygen Enrichment of Room Air at Simulated 5000 m Altitude Improves Neuropsychological Function, High Altitude Medicine & Biology 1, 51; ALMA Memo 302

Holdaway, M. A., & Radford, S. J. E., 1998, Options for Placement of a Second Site Test Interferometer on Chajnantor, MMA Memo 196

Holdaway, M. A., Gordon, M. A., Foster, S. M., Schwab, F. R., and Bustos, H., 1996, Digital Elevation Models for the Chajnantor Site, MMA Memo 160

Holdaway, M. A., 1995, Velocity of Winds Aloft from Site Test Interferometer Data, MMA Memo 130

Holdaway, M. A., Radford, S. J. E., Owen, F. N., & Foster, S. M., 1995, Data Processing for Site Test Interferometers, MMA Memo 129

Radford, S. J. E., Reiland, G., & Shillue, B., 1996, Site Test Interferometer, PASP 108, 441

Radford, S. J. E., 1999, Position of MMA Equipment on Chajnantor, MMA Memo 261

Radford, S. J. E., 2000, Refined Position of ALMA Equipment on Chajnantor, ALMA Memo 312

Sakamoto, S., Handa, K., Kohno, K., Nakai, N., Otárola, A., Radford, S. J. E., Butler, B., & Bronfman, L., 2000, Comparison of Meteorological Data at the Pampa La Bola and Llano de Chajnantor Sites, ALMA Memo 322

Snyder, L. A., Radford, S. J. E., & Holdaway, M. A., 2000, Underground Temperature Fluctuations and Water Drainage at Chajnantor, ALMA Memo 314

West, J. B., Powell, F. L., Luks, A. M., 1997, Feasibility Study of the Use of the White Mountain Research Station (WMRS) Laboratory to Measure the Effects of 27% Oxygen Enrichment at 5000 m Altitude on Human Cognitive Function, MMA Memo 191

ALMA Project Book, Chapter 15: Array Configuration

Antenna Configurations for the ALMA

Min S. Yun
Last changed 2000-12-11

Revision History:
11/11/98: Added summary and milestone tables. Added chapter number to section numbers, tables and figures (Tamara T. Helfer & M.A. Holdaway).
3/30/00: Updated for the ALMA specifications and added sections describing the zoom spiral configuration concept (Min S. Yun).

Summary

Design concepts and sample layouts for antenna configurations for the ALMA are presented.

Table 15.1 Guidelines for Configuration Design

Main D&D Task	Design a set of configurations which allow for a range of angular resolution and sensitivity
Flexible design philosophy	Configurations must allow for graceful expansion through possible collaboration
Costing	Optimize for shared stations to minimize cost
Site placement	Choose specific locations for antenna placement on Chajnantor site

15.1 Number and Size of Antenna Elements

We assume that the ALMA will comprise N = 64 antennas of 12 m diameter. The geometric collecting area is then 7240 sq. m. The ``collecting length" nD, the appropriate measure of the mosaicing sensitivity and for the fraction of occupied cells, is 770 m. The collecting area and collecting length are increased by factors of 3.6 and 2.4, respectively, over the old MMA plan with N = 40 and D = 8 m. We note that the array configuration development plan will need to react to the changes and refinements in the array's concept, particularly with regard to possible collaboration with Japanese partners.

15.2 Limiting Configurations, Number of Configurations, and Resolution Scale Factor

15.2.1 Size of the Most Compact Array

The choice of a compact configuration for the ALMA is driven by the desire to maximize surface brightness sensitivity, which is achieved by placing the antennas as close together as is practical.

If we assume a filling factor of 40%, which is a reasonable compromise between the competing requirements of close packing and the resultant maximum acceptable sidelobes, then

the maximum baseline for the compact array is = 150 m.

15.2.2 Size of the Largest Array

The largest configuration is assumed to have a maximum baseline of 3 km. A separate array of 10 km diameter or larger is also being considered.

15.2.3 Number of Configurations and Resolution Scale Factor

Given the assumed sizes of the minimum and maximum arrays, Holdaway (1998a) has performed a cost-benefit analysis for the number of MMA configurations, which showed that the observing efficiency of the MMA would be close to optimal with 4-8 configurations. The calculation appropriate for ALMA produces a similar conclusion (Yun & Kogan 1999). The strawperson configurations consisting of minimum sidelobe donut/double-rings (Kogan 1998) have the resolution scale factor between adjacent configurations of about 2.1, with the outer diameters of 325 m, 680 m, 1430 m, and 3000 m. By sharing the outer ring of antennas as the inner ring of the next larger configuration, the maximum re-use of pads and minimum antenna moves per reconfiguration can be achieved. The most compact array takes advantage of the existing pads in the central compact configuration. The most compact arrays will provide essentially complete sampling of the (u,v) plane in a snapshot observation. The larger arrays will require longer tracks for good imaging and sensitivity.

The main advantage of having configurations with durations longer than a few days (see below) is that there is a greater chance of accomplishing observations with critical weather requirements (such as total power mosaic and high frequency observations) using the full sensitivity of the array. As weather patterns typically persist over several days or even weeks, any "good" observing weather may come in bursts of this time scale. Requiring reconfigurations only on monthly or bi-monthly basis also means that the full complement of 64 antennas are available for most of the observations, unlike for the continuous reconfiguration case.

The minimum and maximum baselines for each array are listed in Table 15.3, along with the size of a sample beam and the time required for the half of the (u,v) cells to be sampled (FOC = 0.5). Note that the shortest baseline does not correspond to the largest angular structure to which the array will be sensitive, as mosaicing with total power data will permit arbitrarily large sources to be imaged.

Table 15.3 Specifications for the ALMA strawperson configurations.

Array	Minimum Baseline [m]	Maximum Baseline [m]	Array Style	Time for FOC = 0.5 [hours]	Natural Beam at 345 GHz [arcs]
A	30	3000	donut	10	0.050
B	24	1430	donut	2	0.101
C	18	680	donut	0.1	0.22
D	16	325	donut	0.1	0.47
E	16	150	filled	0	0.97

An alternative scheme of a self-similar distribution of pads has also been proposed (Conway 1998, Webster 1998). In such schemes the pad density per unit area falls off as 1/radius squared, and in each configuration the pads over a certain range of radii are occupied. The uv coverages of each configuration are then close to being self-similar, which is useful for projects comparing lines at different frequencies. Because these arrays are continuously self-similar, there is the possibility of accomodating continuous variations in resolution and so these arrays are referred to as 'zoom' arrays. Such schemes also naturally have large pad sharing between configurations. Zoom arrays also naturally give centrally condensed uv coverages, which for the approprate choice of radius over which the pads are occupied in any given configuration (i.e. a factor of about 8) give uv coverages which are close to having a natural Gaussian taper (Conway 1998 and see section 15.3.2.3).

An advantage of zoom arrays is that they can be operated in a highly flexible way. The choice of the number of configurations, resolution step, and the resolutions offered can be left to be decided after construction and can be changed during the lifetime of the array in response to scientific requirements. Depending on the operational constraints, the array may reconfigure continously or in bursts.

The scientific and operational advantages of the continuous reconfiguration mode are further discussed in detail by Conway (1998, 2000a) and by Webster (1998). Observations requiring exact resolution can be better accomodated, and the loss of sensitivity by tapering can be avoided more easily. The total observing efficiency for a continously reconfiguring mode has been argued to be higher than that of a burst mode (Guilloteau 1999, Conway 2000a) although the magnitude of this effect may be small and subject to the type of observations conducted. If only one antenna is moved each day for each transporter, then all antenna reconfigurations may be accomplished within the first few hours of the working day, before high afternoon wind makes moving antennas more difficult.

15.3 Fourier Plane Coverage

15.3.1 Compact Array

The driving considerations for the compact array are maximum surface brightness sensitivity and excellent mosaicing capability. Surface brightness sensitivity is optimized by designing an array with the largest synthesized beam possible, which is achieved by having the shortest baselines possible. The ALMA will be a *homogeneous array*, with total power and interferometric data being collected by the same antennas (Cornwell, Holdaway, & Uson, 1994). Homogeneous array mosaicing image quality is optimized by having a high density of the shortest interferometric baselines and by minimizing the sidelobes in the synthesized beam. Optimizing the short baseline coverage is best achieved with a filled array, which produces a Fourier plane coverage that to first order is a linearly decreasing function of (u,v) distance. The shortest baselines are limited strictly by the minimum safe distance which avoids mechanical collision of the antennas when pointing in arbitrary directions, which depends upon the antenna design, and less strictly by shadowing requirements. Configurations with the highest density of the shortest baselines will be a hexagonal close pack distribution of antennas, which results in a very large grating response in the synthesized beam and is therefore not acceptable. With a minimum distance between antennas of 1.28 D, we can achieve a reasonable sidelobe level of a few percent rms with an array filling factor of 40%. Such a filled compact array will result in complete instantaneous (u,v) coverage, even with only 32 antennas. Some degree of optimization is required, trading off between good short baseline coverage and a large beam on the one hand and minimum synthesized beam sidelobes on the other.

The short spacing requirement for homogeneous array mosaicing, together with the physical inevitability of shadowing, require multiple compact configurations for observations of sources at various declinations. A plan with three compact arrays has been considered (Helfer & Holdaway 1998). The E1 array, with a North-South elongation of 1.2, will cover zenith observations down to somewhat below the onset of shadowing at 50 deg; the E2 and E3 will be progressively more elongated, with elongation of about 1.6 and 3. These three arrays will cover most of the range of declinations available from the Chajnantor site, -90 deg 53 deg. Observations of the small fraction of the sky which is further north and still visible from Chajnantor will need to be conducted in a hybrid configuration.

The E1 and E2 configurations will utilize a mechanical elevation stop which will limit the elevation to be above 20 deg. This limitation will permit a closer packing of the antennas in the E1 and E2 configurations. While it does remove flexibility from the compact arrays, the E1 and E2 arrays would be largely shadowed below this elevation anyway. The elevation stop will be removed from all or most antennas when they are reconfigured into the E3 array, where the antennas will be sufficiently separated that collisions are no longer a possibility. The general specifications for the E1, E2, and E3 configurations are shown in Table 15.4.

Table 15.4 Specifications for the compact configuration N-S elongations.

Array	Min. N-S	Elev. of first	Min. observing	Max. observing	N-S

	Distance	Shadowing	Elevation	Elevation	Elongation
E1	1.3 D	50 deg	40-45	90	1.2
E2	1.9 D	31 deg	30	50+	1.6
E3	3.0 D	19 deg	14	33+	2.9

An important consideration for having several compact configurations is the cost of building the pads, roads, and cables, though we still need to investigate what these costs are. Overlapping the stations will keep the cost and time involved in reconfiguring the antennas to a minimum. Antenna access is another important consideration: maximize the number of antennas that can be moved by a transporter without moving any other antennas.

15.3.2 Intermediate Arrays

Mosaic observations will also be made in some of the intermediate configurations as the almost complete instantaneous (u,v) coverage and good brightness sensitivity can be achieved with an abundance of short baselines. Quickly increasing atmospheric noise with increasing air mass dictates that it is best to observe a source within a few hours of transit (Holdaway, 1998b), and these configurations should be optimized for short tracks observed within a few hours of transit, over a range of declinations.

The larger array require about 2 hours to achieve complete (u,v) coverage (FOC = 0.5, see Table 15.4), so it should be optimized for somewhat longer tracks, but still within a few hours of transit (Holdaway, 1998b).

The largest array requires up to 10 hours to achieve essentially complete (u,v) coverage. At +/- 5 hours off transit, the sensitivity loss due to the atmosphere will be severe at most frequencies; also, some sources are not above the minimum elevation limit for such long tracks. Nonetheless, the largest array should be optimized for long integrations, keeping in mind that it must also have respectable snapshot coverage for those sources strong enough to be observed in this mode.

As a general requirement, we will want to have some of the shortest baselines (i.e., 16-20 m) present in even the largest arrays to permit single configuration mapping of many wide field objects (Braun, 1993). However, if there is a lot of large structure in the object, multiple configuration imaging may be required. At this point, we do not have a coherent strategy for when to combine data from multiple configurations, nor have we considered the impact of multiple configuration observations on the set of configuration designs. Processing multiple configuration data through a data pipeline may pose an added complexity. Designing in short spacing baselines in all configurations is desirable if it can be accomplished without any serious compromise in overall imaging performance.

Several competing philosophies are currently under consideration for the Fourier plane coverage for the intermediate configurations. One philosophy is to achieve as complete coverage in the Fourier plane as is practical; this approach leads to ring-like arrays, as characterized by Keto

(1997) for snapshot observations and by Holdaway, Foster & Morita (1996) for longer tracks. Donut or double-ring arrays with minimum sidelobes in the synthesized beam, as implemented by Kogan (1997, 1998a), offer many highly desirable characteristics as well. A zoom spiral array proposed by Conway (1998,1999) produces a strongly tapered uv-coverage with continuously changing resolution.

15.3.2.1 Reuleaux Triangles

Keto's Reuleaux triangle configurations, and ring-like configurations in general, yield fairly uniform (u,v) coverage plus a narrow peak at small spatial frequencies. They also offer the advantage of achieving the maximal sensitivity for the longest baselines, resulting in smaller naturally weighted resolution than other types of arrays with the same maximum baseline, which is an attractive characteristic. However, true uniform coverage in the Fourier plane has disadvantages as well:

- the sharp cutoff in (u,v) sampling at large spatial frequencies results in large (10-15%) sidelobes close to the central lobe of the synthesized beam (Holdaway 1997), which may complicate an image deconvolution and thereby lower its dynamic range (Holdaway 1996).

- optimization techniques like the elastic net method used by Keto have so far tended to produce large diameters for the central hole in the Fourier plane coverage. It is probable that this problem can be alleviated to some extent, either by using nested rings or Releaux triangles, or by changing the optimization conditions to include some number of short baselines. The nested triangle approach destroys the uniform Fourier plane coverage.

- unpublished simulations by Morita and by Holdaway show that the excess short spacing coverage which a ring array provides is actually more responsible for high dynamic range in wide-field reconstructions than the uniform Fourier plane coverage.

Webster (1998) has investigated the idea of using nested rings or nested Reuleaux triangles to achieve a compromise between uniform (u,v) coverage and sensitivity to extended structure.

15.3.2.2 Minimum Sidelobe Donut/Double-Ring Concept

Kogan's algorithm produces antenna configurations which minimize the maximum sidelobe levels of the point spread function in some region of the image plane. This approach has the advantage of producing PSFs which should introduce fewer problems in image deconvolution. Kogan has also pointed out that in general, sidelobes that are close to the peak of the PSF can be alleviated using a moderate taper (at the expense of small loss in sensitivity). Another attractive feature of Kogan's approach is that it naturally shrinks the hole in the center of the (u,v) plane as the optimization extends over larger and larger regions in the image plane. This produces good coverage at short baselines in the (u,v) plane, which is one of the main shortcomings of the uniform (u,v) coverage optimization described above.

Kogan's code is flexible and can accept a variety of topographical constraints as inputs. Kogan has investigated arrays with the antennas distributed within an annulus with a fixed outer radius and with varying inner radii. Such a ``donut'' configuration can achieve a significantly more tapered beam (1/r distribution of visibility density) than a single ring configuration, thereby

reducing its near-in sidelobes. The configurations can be designed with maximum pad sharing so that each reconfiguration requires moving only 32 antennas. This added dimension is responsible for the decrease in the scaling factor from 4 to 2.1 compared with the MMA strawperson ring configurations discussed by Helfer & Holdaway (1998) -- see Table 15.3. Another natural advantage of the donut configuration is that hybrid configurations with a N-S elongations of 2 and 3 are naturally achieved during the reconfiguration, and fewer still antennas need to be moved during each reconfiguration if such an intermediate hybrid configuration is needed.

The Kogan arrays are optimized for a snapshot in the zenith direction only; however, changing the declination should change only the positions and not the amplitudes of the sidelobes for a snapshot observation (Kogan 1998a). Preliminary tests suggest that these donut/double-ring arrays are robust against random removal of some fraction (e.g. 10%) of antennas. The earth rotation synthesis tends to fill the gaps in the uv-coverage, suppressing the far sidelobes.

One possible disadvantage to Kogan's approach of minimizing the maximum sidelobe within some region of the points spread function is that rather large sidelobes can lurk just outside the region of optimization. Extending the region of optimization to the full width of the primary beam and applying a weighting function which emphasizes the minimization of the close in sidelobes may be desirable.

We plan to study the ramifications of these competing philosophies and ultimately to select a design based on imaging simulations of sources of different size and structure.

15.3.2.3 Zoom Spirals and Gaussian uv Coverage

Conway (1998) has described a three-armed logarithmic 'zoom' spiral array which is self-similar and allows the possibility of having continuously variable resolution. The most recent version of this array is described by Conway (2000a), in which the spiral arm pitch angle gradually changes so the spiral array becomes a ring in its largest configuration, giving close to uniform uv coverage. The largest 3km configuration therefore has the desirable property that it gives close to the maximum resolution from a given limited 3km diamter area. In contrast for smaller configurations the zoom spiral gives a centrally condensed uv coverage, similar in uv density versus radius to the VLA, but not as extreme. It has been argued that such centrally condensed uv coverages may have significant advantages with respect to imaging quality, and optimising all configurations except the largest for such condensed uv-coverages rather than resolution has been proposed (Conway 2000b). Note that the maximum baseline length is twice as large compared to that for a ring array having the same resolution.

Not only do zoom spirals give centrally condensed uv coverages but for appropriate choices of design parameters the uv point density falls off with radius in a manner very close to Gaussian (Conway 1998). The central part of the zenith snapshot beam then has an almost Gaussian shape and near-in sidelobes are minimized. Conway (2000b) has argued that the resulting arrays with minimized near-in sidelobes give better imaging performance than those which have been optimised to reduce far-out sidelobes. For a uv coverage with a natural Gaussian taper the natural beam is also close to the conventional gaussian restoring beam. This opens up the possibility that narrow-field low dynamic range images can be made without deconvolution (i.e. 'Dirty Imaging') although achieving a high dynamic range will ultimately require some form of deconvolution.

Naturally gaussian tapered uv coverages can be considered as having a well sampled core plus outlier points which help constraining the visibilty extrapolation at higher spatial scales. Sparcer uv coverage in the outer radii would require some interpolation in uv space, however. More uniformly filled coverages require less interpolation over the entire sampled region, but some extrapolation beyond the sampled radius may be needed for support during the deconvolution. For more complex images both uv extrapolation and interpolation become more difficult, and simulations involving a large variety of test images may be needed to determine which of these errors are more important in achieving a good imaging performance.

15.4 Sample Configurations

Since the configuration optimization is still in progress, we do not attempt to present optimized configurations in this document. However, to give the reader a feel for the arrays that the Keto and Kogan algorithms produce, we present sample configurations in Figures 15.1 through 15.6.

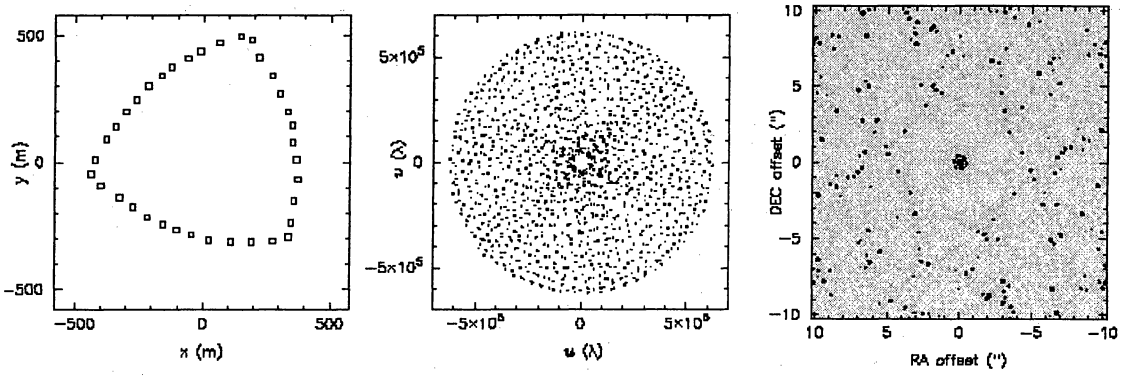


Figure 15.1: Sample Keto array snapshot at 230 GHz. (*left*) Antenna locations in meters, (*middle*) snapshot (u,v) coverage, and (*right*) the resulting synthesized beam, with contours are at 10, 20, 40, 60, 80, 100%. Note the large inner sidelobes.

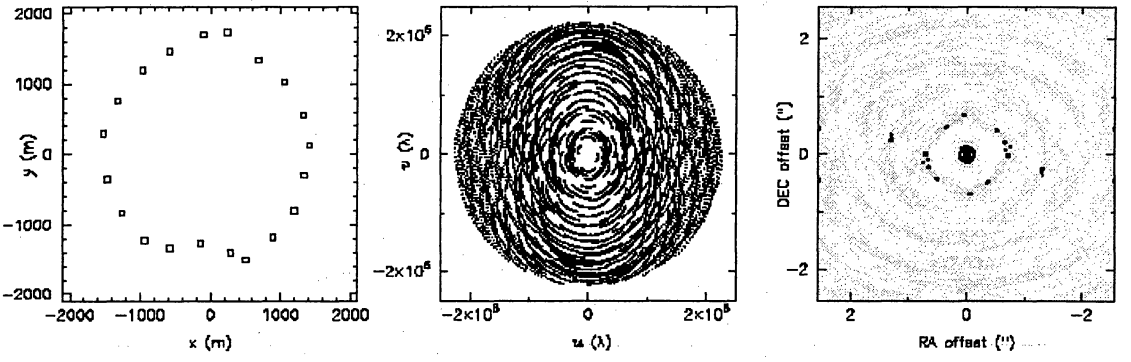


Figure 15.2: Sample 3 km array track for a Keto 20-element array optimized for 4-hour tracks (Holdaway, Foster, & Morita 1996), at 230 GHz. The contours are 0.05, 0.10, 0.15, 0.20, 0.40, 0.60, 0.80, 1.0. The outer sidelobes are reduced for long tracks, but the inner sidelobes remain high.

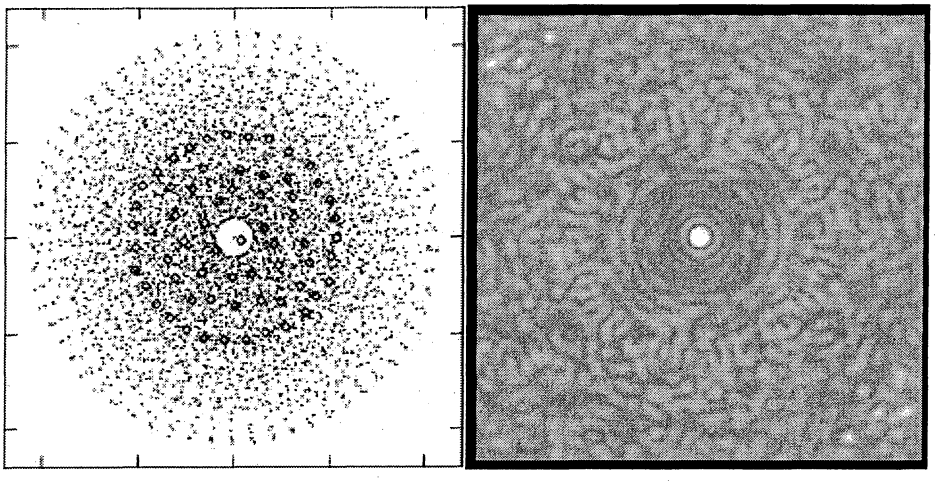


Figure 15.3: Sample compact array pad positions are plotted in diamonds on the left panel along with the zenith snapshot uv coverage. The resulting naturally weighted dirty beam is shown on the right. The greyscale is between -0.05 and +0.10, and the largest sidelobe inside the primary beam is about 5%.

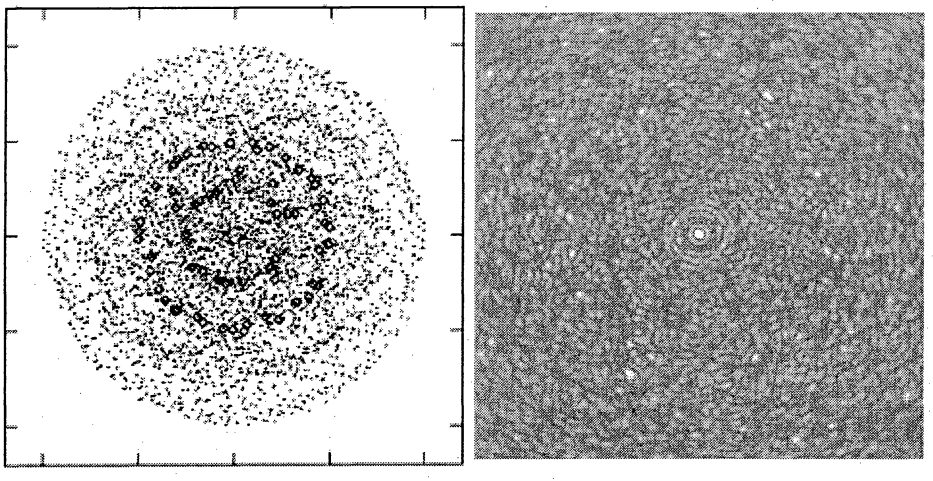


Figure 15.4: Sample Kogan 3 km array pad positions are plotted in diamonds on the left panel along with the zenith snapshot uv coverage. The resulting naturally weighted dirty beam is shown on the right. The greyscale is between -0.05 and +0.10, and the largest sidelobe inside the primary beam is about 10%.

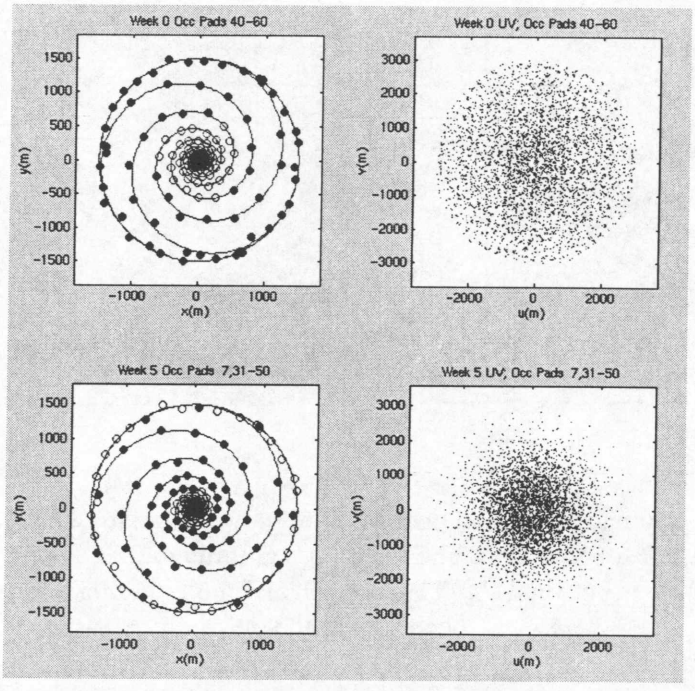


Figure 15.5: The array and zenith uv coverage for the 'zoom spiral' array concept proposed by Conway (1998 and 2000a). Open circles are unoccupied pads and filled symbols are antennas. The top row shows the largest configuration, most antennas lie on the outer ring and so the resolution is maximized. The bottom row shows the array and uv coverage after moving inward approximately half of the antennas. The uv coverage is now centrally condensed and close to gaussian distributed. The resolution is about a factor of 2 less than in the largest configuration. For a 5 month cycle time this array would be reached at week 5.

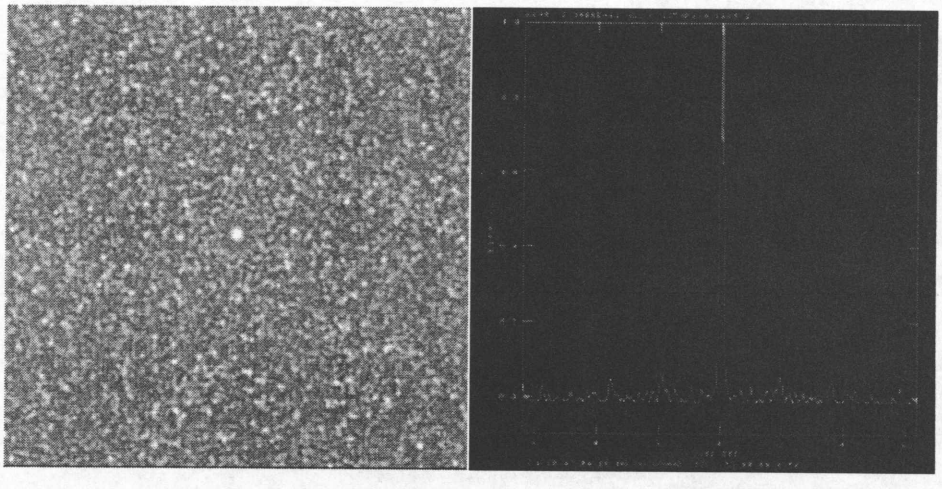


Figure 15.6: Dirty Beam plots for zoom spirals (see Conway 2000b), for zenith snapshot (left) and the N-S slice through the resulting dirty beam (right). The grayscale is between -0.05 and +0.10.

15.5 Hybrid Arrays and Optimal N-S Elongation

From a study of the deviation of synthesized beams from circular as a function of source declination, Foster (1994) concluded that the optimal North-South elongation for tracks of varying length was in the range 1.1-1.3. In order to optimize the elongation of all of the arrays, it is important to know the expected source distribution with declination. Holdaway et al. (1996) assumed a model source distribution in order to estimate the pointing errors for the MMA antenna design. We are now in the process of looking at IRAS source distribution with declination in order to get a better estimate of this function.

Hybrid arrays, made by using stations in adjacent configurations, can be used to help minimize the shadowing and to achieve more circular beams for low-elevation sources. As stated above, a set of hybrid arrays is absolutely required for the compact configurations, but not so crucial for the larger arrays. Hybrid arrays will be studied more in the future when the basic arrays are better determined.

15.6 Topographical Constraints

All of the arrays need to be compatible with the topographic limitations of the Chajnantor site. The basic constraints of the site topography have already been implemented in both the Keto and Kogan algorithms. Digital elevation models (DEMs) of the entire science reserve is now available, and digital topological masks are now being developed (e.g. Butler et al., in preparation).

Radially symmetric configuration designs such as the zoom spiral array or donut/double-ring arrays are intrinsically more susceptible to the topographical constraints, and their implementation may be costly. Preliminary attempts to fit both types of arrays into the topographical constraints are promising (Conway 2000c, Kogan 2000). Development of a cost-mask is highly desirable.

15.7 Interfaces With Other Parts of the ALMA Project

- Antenna: minimum distance for close packing, hard elevation stops
- Antenna: transporter issues, such as intervening antenna clearance, road grade, etc.
- Site Development and Antenna: Antenna Pad Design.
- Site Development: Road Design.
- Local Oscillator/System: underground cables, subarrays for post-move calibrations.

15.8 Other issues to be addressed

There are other issues which have not been examined in this document which deserve closer attention. For example, the arrays need to be optimized for different source declinations, or simultaneously for multiple declinations. Other issues that require further considerations include the effect of earth rotation synthesis (versus snapshots), random losses of antennas, and multi-configuration observations.

Operational concerns such as the frequency and the mode of reconfiguration have also received much attention lately (e.g. Guilloteau 1999, Radford 1999, Yun 1999, Conway 2000). While the ongoing configuration design studies are mostly concerned with the scientific and imaging requirements, weighing the practical concerns such as the cost and maximizing the observing efficiency will ultimately be included in the evaluation process.

References

Braun, R., 1993, ``Telescope Placement at the VLA for Better Single Configuration Imaging", VLA Scientific Memo 165.

Conway, J., 1998, ``Self-Similar Spiral Geometries for the LSA/MMA", MMA Memo #216.

Conway, J., 1999, ``A Comparison of Zoom Arrays with Circular and Spiral Symmetry", MMA Memo #260.

Conway, J., 2000a, ``Observing Efficiency of a Strawperson Zoom Array", MMA Memo #283.

Conway, J., 2000b, ``First Simulations of Imaging Performance of a Spiral Zoom Array; Comparisons with a Single Ring Array", MMA Memo #291.

Conway, J., 2000c, ``A Possible Layout for a Spiral Zoom Array Incorporating Terrain Constraints", MMA Memo #292.

Cornwell, Holdaway, and Uson, 1994, ``Radio-interferometric imaging of very large objects: implications for array design", A&A 271, 697-713.

Foster, S.M. 1994, ``The Optimum Elongation of the MMA A Configuration", MMA Memo #119

Guilloteau, S. 1999, ``Reconfiguring the ALMA Array", MMA Memo #274

Holdaway, M.A., 1998a, ``Cost-Benefit Analysis for the Number of MMA Configurations", MMA Memo #199

Holdaway, M.A., 1998b, ``Hour Angle Ranges for Configuration Optimization", MMA Memo #201

Holdaway, M.A., 1997, ``Comments on Minimum Sidelobe Configurations", MMA Memo #172

Holdaway, M.A., 1996, ``What Fourier Plane Coverage is Right for the MMA?", MMA Memo #156

Holdaway, M.A., Foster, S.M., Emerson, D., Cheng, J., & Schwab, F. 1996, ``Wind Velocities at the Chajnantor and Mauna Kea Sites and the Effect on MMA Pointing", MMA Memo # 159

Holdaway, M.A., Foster, S.M., & Morita, K.-I. 1996, ``Fitting a 12 km Configuration on the Chajnantor Site", MMA Memo #153

Keto, E. 1997, ``The Shape of Cross-correlation Interferometers", ApJ, 475, 843

Kogan, L. 1997, ``Optimization of an Array Configuration Minimizing Side Lobes", MMA Memo # 171

Kogan, L. 1998a, ``Opimization of an Array Configuration with a Topography Constraint", MMA Memo #202

Kogan, L. 1998b, ``Opimization of an Array Configuration with a Donut Constraint", MMA Memo #212

Kogan, L. 1998c, ``A, B, C, and D Configurations in the Shape of Concentric Circles", MMA Memo #217.

Kogan, L. 1999, ``The Imaging Characteristics of an Array with Minimum Side Lobes", MMA Memo #247.

Kogan, L. 2000, ``Fitting of the Largest Configuration (10 km) into the Terrain at the Chajnantor Site", MMA Memo #296.

Radford, S. 1999, ``Antenna Transport Times and Reconfiguration Schedule", MMA Memo #280.

Webster, A. 1998, ``Hybrid Arrays: The Design of Reconfigurable Aperture-Synthesis Interferometers", MMA Memo #214.

Webster, A. 1999, ``Hybrid Arrays: I. The Inner and Outer Hybrids", MMA Memo #239.

Woody, D. 1999, ``ALMA Configurations with Complete UV Coverage", MMA Memo #270.

Yun, M. 1999, ``Sensitivity Loss vs. Duration of Reconfiguration and ALMA Array Design", MMA Memo #276.

Yun, M. & Kogan, L. 1999, ``Cost-Benefit Analysis of ALMA Configurations", MMA Memo #265.

ALMA Project Book, Chapter 16
By: Robert Brown and Daniel Hofstadt

Site Development

Summary

The site development task is specified in its entirety by all the subtasks to the level-1 task 2.0 of the ALMA WBS.

The ALMA site development task includes development of four functionally, and geographically, separate sites. These are:

1. The array site in the Altiplano of the second Region of Chile near Cerro Chajnantor;
2. The Operational Support Facility (OSF);
3. A communications link, or *highway*, connecting the OSF to the Chajnantor array site;
4. The Administrative office in Santiago.

16.1 Requirements

Development of the array site includes construction of all the civil works on the array site. This includes: the concrete antenna foundations and antenna attachment points (which together are referred to as *antenna stations*); provision for electric power and distribution of that power to all the antenna stations; an array control building for array operation and all centrally-located electronic equipment; an optical fiber communications network between the control building and all the antenna stations; water and sewer facilities for the control building; a maintenance building for repairs that must be done on-site; a residence facility for use in emergencies or in situations where personnel travel from the site cannot be done safely; all needed site and access roads; all site furnishings, maintenance equipment and vehicles. The task does not include the process by which the necessary permissions are secured to make use of the site.

Development of the Operations Support Facility includes construction of the civil works needed for ALMA personnel to operate and maintain the array. The OSF, to be built near the village of San Pedro de Atacama, will be sized to the requirements of a turno staff that is responsible for scientific operations, the technical maintenance and logistic support activities. The task includes construction of an office building, laboratories, residence facility and a building for antenna assembly, along with the equipment and furnishings needed. It includes provisions for adequate electric power, water, sewage disposal, and access roads. Provision for the construction base camp to be used by the civil works contractor is part of this task. It does not include the process by which the necessary permissions are secured to make use of the land.

The OSF-Chajnantor link is an optical fiber communication between these two sites. The development task includes trenching for the link and burial of the optical fiber cable. It also includes procurement of the optical fiber.

Development of the Administrative office in Santiago involves provision for office space for the Administrative personnel needed for ALMA in Chile and the scientific offices and facilities needed for the scientific staff that are resident in Santiago and have functional duties at the OSF. It includes the necessary furnishings but does not include the cost of whatever land may be required to erect the facilities.

16.2 Development Overview

The three ALMA development tasks in northern Chile will be done by commercial contractors under competitive bid. Beginning in 2002, architectural and engineering plans will be developed with commercial firms working in concert with ALMA staff. Those plans will be executed under contract by the successful bidders in 2003 and successive years. The work will be scheduled to meet the requirement that sufficient facilities must be in place, both at the OSF and at the array site, to receive the first ALMA antenna in Chile from the antenna contractor in the third calendar quarter of 2004. That antenna and its successors will be *outfitted* (equipped with ALMA-supplied instrumentation) at the OSF and transported to the array site for commissioning. This means that the first phase of the necessary ALMA infrastructure at the OSF and on the array site, must be in place by the third quarter of 2004. Temporary facilities will be provided to meet the increasing ALMA infrastructure demands until the permanent facilities are completed and available for occupancy.

16.3 Development of the Four ALMA Sites

16.3.1 The Array Site

The telescope site lies at an elevation of 5000 meters (16, 400 feet) in Region II of Chile, at a latitude of 23°S. Geologically, the site is a "bench" on the western side of the Andes range, with excellent drainage and a line-of-sight to nearby community.

Logistically, the site has three important advantages: easy access, proximity to developed communities, and a gas pipeline that traverses the site. It lies near an international highway, the Paso de Jama, that is wide and paved with asphalt sufficient for traffic by heavy commercial vehicles. The site is within a one-hour and a half drive (55 km or 34 miles) east of the tourist village of San Pedro de Atacama; it is within a two and a half-hour drive (180 km or 110 miles) southeast of the mining support city of Calama. Calama is served daily by three Chilean airlines with flights from Santiago. The nearest port to the site is Antofagasta. Antofagasta is a large (population 230,000) city with a wide range of industrial facilities to support heavy industry.

Access to the site will be provided by a highway link to the Paso de Jama. Three options for this access road are being studied: (1) connection from the Paso de Jama on the western side of Cerro Toco and Cerro Chajnantor; (2) connection from the Paso de Jama on the eastern side of Cerro Toco and Cerro Chajnantor; (3) direct connection to the OSF from the array site, without using the Paso de Jama. This latter option is most desirable in the event the OSF is located at some distance intermediate between the array site and San Pedro de Atacama. The ALMA Project baseline is option (1). A second option will be maintained as an emergency link.

Electric power on site can be provided either by local generation from natural gas obtained from the gas pipeline running across the site that is operated by Gas Atacama or from commercial power brought to the site from the power grid west of the Salar via overhead power lines. The ALMA Project baseline is local generation of power with gas turbines purchased and operated by ALMA. The commercial option remains under study. Such an option would allow for an easy connection to the Chilean communications network along the power line.

Potable water is a difficult but solvable problem. There are two options. Water can be obtained from wells dug at the OSF and trucked to the site. Or, according to a consultant's analysis, potable ground water is available from wells dug on the site itself. The ALMA Project baseline is to transport water from the OSF; the option of a local well remains under study.

Development of the buildings and roads on site will be done under competitive bid by commercial firms with expertise building at high elevation in Chile. The site civil works will include:

- Antenna foundations, 250 total. Design of the antenna foundation is a deliverable from the antenna contractors that are building the prototype ALMA antennas;
- Connections among the antenna stations and from the antenna stations to the control building. This includes trenching and laying of the cables. The contractor will supply the electrical cables and the optical fiber cables as specified by the ALMA Project;
- Access roads on site and to each of the antenna stations;
- An antenna assembly hall for maintenance of the antennas on site;
- Array control building with office and laboratory space (1000 square meters total);

The ALMA Project, not including the site civil works contractor, will provide furnishings and equipment for the site buildings and it will provide all needed site vehicles and safety equipment.

16.3.2 Operations Support Facility (OSF)

The OSF is the main focus of ALMA. In the construction phase, it is the location where the antennas will be erected by the antenna contractors and accepted by the ALMA Project. Once accepted, the antennas will be outfitted with instrumentation provided by the ALMA Project. From there they will be transported, by the ALMA Project, to the high array site. In the operations phase of ALMA, the scientific operations of the array will be based at the OSF. The data quality will be verified by ALMA scientists at the OSF and plans will be made for the future observing program. The engineers and technicians assigned to maintenance of the array instrumentation and software will be based at the OSF. Repaired instrument modules will be taken to the array site and defective ones returned from the array site to the OSF for repair. The local ALMA managers will have their offices at the OSF also. The purpose of the OSF is to provide an environment at a sufficiently low elevation, but still close to the array site, where operational personnel can carry out their functions efficiently. The OSF serves to minimize the number of people needing to be present at the array site itself.

Development of the OSF site will be done by commercial contractors working to ALMA Project specifications. Their principal tasks include:

- Construction of office, laboratory, and residential space totaling 9500 square meters;
- Utility infrastructure including site preparation, water, sewer and power distribution;
- Construction of an antenna assembly hall for the simultaneous erection of two antennas; an additional two antenna pads will be provided adjacent to the assembly hall;
- Fabrication and assembly of a construction camp for the contract workers.

The ALMA Project will be responsible for furnishing the buildings and providing the equipment needed for array operations and maintenance.

Electric power at the OSF either will be generated locally by the ALMA Project using gas turbine generators supplied from the Gas Atacama gas pipeline or power will be brought to the site via overhead lines from a commercial provider. The ALMA Project baseline is local power generation.

16.3.3 The OSF-Array Site Link

Operation and real-time performance diagnosis of the array from the OSF requires a robust communications link between the two. The OSF-Array site link is an optical fiber connection for both voice and data communication. It is a buried fiber that will be provided to ALMA specifications, trenched and buried under competitive commercial contract. The ALMA Project will be directly responsible for the transmitters and receivers on both ends of the fiber and for obtaining the

necessary construction and environmental permits. In the event that the decision on the site access road is made requiring a road directly between the OSF and the array site (without use of the Paso de Jama), that road contract will become part of this task. The link would, in this case, be a real highway as well as a virtual one.

16.3.4 Administrative Office in Santiago

The ALMA Santiago office will provide the focus for the Project Administration, interaction with officials of the Republic of Chile, business and procurement functions, human resource functions, and matters relating to imports and exports. It will also provide a location where ALMA scientists who reside in Santiago but work turno shifts at the OSF can conduct their professional research activities.

Site development for the Santiago office includes the following:

- Construction of a 600 square meter office building;
- Provision of the infrastructure for that building;
- Furnishings and equipment for the building.

The first two items will be done under commercial contract, the latter is the responsibility of the ALMA Project. This task does not include purchase of the land needed for the building. It is anticipated that the building will be done in collaboration with an existing facility in Santiago.

16.4 Environmental Impact

An Environmental Impact Study for the ALMA project has been undertaken on a voluntary basis, in line with the criteria established by the Chilean legislation. The study addresses the seven main requirements from the law
- physical environment, biotic environment, human impact, construction impact, territorial regulations, cultural impact and the aesthetic perception- .

A preliminary study shows that the project development is rated at a low impact level and that the compliance with the Chilean Regulations can be met without any significant cost overhead.

16.5 Site Development Management

All the civil works will be contracted out, either in Chile and/or on an international competitive basis.
Individual contracts will be supervised by a construction team, under the leadership of a site construction manager. The team will establish the contracts according to the project specification and requirements. It will supervise the construction phase, including the quality control, work schedule and the cost developments. It will be based at SpdA. The team will, essentially, be the

interface between the design and the execution phase and handle the acceptance procedures for the site development.

The design for the site works (road, power and communication network, architecture, etc.) will be sub-contracted separately from the civil work contracts and will provide the basis for the construction tendering documents.

Besides the construction manager, the team will have site construction supervisors, an electro-mechanical engineer, a draftsman, a contract officer and secretarial support. Additional technical support will be provided from the U.S./Europe project office in the area of energy, communications and structural analysis.

Construction, Integration, Interim Operations and Schedule

Robert Brown and Richard Kurz

Revision History
2000-Nov-28: First ALMA version

Requirements

- Establish a project and management structure for construction, commissioning, and interim operations of ALMA;
- Construct ALMA according to the tasks and schedule specified in the ALMA Project Work Breakdown Structure (WBS);
- Ensure that the instruments and software fabricated for ALMA are in compliance with the specifications established by the ALMA Project Book and detailed by the Joint Project Office;
- Test and accept subassemblies fabricated by the European and North American Executives. Provide for the integration of those subassemblies into ALMA project subsystems;
- Test, accept, integrate and commission completed project subassemblies at the array site in Chile;
- Provide for interim science operations as the mechanism to ensure that the delivered ALMA product is in keeping with the needs of the scientists who will use ALMA.

17.1 Overview of the Construction-Phase Project Organization and Management

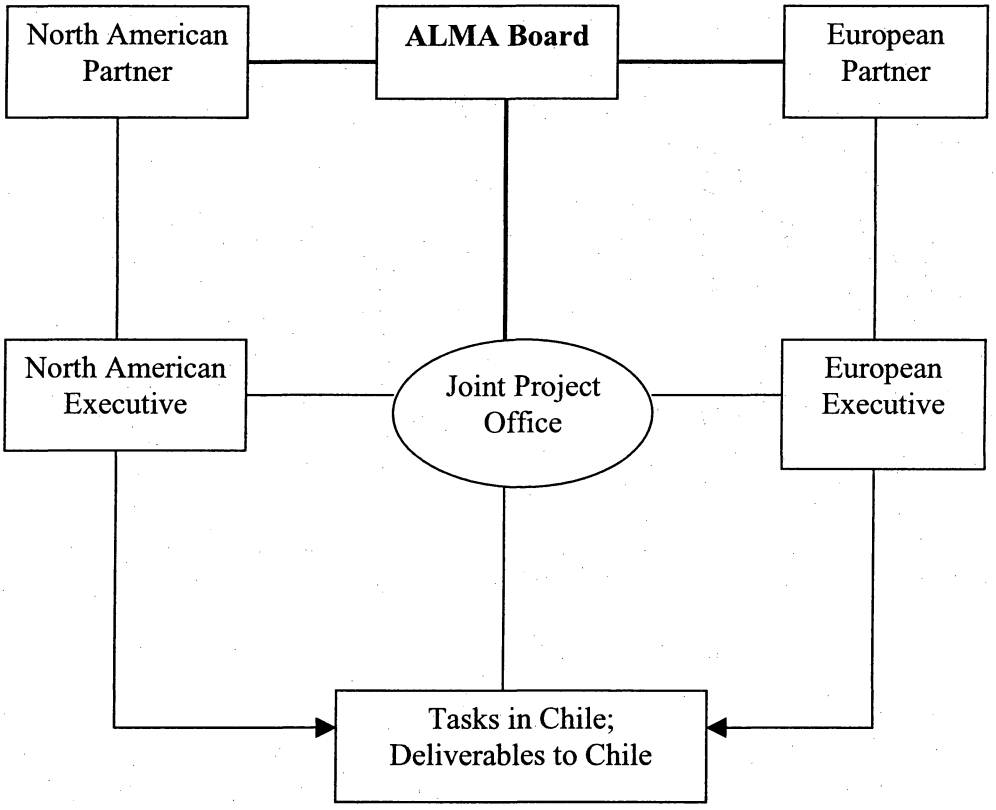
Organization of the management function for the ALMA Project is specified in the ALMA Management Plan. It treats ALMA as a *joint venture* of two Executive bodies that represent the North American and European ALMA partners respectively. It is the desire of both partners that no new legal entities be formed for ALMA; instead all the tasks associated with building, commissioning and operating ALMA will be the functional responsibility of the two Executives. For the European side of the partnership ESO is the Executive; for the North American side of the partnership AUI/NRAO is the Executive. The tasks associated with construction and operation of ALMA will be divided on an equitable basis between the two Executives and thereafter it will be the responsibility of each to perform those tasks, or rovide the deliverables, to the array in Chile.

The ALMA partnership is specified by the ALMA Agreement to be signed by the two partners. The governing board for the ALMA Project, established by the ALMA Agreement, is the ALMA Board whose membership is drawn in equal numbers from the two partners. The coordination of the tasks assigned to the two Executives is the

responsibility of the Joint Project Office. The role of the Joint Project Office is described in detail below. Here it is important to emphasize that the Joint Project Office is not a legal body—it cannot issue contracts, hire employees or establish a bank account. The staff of the Joint Project Office is drawn from the two Executives.

A concept-level organization structure for ALMA is illustrated below. The full organizational structure is described and illustrated in the ALMA Management Plan.

ALMA Construction-Phase Organization



17.2 Construction

The tasks leading to the fabrication of the hardware and software for ALMA are specified in the ALMA Project Work Breakdown Structure (WBS) elements 1.0 – 8.0. These level-1 WBS elements are:

1.0 Administration
2.0 Site Development
3.0 Antennas
4.0 Front Ends
5.0 Local Oscillator
6.0 Backend
7.0 Correlator
8.0 Computing

The responsibility for sub-elements of these level-1 tasks will be shared between the two Executives according to the task division specified in the ALMA Management Plan. A summary of the primary deliverables of each is given below.

Administration: This is a level of effort task that includes the project management at the two Executives and the management of the Joint Project Office. The deliverable here is an effectively managed project that meets its functional objectives, its schedule objectives, and its budgetary objectives.

Site Development: Development of the four sites needed for ALMA in Chile is summarized in Chapter 16 of this Project Book. It includes all the needed civil works (buildings, roads, utilities, antenna stations) on the array site; all the needed civil works at the Operations Support Facility including the equipment needed for integration and commissioning of the delivered hardware and software; the civil works needed to provide a communications link between the array site and the OSF; and the civil works needed to provide administrative and scientific offices in Santiago.

Antennas: The antenna task is specified in Chapter 4 of the Project Book. The antennas will be fabricated by the chosen antenna contractor, or contractors, and will be assembled by that contractor or contractors at the OSF. The responsibility of shipping the antennas to Chile, in pieces or subassemblies as the contractor elects, is the contractor's responsibility. Once erected, acceptance testing will be done at the OSF. Only when the antennas are accepted at the OSF will the ALMA Project take possession. Next, the antennas will be outfitted with ALMA-provided instrumentation (the front ends, cryogenics system, the monitor and control instrumentation, etc.) and the antenna radiometric performance will be verified. At this point they will be transported to the array site and placed on one of the site foundations. Commissioning tests will follow to confirm the mechanical, optical, and radiometric performance. Once the commissioning tests are complete the antenna will be given over to ALMA science operations.

Front-ends: The Front-end task is specified in Chapters 5 and 6 of the Project Book. Owing to the large number of subsystems involved in the Front-end assembly this task will require the greatest amount of coordination and management attention. The task division for the subsystems of the front-end assembly requires a large fraction of the work to be done by each of the Executives, and often done by means of work packages by participating organizations. The ALMA Project baseline plan is to integrate those front-end subassemblies in a hierarchical manner at progressively fewer institutes. This process culminates with the entire front-end assembly being integrated in one or two institutes. At these *front-end integration center(s)* the completed front-ends will be thoroughly tested for compliance with ALMA specifications. The tests will be documented and the front-end assembly will be shipped to the OSF in Chile. Here the important point is that the front-end is shipped from Europe or North America as a completed, functioning and tested unit ready to be installed on an antenna at the OSF.

Local Oscillator: The local oscillator task is specified in Chapter 7 of the ALMA Project Book. The fabrication, testing, and integration of the LO with the front-end assembly is a task entirely done by the Executives in North America and Europe. The LO deliverables are made to the institutes fabricating subsystems for the front-end assembly. These deliveries are shown clearly in the WBS. Once incorporated in the larger subsystems and assemblies the LO subassemblies lose their independent identity. The single exception to this is the central array LO reference which is distributed to each antenna station. The LO reference is fabricated, assembled, and tested in North America and Europe and delivered directly to the ALMA control building on the ALMA site.

Backend: The backend task is specified in Chapter 9 of the ALMA Project Book. The backend subsystem includes all those tasks associated with the down-conversion, digitization and transmission of the signals from the front-end at the antennas to the array control building. Here the instrumentation will be fabricated and tested in North America and Europe and then shipped to the array site in Chile. The final integration and tests must be done in Chile with the optical fiber IF communication system located on site. The site optical fiber is installed under WBS 2.0.

Correlator: The array correlator task is specified in Chapter 10 of the ALMA Project Book. The correlator is built in quadrants, the first quadrant of which will be delivered to the array site in 2004. The ALMA Project baseline is to locate the correlator in the array control building on the array site. The remaining three quadrants will be delivered at the rate of one quadrant every 16 months making the correlator complete and on site in 2007. Here again, the correlator quadrants are delivered to Chile as fully functioning, tested, and documented assemblies.

Computing: The computing task, WBS 8.0, is specified in Chapter 12 of the ALMA Project Book. The task will be accomplished by the common efforts of ALMA computing personnel assigned to this task by the two Executives. It is a joint effort that encompasses the real-time, monitor and control, imaging pipeline, and imaging analysis software that is an integral part of ALMA. At each stage in the construction project the software design and implementation will be reviewed for suitability by future scientist-users of ALMA.

17.3 System Integration

System Integration in Europe and North America: Role of the Test Interferometer: The guiding principle for ALMA construction is to build, assemble and test as much of the ALMA instrumentation as possible at participating ALMA institutes in North America and Europe. The rationale behind this principle is this: at established institutes there are experienced people and resources that can facilitate the ALMA construction tasks. It will minimize risk to the successful technical completion of the task; and it will minimize risk to the schedule because it allows the resource pool to be rapidly increased should the need arise. It is exactly this guiding principle, for example, that is behind the decision to assembly fully the front-end assembly in Europe and North America, and to ship it as complete and functioning units to Chile. But what is to be done with those subsystems

that need the ALMA infrastructure to verify their performance and to assure smooth integration into the system? One answer to this question is the test interferometer.

The test interferometer is an instrument to be assembled from the two ALMA prototype antennas at the ALMA test site adjacent to the VLA in New Mexico. The primary function of the test interferometer is to allow the ALMA staff to do precision tests on the two prototype antennas so that a choice can be made between the two designs. Initially the test interferometer is being built largely from special-purpose interferometer hardware for the antenna evaluation. However, once the antenna evaluation tests are complete, the test interferometer will be retrofit entirely with ALMA prototype instrumentation. This will include all aspects of the hardware except the front-end assembly. Early implementation of the ALMA prototype hardware and software on the test interferometer will permit a thorough system evaluation of those prototypes before production fabrication begins. These tests will assure compliance with instrument specifications as well as with system specifications. And it will do so in a controlled environment with a large resource base of people and test facilities to call upon to solve particularly difficult or subtle problems.

System Integration in Chile: Integration of the production hardware and software will begin in Chile in 2004. Prior to the arrival of the first array antenna, the system integration group will confirm compliance to specifications of the IF transmission system, the LO reference hardware, and the first quadrant of the correlator at the array control building on site. Once the first antenna arrives at the OSF, is accepted and outfitted, the systems integration team will confirm that the interface between the front-end and the antenna is as specified; they will confirm by performance measurements the interface between the software control system and the antenna mechanical servo; and they will verify by radiometric measurements the compliance of the complete antenna system with all performance specifications. At this point the antenna is transported to the ALMA site.

As the first antenna, and each successive antenna, is transported to the site the system integration team will verify all the interfaces with the electrical, data transmission and monitor and control system on site. They will again confirm by radiometric measurements the performance of the antenna and document all their results. At this point responsibility for the antenna is transferred to Array Operations for commissioning observations and interim science observations.

17.4 Interim Operations

There are several reasons to begin interim science operations with ALMA as soon as a few antennas have been accepted, outfitted, and installed on the array site. The following are among the important reasons:

- To use precision astronomical measurements to identify design or implementation deficiencies in the instrumentation or software that should be rectified so that those deficiencies are not replicated multiple times in the hardware still under construction;

- To characterize by careful measurements the affect of different atmospheric conditions on the astronomical imaging performance so as to lead to refinements in the array calibration system or to fine-tune the scheduling algorithms for the prevailing atmosphere;
- To build through experience a model for array operations that is in accord with the desires and expectations of the employees who will be responsible for operations;
- To train and develop a professional team of engineers, scientists, programmers and technicians that will become the ALMA operations staff;
- To expose scientist-users to the growing ALMA so as to receive from them informed advice on development priorities and to train them as the initial cadre of users to whom other prospective users can go for assistance and guidance.

The first few antennas will be on site and available for interim operations by early 2006. In order to make those antennas available for interim science operations it will be necessary to establish the rudiments of an operating observatory also in 2006. This will include:

- An observatory management structure;
- Array operators
- Support scientists, programmers and data analysts;
- Maintenance engineers, programmers and technicians;
- Maintenance support staff (kitchen, housekeeping, maintenance crafts, road maintenance and snow removal, etc.)

Estimates made by M. A. Gordon[1] suggest that an operations staff of 60 people, employees and personnel working under contract, will be needed for interim operations. Given the 5-year overlap between the beginning of interim operations and the end of construction—the period 2006-2010—a thorough operations plan, beginning with interim operations, will be developed in concert with the ALMA construction plan.

17.5 Schedule

The detailed ALMA Project schedule is given in the ALMA WBS. That WBS and schedule has not yet been informed by the funding schedules of the two partners; it awaits a final iteration including commitments by the two partners to provide funding on a particular schedule. In the absence of those commitments, the ALMA baseline schedule drawn from the WBS, at the highest level, is the following.

2002	January	ALMA Construction Project Start
2002	August	Test Interferometer Observations Start
2003	January	Site Civil Works Begin in Chile
2003	April	Prototype Antenna Decision
2003	April	Release Electronics Systems for Production
2003	August	Contract for Production Antennas
2004	October	First Production Antenna Arrives in Chile

[1] ALMA in Chile: A Plan for Operations and Site Construction, May 2000.

2006	January	Interim Operations Start
2010	July	Construction Complete
2010	September	Commissioning; Full Science Operations

POST-CONSTRUCTION OPERATIONS

M. A. Gordon, R. L. Brown
Last changed 2001-Feb-07

Revision History:

1998-10-06	Added Summary
2000-03-28	Revised according to ESO recommendations.
2000-11-30	Irrelevant cost information removed.
2001-02-07	Minor editorial correction [DTE]

SUMMARY

The operating centers for the ALMA will be the instrument itself at the 5,000 m (16,400 ft) site on the Llano de Chajnantor in the Andes mountains near 23° S latitude, the principal operations support facility (OSF) in the nearby village of San Pedro de Atacama at an altitude of 2,450 m (8,040 ft), an administrative support facility in Santiago, and various technical laboratories of ALMA members in the United States and Europe. While a few management personnel should live in San Pedro de Atacama, we expect most of the Chilean support staff will commute from other Chilean communities on a rotating work period basis. Despite the plans described here, we believe that the actual mode of operations will evolve over time as the ALMA staff gains experience operating in Chile.

1. INTRODUCTION

Operating a complex, synthesis radio telescope in Chile will be a new experience. Large optical astronomical observatories have successfully operated in Chile for decades, and the new Very Large Telescope (VLT) is now under construction in Region II. Our operating plan results from the collective experience of these optical observatories and from the experience of synthesis radio telescopes in the United States and Europe.

The plan described below is necessarily tentative. It presumes an operating mode that will take time to perfect. It presumes that ALMA shall be able to find employees willing to live and work in northern Chile, which depends upon the Chilean economy there at the time the ALMA is hiring and in the ambience of the ALMA work environment. To succeed, the ALMA operations management must be analytical, flexible, creative, and willing to build on the experience of the optical observatories. We believe it essential that the ALMA staff in Chile -- that is, the local staff -- should schedule, manage, and operate the radio telescope rather than officials located far away in the United States or Europe.

Chapter 16 above, Site Development, describes the physical plant we believe is necessary to operate the ALMA. These chapters are linked.

2. CONCEPT

The ALMA will operate somewhat like the Very Long Baseline Array (VLBA) headquartered in Socorro, New Mexico. It will be a "service" instrument, observing without astronomers present at the operations center. Astronomers need not travel to Chile to observe, although they may choose to do so. Rather, this observing mode will free them from *having* to travel to the ALMA to observe. In addition, service observing will give the local staff the freedom to juggle observing programs to match the current receiver status and atmospheric transparency. Such a mode requires the ALMA to provide astronomers with the capability to monitor the observing over the Internet, so as to make program changes when necessary.

2.1. Operating Centers

ALMA operations will require three locales in Chile, supplemented by several in the United States and Europe. The instrument itself will be situated on the Llano de Chajnantor, a geologic "bench" at an altitude of 5,000 m (16,500 ft) in the Andes mountains east of the village of San Pedro de Atacama. The operations center will be located near this village because of its proximity and its lower altitude of 2,450 m (8,040 ft). Finally, a small administrative and business office must be located in the capitol of Chile, Santiago, to process duty-free imports, to accomodate high-level administration, to maintain contacts with the national government, and to provide a research environment for the scientific support staff. Sites in the United States and Europe will oversee long-term technical development as well as offer high-level technical support when necessary.

Similar to the Very Large Array (VLA) in New Mexico, the principal operating center of the ALMA may change with time. San Pedro de Atacama is a small village (population 1,000) with few amenities other than those required to support its tourist industry. Few employee's families will want to live there for a long term, especially those with school-age children. As the ALMA evolves into stable operations, we believe it likely that some aspects of its operations will move to a larger community -- probably, Santiago -- with more amenities. Such changes could make long-term employment attractive to skilled professionals. The modern fiber-optic telephone network now being installed in Chile should easily facilitate this relocation. In this case, the San Pedro de Atacama facilities will become principally a maintenance facility similar to the VLA facilities on the Plains of San Augustin in New Mexico.

2.2. Character of Chilean Operations

2.2.1. Management

Management decisions should be local ones. The ALMA director in Chile should make all decisions involving operations in Chile. All employees in Chile should report to the ALMA director, regardless of whether they are "permanent" Chilean hires or ones "borrowed" from related organizations. The sponsoring organizations, the NRAO, ESO, and others, should confine their involvement with the ALMA in an oversight role such as financial support, selection (but not scheduling) of observing proposals, and general policies.

2.2.2. Salaries and benefits

As far as possible, employee salaries and benefits should be "consistent" among all ALMA employees regardless of their place of hire. By the time the ALMA moves into full operation, we expect that Chilean professional salaries will be competitive with the world market. This salary consistency would include the job classifications and the salary steps within them. Exceptions would be temporary employees "borrowed" from other organizations. Because of continuing financial commitments at home, these temporary employees would require larger compensation.. "Benefits" would include medical insurance, pension contributions, educational allowances, housing, and travel allowances where appropriate. Such benefits as well as work rules should be in strict accordance with Chilean law regardless of the eventual diplomatic or international status of the ALMA organization.

2.2.3. Contracting support services

As is customary in the Chilean mining industry, the ALMA should contract for commercial services when they are available. For example, Chile has several large companies that provide food service to remote locations. The employer need only supply a kitchen and dining room, and specify the variety and quality of the food to be served. This situation also applies to medical

services, housekeeping services, payroll, security, and vehicle leasing and maintenance. The ALMA should actually hire only those employees unavailable or inappropriate to obtain from commercial service companies, such as management and administrative personnel, support scientists, engineers, programmers, and telescope mechanics. Not only is this system flexible and cost-effective through competition, but it also frees the ALMA management to expand or contract services as needed without impacting long-term ALMA employees.

3. STAFFING

3.1. Sistema de Turno employment for the ALMA and its Operations Center

To operate the ALMA in Chile, *all* consultants recommend a rotating shift system known in Chile as the "Sistema de Turno" for staffing the operations center and the maintenance of the ALMA itself. In Chile the Turno system is used by all international observatories and most mining operations. It complies with Chilean labor laws.

Turno work arrangements include a range of schedules. Variations are common. A construction project in a remote area east of Iquique operates on a two week "on" and a 10 days "off" system. ESO uses two schedules: known as "5/2" and "8/6." The "5/2" schedule is appropriate for office staff. The work begins at 3PM on Monday, consists of 9.5-hour days Tuesday through Thursday, and ends at 1PM on Friday. The "8/6" schedule is more appropriate for skills needed every day. It provides approximately 88 work hours over a two week period. It begins at 3PM on day 1, consists of 9.5 hours each on days 2 through 7, and ends at 1PM on day 8. Sunday is compensated at 1.75 x the basic rate. Replacement personnel overlap on days 1 and 8. Customarily, the employer provides room, board and transportation to and from an urban assembly point.

An effective Turno system must be appropriate to the specific operation of the ALMA. This system is not appropriate for highest level management people who need to be continually available. It is also inappropriate for employees responsible for creating new systems or equipment. However, it works well for many office positions and for "interchangeable" personnel like telescope operators and maintenance people who must be available seven days a week, 24 hours a day. There is extra cost involved. ESO statistics show that the "8/6" arrangement requires about 2.4 employees for every Turno position to ensure overlap and continuity.

Given the difficulties of staffing a location like San Pedro de Atacama, the Turno system may prove to be the only practical solution.

To accommodate a Turno system, the ALMA would need to provide dormitories at its operations center near San Pedro de Atacama. Our advisors recommend that the dormitories be sized so that each Turno-employee could have the same room and the same bed each visit. In this way, that employee could leave personal effects in the room and could decorate the room to suit his or her preferences.

The ALMA should establish pickup points for Turno employees only in Calama and Antofagasta, at first. Region II has a network of modern, commercial buses linking its cities. Some of these buses serve San Pedro de Atacama more than once daily and, of these, a few continue on the Paso de Jama road into Argentina. The principle would be that commuting employees need to get themselves to the collection points by the most appropriate means and at their own expense.

Professional employees would either live in San Pedro and take substantial holidays as compensation for long hours on the job, or commute from elsewhere in Chile with some of the commuting time being considered working time.

3.2. Support offices in Santiago

The Santiago offices would not require a Turno system, nor would one be appropriate to their function. Because of the added expense and inefficiency of commuting to and from San Pedro de Atacama, all positions not required in SpdA should be located in Santiago. These would include certain senior administrative management, contract and procurement personnel, accounting functions, and an import/export office.

Effective supply for ALMA will require an import/export office in Santiago. The Chilean Foreign Ministry (Ministerio de Relaciones Exteriores) in Santiago must process all papers for duty-free imports regardless of which Chilean port is used. Furthermore, even after 30 years of operating in Chile from La Serena, the CTIO has chosen to use only ports of entry near Santiago even though the city of La Serena is contiguous with the port of Coquimbo. The CTIO has found that the high traffic levels at the Valparaiso seaport and the Santiago International Airport give the widest opportunities for shipping. Equally important, they have found that, in most cases, these ports are less expensive to use than the port of Coquimbo even though the Santiago goods must be trucked to and from La Serena. The same situation may apply to Antofagasta – the closest port to ALMA.

The Santiago facilities would also include offices for support astronomers. ESO experience has shown that their support scientists had little success conducting research while at the ESO observatory at La Silla. Yet, opportunities to conduct research is essential to retaining quality astronomers. Therefore, ESO instituted schedules for their support astronomers of a week at La Silla, a week in Santiago, and a week off – or some variation on this theme. This system has worked well, and ALMA should adopt it.

The roles of the Santiago office may change with time. As ALMA operations mature, it may be possible to withdraw all but essential operations support from San Pedro de Atacama to Santiago. Specialized support staff would then fly to SpdA (Calama) when needed. This could save money without impacting the reliability of ALMA.