

# Software Review

## ALMA Test Interferometer Raw Data Format

2001-06-25

This is to announce that the document: *ALMA Test Interferometer Raw Data Format* by R.Lucas and B.Glendenning, (Rev.1, 2001-04-26) has been reviewed successfully on June 25th.

The Telecon part of the review was done on June 25th (Monday) at 15.h UT and attended by:

E.Greisen  
P.Grosbol  
A.Kemball  
K.Nakanishi  
G.Raffi (Chair)  
P.Schilke

B.Glendenning (Author)  
R.Lucas (Author)

### Conclusion:

The document got 66 written comments. All of them received written replies. The most relevant and controversial ones were discussed at the telecon meeting. The document is approved and will be edited according to the final set of comment/replies (see below for procedure steps).

### ***Minutes of telecon discussion:***

The comments discussed were extracted from the list prepared by R.Lucas on the main comments, grouped by subject.

- *scope of the data format: 1, 6b, 7*

6b: a more careful separation of data will be needed for the final ALMA format.

7 : It is important to give (at least one) real science example based on real science data, even if the purpose of the TI is not to do science.

It will show that science data can be taken with this format and with the prototype antennas.

- *format of document itself: 5, 9, 59*

59: reply is OK

- *data organization: 10, 12, 13, 13b, 14, 15, 17, 18, 20, 21, 23, 24, 24b, 24c, 32, 38, 44, 45b, 46, 47, 50b, 56*

10: There is a heap of data. A better format is recommended as final format. This should have a link to the SBs.

12: Format to be improved for final ALMA format.

13b: Reply is OK

20,24, 24 c: Date, file name in file header. Structure of the file should

instead NOT be part of the header, as it could be wrong. After discussion there is

agreement with the reply to comment 24.

45b: There is a historical constraint on this format for the TI, to be able to use PdB software. This should be removed and the format changed for the final ALMA format.

46: The mixing of non-time dependent variables with time-dependent ones is no good.

==> There is an action on Robert and Brian to rediscuss this part, having become aware of the problem and to come up with a solution to this problem. The Board entrusts the authors to come up with a suitable solution, which for reasons of time could not be discussed in detail at the meeting.

- *table names, keyword names:* 8, 11, 16, 19, 22, 25, 7, 27, 28, 29,  
30, 33, 34, 34b, 35, 36, 37, 39, 40, 50c, 51, 54, 57, 60, 60b,  
61, 65, 66

11: It is agreed that the table names should be inverted, like e.g. in DATAPAR-ALMATI, putting ALMATI last.

34: Reply is OK for TI data

34b: Absolute coordinates will be used for moving objects

39: fine for TI data

33,40: Time will be given in secs (Modified Julian Date), as proposed in 33.

51: replies are OK

54: It is safe to go on with latest Draft WCS convention for FITS format. So one should conform to it, even if it is a Draft.

60b,61: The order should be changed, assuming there is a way to do it with FITSIO. A report should be sent to FITSIO to this effect.

16: EXTVER should be corrected, as it is used with the wrong meaning.

- *data contents, units:* 43, 45, 48, 49, 50, 52, 53, 55, 58, 62, 63,  
64

58: Reply is OK

62: TELESCOPE to be checked against spec's.

43: There should be an AtoD converter unit in counts in the FITS spec.

***Procedure for review:***

- written comments by reviewers (and others who might wish to do so)
- replies in writing by author(s)
- telecon review meeting (with discussion of controversial replies).
- minutes/editing of replies
- final editing of document
- document on Web (as officially reviewed Joint Memo)

## *Comments with Final Replies*

---001-----  
(egreisen)

Comments on "ALMA Test Interferometer Raw Data Format"

Let me begin by stating that I know little about ALMA and that I have also not had the time to delve deeply into some of the issues raised by this document.

1. My first and probably most significant impression is that the format described is likely to be quite useful for the Test Interferometer, but that one would not want to use it as the science data format given to scientist users of ALMA. Those users are not interested in the nth LO frequency, the 1000 monitor points, etc. Time, u,v,w, baseline, pre-calibrated data and source info are about all that the scientists want and need from ALMA. The real-time calibration/flagging system and technicians, including staff scientists checking and improving the instrument, will need everything nearly always with the TI and sometimes with ALMA. But not the scientists.

Reply:

The described data format aimed at the Test Interferometer. In the more general frame of ALMA it could evolve in a raw data format, that is a format not directly oriented to the end-user but to observatory-supported software that would read it, calibrate it, and produce science data (images). Monitor points are only needed here to be able to plot science data together with monitor data for antenna evaluation purposes.

---002-----  
(fviallefond)

Questions: 1) How many bits for visibilities

- 2) For phase corrected and uncorrected data where would you put corrected /uncorrected data? Which keywords to specify it?

Reply:

- 1) 32 bit for each component

- 2) not described here for the TI since it is not needed, it could be included as a fifth dimension in the visibility data set.

---003-----  
(graffi) dummy comment number - no General Comments

---004-----  
(pgrosbol) dummy comment number - no General Comments

---005-----  
(pschilke)

General comments:

I couldn't find a bibliography section, on p.6 for example there is a

reference to Cotton et al. (1995), which isn't resolved.

Reply:

We'll add a References Section.

---006-----  
(sichikawa) dummy comment number - no General Comments  
---006b-----  
(akemball)

I think this data format is fine for the test interferometer. If it can be used by CLIC then it is fine for the ALMA AIPS++ tests also, as far as I am concerned. I would have reservations about using this format for the full deployed version of ALMA later however, but I understand that the review at present is not in that spirit.

Reply:

None needed.

---007-----  
(egreisen)  
p.5

8. Section 2 "definitions" includes several observing examples none of which include taking scientific data. I began to fear that you did not plan to do so. The OBSMODE list on page 8 suggests otherwise, but it would be good to include real science examples.

Reply:

The TI is not aimed at taking true scientific data, only to testing the antennas. We will nevertheless give other examples.

---008-----  
(graffi)  
p.5 prefix ALMATI-

Is the use of this prefix necessary all over the place? This embeddes the name TI along with the data structures, but it could probably be avoided. If one could use instead a more stable name, e.g. just DATAPAR, AUTODATA etc and confine the ALMATI to some Primary Header Keyword, this would still characterize data as TI data.

Reply:

The prefix was inserted to uniquely identify those extensions from other with similar names e.g. in FITS-IDI.

---009-----  
(pgrosbol)  
p.5, s.1: It would be appropriate to list 'NOST 100-2.0' as an applicable document.

Reply:

Sure.

---010-----  
(pgrosbol)  
p.5, s.2: 'observation' / 'scan'

The definitions are not fully clear as they rely on 'an elemental pattern' and 'common goal'. This makes it difficult to define the relation between observations/scans and physical raw data files. For archival research and sharing of calibrations, it would be better to have a more clear relation between data files and individual target.

Reply:

We intend to put several scans -- and observations -- in a raw data file.

---011-----  
(pgrosbol)  
p.5, s.3.2a: 'ALMATI-DATAPAR' ...

The EXTNAME always has the prefix 'ALMATI-' which is redundant as the keywords ORIGIN and TELESCOPE defines the origin. The prefix should be omitted.

Reply:

See 008. May be you are right. However our thinking was that more "standard" existing table extensions may migrate to this format, and moreover we may want to make some changes for ALMA. Fully qualifying the name makes it unambiguous.

As a result of discussion we decided to exchange the order:  
e.g. 'DATAPAR-ALMATI', and so on.

---012-----  
(pschilke)  
p. 5

### 3: FITS ALMA-TI File Structure

- This is maybe not yet relevant for the test interferometer, but shouldn't there be two ALMATI-CORRDAT tables for the data with and without phase correction?

Reply:

See 002. Data with and without phase correction could define an extra dimension in the data array.

---013-----  
(sichikawa)  
p 5. and 6.

1) (pp.5-6)

I could not understand how many binary tables one FITS file contains. Does it contain more than one binary tables which have the same 'EXTNAME' ?

Reply:

Sorry this should be more explicit. The answer is "Yes". Some binary tables are present for each observation.

---013b-----

(knakanishi)

p. 5

1) (pp. 5, Section 3)

I suppose too many binary tables may be contained in one FITS file when a set of binary tables is created for each observation. What is the advantage from creating a set of binary tables for each observation?

I think it is also reasonable to create a set of binary tables for each scan (not for each observation). Is this style also allowable?

Reply:

We have tried to have each individual observation described in a complete way. This is at the price of some repetition (source information, antenna positions, ...) but it saves some complexity in the writing and reading software. Thus the scan based parameters (like OBSMODE, object positions, ...) will be repeated for each observation in the scan.

---014-----

(graffi)

p.6 Primary header

Is TELESCOP where the TI array gets defined? If not one could have a name like ARRAY="ALMATI", but I suspect that ARRAY might not be any FITS standard name.

Is TELESCOP a list of variable length? I wonder if the case of any two antennas out of three is covered here, like Vertex vs.EIE or Vertex vs. Japanese.

Reply:

We may build a mnemonic name other than INTER for those cases now that there will be three prototypes. But see 015.

---015-----

(graffi)

p.6 TELESCOP

For the single antennas it would be good to have a number, a type and an antenna foundation position number. While less relevant

for TI, one could say that antenna 1 is type Vertex and on foundation 1, antenna 2 is type EIE and 3 is type Mitsubishi. The case of movement of antennas we should cover, as movement will be tested. So antenna 1 could go to foundation 2 or 3.

Reply:

We have ANTENAME and STATNAME columns in ALMATI-ANTENNA table. We just included TELESCOP here as an aid to quick identification if all the data in the file belong to one and only antenna, or if this is only interferometric data.

---016-----  
(pgrosbol)  
p.6, s.3: EXTVER

The FITS standard defines another meaning for this keyword. Thus, the software version should be given by another one.

Reply:

After checking: you're right. We now use TABLEREV for the iterations in table format definitions (if needed).

---017-----  
(pgrosbol)  
p.6, s.3: OBS-NUM and SCAN-NUM

These numbers are unique within a file but not if FITS extensions are removed from the original file (which is trivial to do). It would be safer to have a unique identification within each extension header e.g. including DATA-OBS, ORIGIN, TELESCOPE or SB-ID.

Reply:

These number are not unique within a file, but each table extension refers uniquely to a DATAPAR table by those numbers, and each DATAPAR in a file is uniquely identified in a file by these two numbers. So yes, we could repeat DATE-OBS in all extensions to protect us against extensions wandering from file to file.

---018-----  
(pgrosbol)  
p.6, s.4: GROUPS, GCOUNT, PCOUNT

This defines the file as a random groups file which is not the case. Thus they should be removed.

Reply:

I thought setting them to zero was OK; I took that from FITS-IDI.

---019-----  
(pgrosbol)  
p.6, s.4: ORIGIN, TELESCOP

The ORIGIN keywords should specify the organization or institute which generated the file and not the software. Thus, it would be better to have e.g. ORIGIN = 'ALMA-TI' and TELESCOP = 'VERTX'

The software identification and version should be given in other keywords.

Reply:

You're right. Any standard for software?

---020-----  
(pgrosbol)  
p.6, s.4: Information in prime header

The prime header is close to the minimum. The VLT experience has shown it to be convenient to have somewhat more information in the prime header as it is easy to find it even for simple application. Also when calibrations are in several physical file, it is easier to find associations between science and calibration files if the appropriate information is in the prime header. This may be items such as bands used, array configurations, time, program and SB id's etc.

Reply:

We could have some comment information but if they are entered by an operator one has to make sure that they are entered in a uniform way.  
See reply to comment 24 below.

---021-----  
(pgrosbol)  
p.6, s.5: 'whole observation'

Other places in the document it is suggested that a file can contain one or more scans which each consists of many observations. The term 'whole observation' seems not to refer to the definition in Sec.2 and should therefore be re-phrased.

It would be good to have a clear definition of the relation between a single raw file and observations/scans it can contain.

Reply:

It does refer to the definition. There is one and only one DATAPAR for each observation. We will clarify the relation between files, scans and observations.

---022-----  
(pschilke)  
p. 6



5: The ALMATI-DATAPAR Binary Table

- Shouldn't there be an entry about how many integrations are contained in the table?

Reply:

That's in the NAXIS2 for the table?

---023-----  
(sichikawa)  
p.6  
2) (p.6)  
I insist that it is necessary 'BITPIX' keyword  
in the primary header even if it has no meaning.

Reply:

The standard says you're right. We'll mention it.

---024-----  
(sichikawa)  
p.6  
3) (p.6)  
There should be more information in the primary header.  
I strongly recommend that there should be keywords  
in the primary header, which describe date (file creation date  
or observation date), file identification (file name or sequential  
number), and the structure of the file (list of binary tables  
and number of rows of each binary table).

Reply:

That's a good idea for the date and file name, but the list of  
binary table extensions can be long, and it requires that the  
header be rewritten which might not be convenient for the on-line  
system (it would have to know how much space to leave initially).

---024b-----  
(knakanishi)  
p. 6  
2) (pp. 6, Section 3)

The order of the binary tables is specified, but the table  
ALMATI-MONITOR is not appeared.

Reply:

We will add this.

---024c-----  
(knakanishi)  
p. 6  
3) (pp. 6, Section 4)

I recommend that much more keywords which give detailed definition

about file structure are added to primary header. In most cases, many (not one) sets of binary tables are contained in one FITS file. For another common case, numbers of ALMATI-AUTODATA and/or -CORRDATA binary tables are contained in a binary table set using plural basebands simultaneously. In these situation, structure of a FITS file should be very complicated. I suppose it is useful to describe using some new keywords in Primary Header that how many extensions and/or how many sets of binary tables are existing in the FITS file.

Reply:

See reply to 024 above.

---025-----  
(egreisen)  
p.7

9. In DATAPAR, TIMESYS is of A format with no units.

Reply:

We'll correct this error.

---026-----  
(egreisen)  
p.7

10. WCS Paper III proposes OBS-LONG, OBS-LAT, OBS-ELEV for the keywords you call SITE.... It might be better to use these (tentative) standard ones.

Reply:

We'll do it.

---027-----  
(egreisen)  
p.7

11. In DATAPAR: NO\_POL is usually used for the number of polarizations correlated - called NO\_STK here. If there is 1 feed there is 1 Stokes, if there are 2, then there is 1, 2, or 4 Stokes. Who cares how many polarizations the antenna has? We only care what was used and how.

Reply:

We may change NO\_STK to NO\_POL and NO\_POL to somethinh like NO\_FEED. But some parameters like the calibration parameters are total powers measured by each feed, or efficienciers for each feed which have no or little meaning in the Stokes parameters representation. So if there are two feeds and four correlation products: these are two different numbers. We may - or may not - forget about the other feed if only one is used for correlation. Total power from the other feed, if running, might

be used to add information on the atmosphere status.

---028-----  
(fviallefond)  
p 7.

EXPOSURE INtegration time for one observation i.e. EXPOSURE=INTEGNUM\*INTTIM

Reply:

Or rather EXPOSURE=sum\_of(INTEGNUM\*INTTIM).

---029-----  
(fviallefond)  
p 7.

SOURCE OBJECT seems more widely used in FITS files.

Reply:

We may change that.

---030-----  
(graffi)  
p.7 Number of antennas

This gives me the idea that there might be 3 antennas. Would it make sense, at least technical? Should we consider this? Form a formal point of view Na must correspond to a list of length Na. The only one I see is TELESCOP, where there is no Na factorization (see also previous comment).

Reply:

It will make sense when the three prototypes are available and a correlator with 3 inputs is available. See the reply to 015.

---031-----  
(graffi)  
p.7 Nbd,... Nau

Same comment as before. It should be possible to reproduce at least manually the same observation, but it can only be done if all info related to which base bands, whic side-bands etc is given

Reply:

The information is there in the headers of the data tables.

---032-----  
(pgrosbol)  
p.7, s.5.1: OBS-NUM and SCAN-NUM

It may be better to give a full unique id instead of just an internal counter.

Reply:

With DATE-OBS we have a full unique id ?

---033-----  
(pgrosbol)  
p.7, s.5.1: DATE-OBS

The full date/time should be given in ISO format. ESO is using the MJD-OBS (Modified Julian Date) as an easy reference to the start of observations.

Reply:

We use that for date and time (see the sample header page 19).

---034-----  
(pgrosbol)  
p.7, s.5.1: RA, DEC, ...

This set of keywords suggests that only one sky position is considered, however, if observations on different sources are made the relation between source positions and table extensions is unclear.

The FITS IDI format uses SOURCE extensions for this purpose. Although I am concerned by the usage of internal index number to cross-reference between table extensions, it is a way to specify multiple sources in a single file. It is unclear how the current proposal address this problem.

Reply:

These refer to one central position; there is only one in each observation. Observations are short enough so that this is not a restriction. Inside an observation scanning is possible, but only measures in an offset coordinate system. However all the observations in a file could refer to different sources. If they refer to only a few sources. The source information will be repeated in all observations of the same source. This is for the sake of simplicity.

---034b-----  
(knakanishi)  
p. 7  
4) (pp. 7, Section 5)

If the observed source is moving object (such as planets, comets, asteroids), the source coordinate vary from time to time. Informations about absolute pointing direction (not Az/El) are useful and necessary for reducing observational data of mapping or scan across a moving object. I suggest adding new optional column in ALMATI-DATAPAR which give coordinates of telescope pointing corresponding integration. There is two possible expression of telescope pointing coordinate for observing moving object; one is in absolute coordinates (equatorial, Galactic, or Ecliptic), another

one is an arbitrary coordinate referring to moving object itself.

Reply:

This is a good suggestion; it enables a posteriori checks.

---035-----  
(egreisen) p.8

2. The DATAPAR header contains a lot of parameters that can be computed from each other. This is fine in the real-time system but should not appear in off-line systems. What always happens there is that one program works with RA and forgets Galactic Coords and the next does the reverse, after which neither are correct. In the online system this should not be a problem and may help with editing, real-time displays to operators, etc.

Reply:

On one hand we should avoid that in the final system. On the other hand as this is raw data from a test interferometer, a little redundancy might enable to track down software errors?

For coordinates only the set that define the source direction should be present (see note 3 page 8).

---036-----  
(pgrosbol)  
p.8, s.5.1: UT1UTC, IATUTC

I understood that UTC will no longer apply leap seconds. If this is the case, these two keywords will always be constant and therefore redundant.

Reply:

Is that decided? I thought it was only proposed.

---037-----  
(pgrosbol)  
p.8, s.5.1: POLARY, POLARY

The meaning is not fully clear to me. Either it is redundant information or could be replaced field rotation.

Reply:

It's the Pole motion around whatever flag they have there. It's important to know what value was used if you want to reprocess the data later in an astrometric way. May be there is a standard keyword for that? I took that one from FITS-IDI.

---038-----  
(pgrosbol)  
p.8, s.5.1: Note 8: since ALMATI-DATAPAR is the first binary table.

In principle, a program should not assume a specific order

of extension in a FITS file. If basic informations are needed it may be better to place them in the prime header.

Reply:

The order is only there I think for efficiency of the reading program. In principle there is enough information for it to work if all binary tables are in random places (thanks to OBS-DATE, SCAN-NUM and OBS-NUM). But it might mean many random accesses in disk or memory.

---039-----  
(egreisen) p.9

13. UUVVWW probably should be in CORRDATA rather than DATAPAR. They are always essential with the data. What are the units of UUVVWW? You will need them in wavelengths in imaging and that varies with baseband and spectral channel.

Reply:

Most columns in DATAPAR are essential with the data. We set them antenna-based and in seconds so that they can be shared between several CORRDATA tables at different central frequencies, and anyway these tables have many frequency channels. The raw data are far from ready to produce images.

---040-----  
(egreisen)  
p.9

14. TIME in CORRDATA is in A format - not very useful. Perhaps a time in days in D format from the start of the observing run would be helpful in the CORRDATA, AUTODATA, and HOLODATA tables instead or in addition.

Reply:

You meant in DATAPAR. We had put it in seconds (or days?) then switched to A format for consistency. We will use seconds (see 33).

---041-----  
(fviallefond)  
p 9.

INTEGNUM Is this the number of correlator dumps in one observation

Reply:

No it's the sequence number of each integration in the observation

---042-----  
(fviallefond)  
p 9.

INTTIM Is this the integration time of a dump or of one observation  
INTEGTIM suggested.

Reply:

It's the integration time of the ddump; INTEGTIME is better.

---043-----  
(pgrosbol)  
p.9, s.5.2: TOTPOWER

Using Volt as the unit for a total power is unusual, at least  
for an optical/infrared astronomer like me.

Reply:

You're right it looks strange. "Counts" is probably better (it's  
uncalibrated so you cannot use watts, and the detected signal  
goes through some ADC.

The actual unit will be 'adu', to conform with standard.

---044-----  
(egreisen)  
p.10

3. The MONITOR data will become very hard to use in fast OTF  
single-dish observations. Given that FLAG appears in the DATAPAR  
table, the science data set will not need the monitor. The TI data  
set probably will benefit from keeping it around in the present  
form. Eventually tables kept on the natural monitor data rates  
should be used.

Reply:

You are right. The Monitor is an ad hoc inclusion for the Test  
Interferometer.

---045-----  
(pgrosbol)  
p.10, s.6: 'Missing character string ... as spaces'

The FITS standard specifies that ASCII NULL should be used.

Reply:

Agreed

---045b-----  
(knakanishi)  
p.10  
5) (pp. 10, Section 5, Notes 9, for ALMATI-DATAPAR binary  
table columns)

In this example, both the 1-second interval data and 60-second  
integrated data are written in same binary tables. I suppose this  
make structure of binary tables much complicated. It seems to be  
more simple and easy to handle file in the case that we create two  
independent FITS file, put only 1-second interval data in one file,

and put 60-second integrated data in another. What is the advantage from putting differently sampled data into one FITS file?

Reply:

There is a practical reason: these files will be converted to be treated by the Plateau de Bure software that will be used for simple data reduction. What you suggest would make the conversion more complicated, as we would need to keep track of two or more files for each observation. This constraint will not be present for the final ALMA format.

---046-----  
(egreisen)  
p.11

4. The ANTENNA tables mixes things that are stable with things that depend on time and frequency. The system temperatures should be moved to a TEMPERATURE table and kept on the time scale of the data records (i.e. INTEGNUM). Editing (and instrument studies) based on aberrant temperatures will be useful but only if the Tsys data are at a finer time resolution. The frequency dependence could be treated by recording a Tsys for each baseband. One per data channel is gross overkill and, depending on the type of correlator, probably meaningless.

Reply:

We didn't want to multiply tables (at least for the TI). Tsys (p.12) is only here one per baseband and feed. But if we need this at higher resolution (in ALMA a baseband can be 2 GHz) is open. As Antenna table lists only calibration information we decided to rename it 'CALIBRATION-ALMATI'.

---047-----  
(pgrosbol)  
p.11, s.6.2: 'file may be dominated by monitor data'

If this is really the case, it would seem better to have separate data streams for science and monitor data.

Reply:

That's only for the TI, with a low data rate anyway. If for ALMA it is held to be valuable to keep the monitor data we would write it at its "natural" rate. The problem comes about because for the TI we are resampling the monitor data to the integration times (which can be 1ms for ALMA autocorrelations).

---048-----  
(pgrosbol)  
p.12, s.7.2: POLA

The unit should be in deg.

Reply:



OK.

---049-----  
(pgrosbol)  
p.12, s.7.2: PRESSURE

The unit is given in hPa which in principle is okay. However, when frequencies are specified in Hz and not GHz, it would be better to use SI base units all places.

Reply:

We'll switch to pascals.

---050-----  
(pgrosbol)  
p.12., s.7.2: PHOT, PCOLD, PSKY

The unit of Volt for power is unusual.

Reply:

See above. -> counts.

---050b-----  
(knakanishi)  
p.12  
6) (pp. 12, Section 7.2)

In ALMATI-ANTENNA Binary Table Columns, there is numbers of variable quantities within one observation or scan, from HUMIDITY , AMBIENT, PRESSURE, to TDEWAR. I suppose those variable quantities should be putted into ALMATI-MONITOR binary table.

Reply:

I have put there all coefficient that are actually needed for data calibration (in CLIC software), while MONITOR contains all monitored data for plotting purposes.

---050c-----  
(knakanishi)  
p.12  
7) (pp. 12, Section 7.2)

For ALMATI-ANTENNA Binary Table Columns, only eleven antenna pointing coefficients are defined. It is probable, however, that much more pointing coefficients are used for antenna pointing correction. Can we add new column(s) to give pointing coefficients which is not presented in table currently?

Reply:

I have been following the coefficients defined in Jeff Mangum's ALMA memo. We can add more coefficients as the needs arise, in further versions of this document.

---051-----  
(egreisen)  
p.13

5. The CORRDATA header should probably set CRVAL1 to 1.0 = real part, and should use NAXISn words to specify the number of complex, frequency, sideband, Stokes. What is the use of all these LO frequencies, especially if they do not include significant parts? The source flux is to be determined from the data, why is it in the header? Scanned observations, resolved visibilities, etc do not have much relationship with a single polarization set of numbers. Aren't sidebands always separated on modern interferometers? They have to be to do spectral-line work and field-of-view considerations will force that for broad-band continuum as well.

Reply:

- CRVAL1=1.0, OK.
- NAXIS values were not listed (they are on page 22).
- LO frequencies are just to keep track of the actual values rather than having to recompute them afterwards, if trying to understand unexpected behaviour in the visibilities (spurious lines, ...). Remember this is a test system and a raw data format.
- Source fluxes here mean they were previously measured by the system and made available for the present observations
- ALMA can either separate bands (by  $\pi/2$  phase switching) of reject one of them (by frequency offset).

---052-----  
(egreisen)  
p.14

12. In CORDATA CRPIX2 is usually E format and CDELTA2 doesn't usually require D.

Reply:

Yes: CRPIX2 and CDELTA2 can be E.

---053-----  
(fviallefond)  
p 14.

ALMATI-CORRDATA:

CTYPE2      Often includes the velocity rdefinition and reference system like FRE-OLSR for optical LSR or FRE-RHEL for radio heliocentric. FITS files are not well homogeneous on this subject. Can you list the possible values of VELDEF. What relation with VELTYP in the sample ALMATI-DATAPAR table header?

DELTAVEL    This is not really meaningful since sampling is made with constant DELTANU. Should be suppressed.

Reply:

- DELTAVEL can be suppressed;
- In principle object velocity and type should be here and not in DATAPAR since it may be different for different basebands if different transitions are observed; but it is academic for the Test Interferometer.

---054-----  
(pgrosbol)  
p.14, s.8.1: CTYPE1, ... CDELTA4

These keywords do not conform to the WCS proposal and should be changed.

Reply:

What should they be? We pragmatically delayed the response to those questions until we define the actual data format for ALMA. One of us (BEG) has had a bad experience implementing a previous draft of WCS only to have it change.

After discussion we decided to use WCS in the revised version of the data format.

---055-----  
(pgrosbol)  
p.15, s.8.2: Note 2

There seems no good reason to have 'dummy' values in the data matrix.

Reply:

The word dummy is may be not well chosen. The values are real and imaginary, but 'real' and 'imaginary' are not coordinates on an intuitive axis. We just say real=1 and imaginary=2.

---056-----  
(egreisen)  
p.16

6. Autocorrelation data may be treated the same as interferometer data with antenna 2 = antenna 1 and the number of pixels on the COMPLEX axis = 1 (or 2 with imaginary part always 0). In polarization single-dish the cross-hands may have phase also so treating it as only 1 may not be a good idea.

Reply:

Good suggestion. Let's see if it can be easily implemented.

---057-----  
(pgrosbol)

p.16, s.9.1: CTYPE1, ... CDELT2

These keywords do not conform to the WCS proposal and should be changed.

Reply:

See 054.

---058-----  
(egreisen)  
p.17

7. I see no place in the HOLODATA section to tell me which antenna is being measured. I know little about holography, but I know that it is done in both single and interferometer modes. Does HOLODATA refer to both modes? In other areas you provide data that can easily be inferred from other provided data, but here one must infer things in ways that are not obvious. The antenna and perhaps other data too should be included here for clarity even if it is not essential.

Reply:

HOLODATA refers only to single-dish mode, with a dedicated receiver in the prime focus. The associated DATAPAR should say NO\_ANT=1 and the antenna name and station are there. It should be better explained.

---059-----  
(sichikawa)  
p 18.  
4) (pp.18-26)

As for comment parts (after '/') of each line of the header, there should be more usefull description.

Reply:

We should improve that.

---060-----  
(sichikawa)  
p 18.  
5) (pp.18-26)

I think it is more convinient if 'EXTNAME' line could be next of 'TFIELDS' line.

Reply:

That's due to the way FITSIO works (or to the wrong way I've been using it).

---060b-----  
(knakanishi)  
p.18

8) (pp. 18-26, A-E)

I recommend that the keyword EXTNAME is appeared prior to other keywords except keywords in fixed order (from XTENSION, BITPIX, to TFIELDS), in order to be easily recognized the type of the binary table by off-line analysis software and also human eyes.

Reply:

There is no really need for that other than more human readability. I have been using FITSIO to create these examples and this library creates the column descriptions first.

---060c-----  
(knakanishi)  
p.18  
9) (pp. 18-26, A-E)

I recommend that the keyword TDIM is in order of just after TTYPE, TFORM, and TUNIT.

Reply:

We show just a sample header; however the order of keywords here is not mandatory.

---061-----  
(pgrosbol)  
p.19, s.A: Header example

It would be better, for visual inspection of the table headers, if keywords like EXTNAME, DATE-OBS SCAN-NUM, OBS-NUM, LST, ... would be placed in the start i.e. in front of the T\* keywords.

Reply:

That's due to the way FITSIO works (or to the wrong way I've been using it).

---062-----  
(pgrosbol)  
p.19, s.A: TELESCOPE

The value of this keyword does not correspond to the specifications.

Reply:

I will check that.

---063-----  
(pgrosbol)  
p.20, s.B: TUNIT6

The unit should probably be 'deg'.

Reply:

OK

---064-----

(pgrosbol)

p.23, s.C: TRANSITI

As all other strings have been specified to 8 characters,  
it would also be reasonable for this keyword to give the  
empty string as ' '. Alternatively, the keyword  
should be omitted.

Reply:

OK

---065-----

(pgrosbol)

p.23, s.C: CTYPE1 ... CDELT4

Keywords does not conform to FITS WCS proposal.

Reply:

See 054.

---066-----

(pgrosbol)

p.24, s.D: CTYPE1 ... CDELT2

Keywords does not conform to FITS WCS proposal.

Reply:

See 054.

\_\_\_\_\_oOo\_\_\_\_\_