ALMA Test Interferometer Raw Data Format

Robert Lucas and Brian Glendenning

October 8, 2000

$Id: AlmaTIFormat.tex,v 1.7 2000/10/08 11:16:20 lucas Exp $

Abstract

This draft document is an attempt to define the raw data format for the ALMA Test Interferometer.

Contents

1 Introduction

This draft document is an attempt to define the raw data format for the ALMA Test Interferometer. This is nevertheless done while keeping in mind that the format might be later extended to be a candidate for the ALMA raw data format.

Another consideration was that the raw data will be converted to be input to Plateau de Bure Calibration software (CLIC, CLASS) to be reused for the test interferometer.

2 Some definitions

Extracted from the ALMA Software Glossary:

dump The smallest interval of time for which a set of correlated data can be accumulated and output from the correlator.

integration A set of dumps, all identical in configuration (except for the antenna motion and some others), that is accumulated and forms the basic recorded unit.

observation A set of integrations while the antennas complete an elemental pattern across the source, and possibly while frequency switching, nutator switching, etc.

scan A set of observations with a common goal, for example, a pointing scan, a focus scan, or an atmospheric amplitude calibration scan.

For instance in the case of holography measurements an observation would be a drift across the transmitter or a boresight measurement, while a scan could be the whole set of observations needed to acquire a beam map. Or a scan could be a pointing scan with two observations (an azimuth drift and an elevation drift across the pointing calibrator) or an atmospheric calibration scan with three observations (autocorrelations on the sky, and two loads at differenttemperatures, ...).

3 FITS ALMA-TI File Structure

A FITS ALMA-TI File shall contain:

1. One primary header that will identify the file as being of the ALMA-TI file format.
2. For each observation one will find:
   - One or more binary tables for data associated parameters `DATAPAR`. This table will contain parameters that vary with time during the data set (except the data itself, contained in `AUTODATA` and/or `CORRDATA` tables. An new `DATAPAR` table is written for each observation.
   - Zero or mode binary tables for calibration parameters: `CALIBRATION`
   - Zero or more binary tables for data values (autocorrelation): `AUTODATA`
   - Zero or mode binary tables for data values (correlation): `CORRDATA`
   - Zero or mode binary tables for data values (single dish holography): `HOLODATA`
3. Other binary tables for header parameters that are valid through the whole data set time range. These are not repeated for each observation.

The use of binary tables introduces some flexibility by allowing columns to contain arrays. We make use of the “Multidimensional Array” convention in Appendix B of Cotton et al (1995).

Also some binary tables may be reused from the FITS-IDI format.

All tables have an `EXTVER` keyword so that software may support format changes.

The tables shall be written in the following order: `ANTENNA`, `FREQUENCY`, ... , `DATAPAR`, `CALIBRATION`, `CORRDATA`, `AUTODATA`.

All table `EXTNAME`s are internally prefixed in the FITS files by ‘ALMATI-’. That prefix is omitted in text for clarity.

The general header and the table headers shall include all keywords needed to conform the the FITS standards; for clarity reasons not all of them are listed here.

4 The Primary header

Keyword	Type	Value	Description
NAXIS	I	0	
SIMPLE	L	T	
EXTEND	L	T	
GROUPS	L	T	
GCOUNT	I	0	
PCOUNT	I	0	
ORIGIN	A		Software (including version)

Notes:

- Add a keyword here to make sure the format is uniquely identified ?
- ORIGIN contains the name and version of the software that produced the FITS file.

5 The DATAPAR Binary Table

The DATAPAR binary table contain most header information. Parameters that are valid for the whole data set appear as header keywords, while parameters or arrays that change with time (every data integration) are in table columns. There is a new DATAPAR binary table for each observation. The table has one line per integration point. Each integration point in the file is referred to by a unique number INTEGNUM which appears in the DATAPAR table and in the data tables.

5.1 DATAPAR Binary Table Header Keywords

name	Type	Units	Value	Description
EXTNAME	A		'ALMATI-DATAPAR'	
EXTVER	J		1	
TELESCOP	12A	-	'ALMATI-INTER'	Telescope Name
	12A	-	'ALMATI-VERTX'	
	12A	-	'ALMATI-EIE'	
SCAN-NUM	J	-		Scan number
OBS-NUM	J	-		Observation number
DATE-OBS	D	-		observing date
UT	D	s		Universal time (start)
LST	D	s		Sidereal time (start)
OBSMODE	4A	-		Scan type
PROJID	4A	-		Project
AZIMUTH	D	deg		Azimuth at start of observation
ELEVATIO	D	deg		Elevation at start of observation
LATSYS	8A	-		type of latitude-like offsets
LONGSYS	8A	-		type of longitude-like offsets
OBSTIME	E	s		Total integration time
SOURCEID	J	-		Source number
NO_ANT	J	-	$N_{\rm A}$	number of antennas
NO_BAND	J	-	$N_{\rm BD}$	number of basebands
NO_STK	J		$N_{\rm ST}$	number of pols.
NO_SIDE	J		$N_{\rm SB}$	number of sidebands
VFRAME	D	m/s		radial vel. correction
SITELONG	D	deg		site longitude
SITELAT	D	deg		site latitude
SITEELEV	E	m		site elevation

Notes:

- LATSYS and LONGSYS describe the meaning of the LATOFF and LONGOFF columns in the table. They take values like 'RA—SIN', 'DEC–SIN', 'AZIM-SIN', 'ELEV-SIN' (we have the goal of following the World Coordinate System (WCS) convention).
- VFRAME radial projection of $V_{\rm FRAME} - V_{\rm OBS}$
- SOURCEID=0 means we are observing at fixed AZIMUTH, ELEVATIO (plus possible varying offsets); or introduce new keywords in the SOURCE table for horizontal coordinates (like the holography transmitter).
- Foreseen OBSMODE values:

POIN	Pointing measurement
FOCU	Focus measurement
CALI	Autocorrelation on load
AUTO	Autocorrelation on sky
CORR	Correlation on sky
SKYD	Skydip
HOLO	Holography
ONOF	Total power ON-OFF
FSWI	Frequency switch
...	...

– `SITELONG`, `SITELAT` and `SITEELEV` gives informatively the site coordinates in a geocentric system. They should correspond to the `ARRAYX`, `ARRAYY`, and `ARRAYZ` coordinates in the `ARRAY_GEOMETRY` table which define the origin of the `UUVVWW` vectors for each antenna.

5.2 `DATAPAR` Binary Table Columns

name	Type	Units	Description
`INTEGNUM`	1J	-	Integration point number
`INTTIM`	1E	s	Integration time
`DATE`	1D	days	Julian Date at 0h
`TIME`	1D	days	time since 0h
`UUVVWW`	D(3,$N_{\rm A}$)	s	u,v,w ant. coords.
`AZELERR`	E(2,$N_{\rm A}$)	deg	Az,El point err
`SOURDIR`	D(3,$N_{\rm A}$)	-	Source direction cos.
`DELOFF`	E($N_{\rm A}$)	s	Delay offset
`FOCSOFF`	E($N_{\rm A}$)	m	Focus offset
`LATOFF`	E($N_{\rm A}$)	deg	lat.-like offset
`LONGOFF`	E($N_{\rm A}$)	deg	long.-like offset
`TOTPOW`	E($N_{\rm BD}, N_{\rm A}$)	V	m basebands, Total Pow.
`WINDSPEE`	1E	m/s	Wind speed
`WINDDIRE`	1E	deg.	Wind direction (E from N)
`FLAG`	J($N_{\rm ST}, N_{\rm BD}, N_{\rm A}$)	-	Flag words

Notes:

– `INTEGNUM` refers to corresponding integration periods in the data tables.

– `SCAN-NUM` and `OBS-NUM` uniquely identify a given `DATAPAR` table in the file. `OBS-NUM` is needed to identify separate observations within a scan. `SCAN-NUM` sequentially increases with time but may be recycled (e.g. *modulo* 10000). These two numbers are repeated in the data table headers to avoid confuion.

– All table columns refer to integration midpoint.

– `LATOFF` and `LONGOFF` give the antenna pointing direction relative to the source direction as defined in the `SOURCE` table.

– `SOURDIR` gives the phase tracking center direction in the geocentric system of rectangular coordinates (the one used to define the baseline coordinates).

– *to be added here:* column for delays actually applied in hardware and software to each baseband, and a description of their variation with time during integration.

– `FLAG` is an array of flag 32-bit words, one for each combination: antenna, baseband, polarisation product. Each bit is assigned a different meaning:

· 32: DATA (reserved for data reduction)
· 31: POINTING (tracking errors too large)
· 30: TSYS (system noise too high)

· 29: LOCK (phase-lock loop open)
· 28: SHADOW (antenna was shadowed)
· 27: SATURATION (to much signal ot detectors)
· 26: TIME (was wrong)

6 The CALIBRATION Binary Table

There is one CALIBRATION binary table for every scan/observation. It includes the calibration parameters used / to be used for amplitude calibration of the data in that scan/observation.

6.1 The CALIBRATION Binary Table Keywords

name	Type	Units	Value	Description
EXTNAME	A		'ALMATI-CALIBRATION'	
EXTVER	J		1	
CALMODE	12A	-		Calibration mode
NO_BAND	J		N_{BD}	number of basebands
NO_STK	J		N_{ST}	number of pols.
NO_SIDE	J		N_{SB}	number of sidebands
SCAN-NUM	J	-		Scan number
OBS-NUM	J	-		Observation number

6.2 The CALIBRATION Binary Table Column List

name	Type	Units	Description
ANTENNID	1J	-	Antenna number
FREQID	1J	-	Frequency setup number
APEREFF	E(N_{ST})	-	Aperture efficiency
BEAMEFF	E(N_{ST})	-	Beam efficiency
ETAFSS	E(N_{ST})	-	Forward efficiency
ANTGAIN	E(N_{ST})	K/Jy	Antenna Gain
HUMIDITY	E	-	rel humidity (0.-1.)
TAMBIENT	E	K	Ambient temperature
PRESSURE	E	hPa	Ambient pressure
THOT	1E	K	Chopper temperature
TCOLD	1E	K	Cold Load temperature
PHOT	E(N_{ST})	V	Total power on Chopper
PCOLD	E(N_{ST})	V	Total power on Cold Load
PSKY	E(N_{ST})	V	Total power on Sky
GAINIMAG	E(N_{ST})	-	Gain ratio image/signal
TRX	E(N_{ST})	K	Receiver temperature
TSYS	E(N_{ST})	K	System temperature
TSYSIMAG	E(N_{ST})	K	System temperature (Image)
TAU	E(N_{ST})	-	Opacity
TAUIMAG	E(N_{ST})	-	Opacity (Image)
TCABIN	1E	K	Receiver Cabin temp.
TDEWAR	1E	K	Receiver Dewar temp.

Notes:

- There is one table line per antenna.
- CALMODE specifies the atmospheric calibration mode that has been applied in quasi real time (if any). Under different modes, some parameters are actually measured, others are determined from them, e.g. using a model atmosphere. The atmospheric calibration modes for ALMA are TBD.
- Some or all of the parameters may have to be differenciated by baseband too.

7 The CORRDATA Binary Table

There is one CORRDATA table for each baseband. As a difference with FITS-IDI, the basebands have different number of channels.

7.1 CORRDATA Table Header Keywords

name	Type	Units	Value	Description
EXTNAME	A		'ALMATI-CORRDATA'	
EXTVER	J		1	
NO_CHAN	J		$N_{\rm CH}$	number of channels in baseband
NO_STKD	J		$N_{\rm ST}$	number of pol. products
NO_SIDE	J		$N_{\rm SB}$	number of sidebands
SCAN-NUM	J	-		Scan number
OBS-NUM	J	-		Observation number
BASEBAND	J	-		Baseband number

7.2 CORRDATA Binary Table Columns

name	Type	Units	Description
INTEGNUM	I	-	Integration point number
STARTANT	I	-	Start Antenna
ENDANTEN	I	-	End Antenna
DATA	E	-	Data

Notes:

- INTEGNUM refers to corresponding integration period in the data parameter table.
- TDIM for the DATA column is '(2,$N_{\rm CH}$,$N_{\rm SB}$,$N_{\rm ST}$)'
- $N_{\rm SB}$ is 1 unless phase switching is used to separate sidebands from the same baseband.
- A sample CORRDATA header is given in Appendix.
- Units should be relative correlation (1=100% correlated)

8 The AUTODATA Binary Table

There is one AUTODATA table for each baseband. The basebands may have different number of channels (there is only one baseband for the test interferometer).

8.1 AUTODATA Binary Table Header Keywords

Name	Type	Units	Value	Description
EXTNAME	A		'ALMATI-AUTODATA'	
EXTVER	J		1	
NO_CHAN	J		$N_{\rm CH}$	number of channels in baseband
NO_STKD	J		$N_{\rm ST}$	number of pol. products
SCAN-NUM	J	-		Scan number
OBS-NUM	J	-		Observation number
BASEBAND	J	-		Baseband number

8.2 AUTODATA Binary Table Columns

Name	Type	Units	Description
INTEGNUM	I	-	Integration point number
ANTENNA	I	-	Antenna
DATA	E	-	Data

Notes:

- INTEGNUM refers to corresponding integration period in the data parameter table.
- TDIM for the DATA column is '($N_{\rm ST}$,$N_{\rm CH}$)'
- Units: TBD. Fraction of system temperature ?

9 The HOLODATA Binary Table

There is one HOLODATA table for each baseband. The basebands may have different number of channels (there is only one baseband for the test interferometer).

9.1 HOLODATA Binary Table Header Keywords

Name	Type	Units	Value	Description
EXTNAME	A	-	'ALMATI-HOLODATA'	
EXTVER	J	-	1	
NO_CHAN	J	-	$N_{\rm CH}$=1	number of channels
NO_STKD	J	-	$N_{\rm ST}$=1	number of pol. products
SCAN-NUM	J	-		Scan number
OBS-NUM	J	-		Observation number
BASEBAND	J	-	1	Baseband number
TRANDIST	E	m		transmitter distance
TRANFREQ	D	Hz		transmitter frequency
TRANFOCU	D	m		offset from prime focus

9.2 HOLODATA Binary Table Columns

Name	Type	Units	Description
INTEGNUM	I	-	Integration point number
ANTENNA	I	-	Antenna
DATA	E	-	Data

Notes:

- INTEGNUM refers to corresponding integration period in the data parameter table.
- TDIM for the DATA column is '(6,$N_{\rm ST}$,$N_{\rm CH}$)' (=6) since 6 receiver data products are recorded (s*s, r*r, q*q, s*r, s*q, q*r).

10 The SOURCE Binary Table

10.1 The SOURCE Binary Table Keywords

name	Type	Units	Value	Description
EXTNAME	A	-	'ALMATI-SOURCE'	
EXTVER	I	-	1	
NO_BAND	J	-	$N_{\rm BD}$	number of basebands
NO_GEO	J	-	$N_{\rm GEO}$	number of ephemeris values

10.2 The SOURCE Binary Table Columns

name	Type(Dim)	Units	Description
SOURCEID	1J		Source number
SOURCE	12A		Source name
CALCODE	4A		Calibrator code (OBJ,PHA,...)
IFLUX	E($N_{\rm BD}$)	Jy	flux
QFLUX	E($N_{\rm BD}$)	Jy	
UFLUX	E($N_{\rm BD}$)	Jy	
VFLUX	E($N_{\rm BD}$)	Jy	
RAEPO	1D	deg	RA at mean equinox
DECEPO	1D	deg	DEC at mean equinox
PMRAEPO	1D	deg	RA proper motion at mean equinox
PMDECEPO	1D	deg	DEC proper motion at mean equinox
EQUINOX	8A		mean equinox
SYSVEL	D($N_{\rm BD}$)	m/s	systemic velocity
VELTYP	8A		e.g. 'VELO-LSR'
MJDGEO	1D	s	MJ date
UTGEO	D($N_{\rm GEO}$)	s	times
RAGEO	D($N_{\rm GEO}$)	s	RAs
DECGEO	D($N_{\rm GEO}$)	s	DECs
VELGEO	D($N_{\rm GEO}$)	s	radial velocities

Notes:

- CALCODE indicates if the source is a target OBJECT, a PHASE calibrator, ...
- We use NaN for unknown fluxes.
- NO_GEO is non-zero if we define the source by its geocentric ephemeris (times, positions, radial velocities). Then the columns referring to EQUINOX are not present.
- *Actual ephemeris keywords and columns to match the input catalog format.*

11 The ANTENNA Binary Table

This table should mainly contain the pointing model coefficients for the antenna. The exact format (keyword names) is still TBD.

12 The INTERFEROMETER_MODEL Binary Table

See FITS-IDI

13 The ARRAY_GEOMETRY Binary Table

13.1 The ARRAY_GEOMETRY Binary Table Keywords

[see FITS-IDI description pp 37-38]

13.2 The ARRAY_GEOMETRY Binary Table Column list

name	Type	Units	Description
ANTENNID	J		Antenna number
ANNAME	8A		Antenna name
STABXYZ	3D	m	Coordinates
STAXOF	1D	m	Axis offset
STANAME	4A		Station name

14 The FREQUENCY Binary Table

Use FREQ-OBS as FITS built-in axis in DATA tables if Doppler effect was not tracked, VELO-LSR if the Doppler effect was tracked (either in hardware or in software)

14.1 FREQUENCY Binary Table Keywords

name	Type	Units	Value	Description
EXTNAME	A		'ALMATI-FREQUENCY'	
EXTVER	J		1	
FREQID	J	-		Frequency setup number
NO_BAND	J	-	N_{BD}	number of basebands

14.2 The FREQUENCY Binary Table Columns

name	Type	Units	Value	Description
FREQID	1J	-		Frequency setup number
NO_BAND	1J	-	N_{BD}	number of basebands (1-4)
FREQLO1	D(N_{BD})	Hz		LO1 Frequency
SIDEBLO1	J(N_{BD})	-		side band LO1
FREQLO2	D(N_{BD})	Hz		LO2 Frequency
SIDEBLO2	J(N_{BD})	-		side band LO2
...	...	...		(as needed)
CHANNELS	J(N_{BD})	-		number of channels
IF_REF	R(N_{BD})	-		ref. channel
IF_VAL	D(N_{BD})	-		IF frequency of ref. channels
IF_INC	D(N_{BD})	-		IF channel increment
FREQRES	R(N_{BD})	-		Frequency resolution
RESTFREQ	D(N_{BD})	-		Rest frequency
TRANSITI	12A	-		line identifier
RECOFF	2D	deg.		beam offsets from boresight (AZ,EL)

Notes:

– FITS-IDI assumes all frequency bands have the same number of channels. This may be wrong for ALMA: 8 basebands may have different number of channels.