IVS Memorandum 2017-001v02

3 March 2017

"Correlator Report"

Alessandra Bertarini, Ed Himwich,
David Hall, Mike Titus

Correlator Report

A. Bertarini, E. Himwich, D. Hall, M. Titus, B. Corey

March 2017

1. Introduction

To ease the readability of the correlator reports for IVS experiments, correlator representatives from Haystack, Washington and Bonn met at the Technical Operation Workshop (TOW) in 2009 to define a standard.

The standard was defined to help both the stations and the analysts to read and understand the correlator output. In the next paragraphs we will describe the standard and recommend all the IVS correlators to follow it with the caveat that the correlator report is strongly connected with the usage of the Haystack HOPS package for fringe fitting the data.

2. Sections

All the sections start with the character "+" followed by the name of the section. If the name of the section is composed of more than one word, then the words are connected using an underscore (i.e. +THIS_SECTION).

Important to mention is that there is a limit of 70 characters per line due to the software used by the analysts. All characters after the 70th are truncated, hence the correlator report becomes unreadable.

2a. Header

```
+HEADER
CORREL   BONN
DATABASE 17JAN17XA
SESSNAME R1775
OBSTIME  2017/01/17 2017/01/18
UTSTART  1700
DURATION 24
DOY      017-018
CORRTIME 2017/01/31
CORRPASS 1
EXPORT   DONE
```

Above is an example of the *header*. This introductory section identifies:

1. <u>COREL</u>: correlator. All the correlators codes can be found at ftp://cddis.gsfc.nasa.gov/pub/vlbi/ivscontrol/master-format.txt (section CORR). In the example above is BONN (please note and keep the capital letters).

2. <u>DATABASE</u>: database name. The database name is written in the format year, month, day of the observation and database code as defined in ftp://cddis.gsfc.nasa.gov/pub/vlbi/ivscontrol/master-format.txt (section DBC Code). In the example above is 16OCT04XA.

3. <u>SESSNAME</u>: experiment name as defined in the master file in the IVS web page (http://ivscc.gsfc.nasa.gov/program/control_files.html, follow the link masterYY.html (e.g. master16.html and read the session code). In the example above is R1760.

4. <u>OBSDATE</u>: Observation date. Start date and end date of the experiment written in the format YYYY/MM/DD. In the example above the start date is 2016/10/04 and the end date is 2016/10/05.

5. <u>UTSTART</u>: start time of the observation in UT in the format HHMM. In the example above is 1700 UT.

6. <u>DURATION</u>: duration of the observation in hours. In the example above is 24.

7. <u>DOY</u>: start and end days of the year in which the observation took place. The two DOY are separated by a dash ("-"). In the example above 278-279.

8. <u>CORRTIME</u>: duration of the correlation in the format YYYY/MM/DD. The two values indicates the start of the correlation and the end of the correlation. In the example above are 2016/10/18 2016/10/18.

9. <u>CORRPASS</u>: correlator passes. A value greater than 1 indicates that more stations participated in the observation than there are playback units available at the correlator to accommodate them, i.e. the correlation has to be run more than once to cross-correlate all the baselines against each other. In the example above the value is 1.

10. <u>EXPORT</u>: indicates whether the databases have been submitted to the IVS data centers. In the example above is DONE (please note and keep the capital letters). Normally the correlator report is submitted once the experiment has also been submitted to the data centres, so the "DONE" is always there. Sometimes, due to delay in data delivery at the correlator, the correlators export to the data centres a preliminary release. In this cases write "DONE as preliminary release without <station>".

2b. Summary

```
+SUMMARY
 Qcode         % of Total    % of Correlated
                  scans            scans

   5-9             72%              -
    0               4%              -
 4-1,A-H,N         24%              -
 Removed            -               -
```

Above is an example of the *summary section*. This section gives the reader the distribution of the quality of the data in percentage. In case one station is largely responsible for over 10 % of the data being removed from processing, the correlators write a note so that the analysts know there might be a sparse baseline.

1. The column "**Qcodes**" represents the fringe quality code as defined by HOPS-fourfit. In a later section (+QCODES) there will be a description of these codes.

2. The row quality codes "**5-9**" represents the percentage of data that will be used for further analysis.

3. The row quality code "**0**" represents the percentage of non-detections.

4. The row quality codes "**1-4,A-H,N**" represents the percentage of the data that showed a problem either during the correlation or fourfit and are not usable by the analysts.

5. The row **"Removed"** indicates the percentage of the scheduled scan, that for some reason were not observed.

6. The column "**% of Total scans**" considers all the scans scheduled within the observation.

7. The column "**% of Correlated scans**" ignores the scans that were scheduled but were not observed.

These percentages are calculated based on the quality code table as produced by HOPS-aedit software, described in the +QCODES section below.

2c. Correlator Notes

```
+CORRELATOR_NOTES
This is the final version with both ZELENCH and SVETLOE baselines.
Baselines IY scans 034-2138 and 035-1614, IN scan 035-1104 and NV
scan   035-0009, the fringe rate passed through 0.
```

This section gives an overview of issues related to the correlation. It gives a hint to the analysts on what went wrong during correlation or e.g. if some tag-along station was correlated but was not originally scheduled.

If the correlation went well, then this section is empty.

2d. Station Notes

```
+STATION_NOTES
HOBART12   (Hb/H): Manual phasecal applied at HOBART12
HOBART26   (Ho/A): Station reports late start due to PCFS failure.
                   Periods of low pcal amplitude leading to H-codes.
KOKEE      (Kk/K): Manual phasecal applied at KOKEE.
```

The section above contains specific comments related to the performance of each station participating in the observation. Some of the comments are simply a copy of the ivs-ops stop mails, some others are issues discovered while correlating and fringe fitting the data. Important for the analysts is always to report if a station was fringe fitted using the manual phase calibration (in the example above: "Applied manual pcal").

The stations names are as written in the vex file at the section $ANTENNA and the two letter codes are the standard two-character identifier (also readable in the vex file section $SITE, site_ID). The one-letter code is something more specific for the HOPS package. Since in a later section of the correlator report the stations are referred using this one-letter code, it is important to specify it in the station notes because the one-letter codes are not unique, although the correlators try to use always the same code for the same station. For the DiFX users, this code is assigned from difx2mark4 while converting the DiFX output files into Mark4 output file (used by HOPS).

Vex2 will re-introduce the concept of the one-letter code.

2e. Drop Channels

```
+DROP_CHANNELS
  Aq SR4U
  Yk SR5U SR6U XR8L
```

This is a systematized list of channels that required flagging during fringe fitting for various reason that are described in the station notes. The first column contains the standard two-character code identifier. The meaning of the codes in the table above are described in the *channel section* of the correlator report (see below), but in short, they are band|polarization|BBC #|sideband. The BBC number starts with the lower frequency channel in each frequency band (i.e. SR1U = first S-band BBC, XR1U = USB of the first X-band BBC).

2f. Manual Phase Calibration

```
+MANUAL_PCAL Gc Yk
```

This is a list of the stations (always using the two-character code identifier) for which manual phase calibration was applied. The reasons are normally found in the station notes. The usage of manual pcal is decided at the moment of fringe fitting; if no pcal was injected or the pcal does not correct the visibility phases, then the correlators apply manual pcals. This is true with only one exception: Kokee Park. It was established by the analysts that applying the pcal to Kokee Park deteriorates the solutions, hence the correlators are requested never to use them for Kokee.

2g. Channels

```
+CHANNELS
 Channels: XR1U = band|polarization|BBC#|sideband

XR1U/L  BBC01 8210.99
XR2U    BBC02 8220.99
XR3U    BBC03 8250.99
XR4U    BBC04 8310.99
XR5U    BBC05 8420.99
XR6U    BBC06 8500.99
XR7U    BBC07 8550.99
XR8U/L  BBC08 8570.99
SR1U    BBC09 2212.99
SR2U    BBC10 2227.99
SR3U    BBC11 2237.99
SR4U    BBC12 2267.99
SR5U    BBC13 2287.99
SR6U    BBC14 2292.99
```

This table assigns unambiguous labels to the BBC numbering and sky frequencies to define the channels clearly. These labels are used in the *station notes* and the *drop channels* section. The sky frequency and BBC numbering can be read in the $FREQ of the vex files.

2h. Clocks

```
+CLOCKS

 Clocks: Bonn

  Station       fmout-gps         Used          rate              Comments
                [usec]            [usec]        [sec/sec]
    Ft           0.42             0.77          0.207E-12
    Ht           8.12            11.16          0.245E-12
    Wn          -0.36             1.50         -0.001E-12
    Wz         -37.55           -37.16         -0.251E-12
    Yg          12.61            14.99          0.936E-12
    Date:    2017/01/17         17:00:00
```

These are the clock offsets used for each station.

1. The line "**Clock**" identifies the correlator that produced the self contained clock section (e.g. Bonn, GSI, Haystack, SHAO, UTAS, WACO..).

2. The column "**Station**" contains the two-character code. Note that a station may appears in more than one line if more than one clock model was required (e.g. clock jumps).

3. The column "**fmout-gps**" is the (smoothed) value reported in the station field system log files corresponding to the epoch of the "Date" line. Positive values indicates that the station clock is early compared to GPS/UTC.

4. The column "**Used**" is the value actually used at the correlator, after correcting all the clocks.

5. The column "**Rate**" is the clock rate (drift) applied to keep the fringes within the search window. A positive rate means that the station clock is running faster than the GPS.

6. The column "**Comments**" can contain information like clock jumps, or report which station was used as reference when correcting the clocks.

7. The line **"Date"** indicates the clock model epoch date in the form YYYY/MM/DD and with the time in the form HH:MM:SS. For reports from Haystack with no time field, the time for a 24-h session is usually 00:00:00 UT on the second day. For other correlators, it is usually the start of the experiment.

Note that for geodesy there are fixed clock offsets added by the correlators to the fmout-gps values. These fixed offsets are necessary for calculating the correct UT1 values within CALC/SOLVE. All the IVS correlators must apply these offsets when correlating IVS experiments.

2i. Quality Codes

```
+QCODES

Qcod  0 1 2 3 4 5 6  7   8    9 A B C D E F  G H  N -   Tot
-------------------------------------------------------------
bQ:X  1 0 0 0 0 0 0  0  10  181 0 0 0 0 0 0  0 0  0 0   192
bQ:S  5 0 0 0 0 0 0  0   8  197 0 0 0 0 0 0  0 0  0 0   192
bK:X  2 0 0 0 1 0 1 27  57   65 0 0 0 0 0 0  0 0  0 0   153
```

```
bK:S   9 0 0 0 0 0 2  3  87    38 0 0 0 0 0 0 14 0  0 0    153
[..]
---------------------------------------------------------------
Tot 520 0 0 0 2 6 75 300 155 653 0 0 0 0 0 0 89 0 2708 0 1178
```

The quality code table summarizes the fringe quality codes as calculated by HOPS fourfit. The table is produced by HOPS aedit (command psfile). The output of aedit exceeds the 70 character max allowance for a line (see Introduction), so the correlators shrink it to match the limit.

1. The first column contains the baseline labelled using the one-letter codes (e.g. in the first line "aj" = baseline between station "a" and station "j").

2. The second column is the frequency (S/X-band).

3. All other columns represent the numbers of scans that have the specific quality code as reported in the first row of the table (from 0 to N).

4. The last column, "Tot", is the number of scans scheduled on that baseline for the experiment.

The last row of the table reports the totals of the quality codes for all the baselines and all the bands. These totals are used to calculate the percentage of data usable from the analysts as reported in the *summary* section of the correlator reports.

Below the quality code tables there is a legend of the alpha-numeric characters used to label the quality codes (see below).

```
 Legend:
QC = 0    Fringes not detected.
   = 1-9 Fringes detected, no error condition. Higher #, better quality.
   = B    Interpolation error in fourfit.
   = D    No data in one or more frequency channels.
   = E    Maximum fringe amplitude at edge of SBD, MBD, or rate window.
   = F    Fork problem in processing.
   = G    Fringe amp in a channel is <.5 times mean amp (only if SNR>20).
   = H    Low Phase-cal amplitude in one or more channels.
   = N    No valid correlator data.
 Minus    Scans in original schedule file for which correlation was not
          attempted, usually because of known station problems.
   Tot    Total number of scans in schedule.
```

2j. Ratio between Observed SNR and Predicted SNR

The table below shows baseline by baseline the observed vs predicted (by sked) signal-to-noise-ratios over all the experiment, all sources.

```
+SNR_RATIOS

 MEAN RATIOS = Observed SNR / Predicted SNR  for exp no. 3981

  ...by baseline, over all sources:
```

```
bl      X       n       S       n
Ni      1.13    32      0.63    34
jW      0.85    93      0.57    84
```

This table is calculated using a standalone C-program, that now has also been integrated into the HOPS fourfit package as downloaded together with the DiFX code. The program requires the alist file (from HOPS alist), the schedule file and an ancillary ASCII file that point the two-character station code to the one-letter code used for the specific experiment. Values around 1 indicate that the SNR was properly calculated. Above 1 that the SNR was underestimated, and below 1 that the SNR was overestimated.

The legend of the table is the following:

bl = baseline (one-letter code)

X = X-band SNR ratio (observed/predicted)

n = number of X-band scans used for the calculation of the ratio

S = S-band SNR ratio (observed/predicted)

n = number of S-band scans used for the calculation of the ratio

2k. Fourfit Control File

This section is copied from the fourfit control file used for the experiment. At the end of the section there is the signature of the responsible at the correlator who correlated and fringe fitted the experiment.

```
+FOURFIT_CONTROL_FILE
    [the control file goes here]
```

2l. End of File

The end of the correlator report is marked by:

```
+END
```