Robert M. Campbell: Recent Results from the EVN Mark IV Data Processor at JIVE, IVS 2004 General Meeting Proceedings, p.123–127
http://ivscc.gsfc.nasa.gov/publications/gm2004/campbell

Recent Results from the EVN Mark IV Data Processor at JIVE

Robert M. Campbell

Joint Institute for VLBI in Europe
e-mail: `campbell@jive.nl`

Abstract

Recent achievements at the EVN Mark IV data processor at JIVE include decreasing the read-out time for the whole correlator (32 Mark IV boards) to 0.25s, improving the end quality of user data (*e.g.,* excising TRM byte-slips, applying an improved 2-bit van Vleck correction), developing new astronomical capabilities (*e.g.,* oversampling), and strengthening liaison procedures with PIs (*e.g.,* through the new EVN archive facility). The move to Mark 5 and fiber-linked VLBI is well underway. 90% of the user experiments from the Oct/Nov'03 EVN session had at least one station recording onto disk, and three stations had moved over to all-disk recording in the Feb'04 session. Ftp fringe-tests have been included as a regular feature of EVN sessions, allowing for faster feedback to stations.

1. Background

A key item in the Mark IV upgrade of the EVN was the construction of the EVN Mark IV data processor at the Joint Institute for VLBI in Europe (JIVE). JIVE is hosted by ASTRON in Dwingeloo, the Netherlands, and is funded by science councils of a number of European countries. Special projects have been funded directly by the European Commission. The EVN Mark IV data processor [1], [2] was constructed in the context of the International Advanced Correlator Consortium through which the other Mark IV geodetic correlators were also built, with significant contributions, from European members, in hardware from CNR/IRA, control software from Jodrell Bank, and correlator software from ASTRON.

The first fringe on the EVN Mark IV data processor was seen on 21 July 1997 and its official inauguration took place on 22 October 1998. The "first science" resulting from data correlated on the EVN Mark IV data processor detected HI absorption against the counter-jet very close to the nucleus of NGC 4261 [3]. The first paper published in *Nature* based on EVN observations correlated at JIVE has also appeared [4], using OH mega-maser emission to study the kinematics and spatial structure of the circumnuclear torus in Mrk 231. The EVN Mark IV data processor now correlates the vast majority of astronomical EVN experiments, and about half of the global experiments. More information about the EVN and JIVE can be found at the websites `www.evlbi.org` and `www.jive.nl` .

2. Current Capabilities

Basic characteristics of the design of the EVN Mark IV data processor include: simultaneous correlation of up to 16 stations with 16 channels per station, each having a maximum sampling rate of 32 Ms/s (thus a total of 1 Gb/s per station for 2-bit recordings). The correlator houses 32 Mark IV boards. The principal science drivers behind the development of the data processor and associated software include the ability to handle continuum dual-polarization observations, spectral line experiments, and phase-reference mapping. Figure 1 shows the data playback units, the station units, the data distributor unit, and control computers. The correlator itself is housed in a separate room out of this view to the right.

Figure 1. A view of the main control space for the JIVE Mark IV Data Processor.

2.1. Features Snapshot

The EVN Mark IV data processor can currently correlate/provide:

- 1- and 2-bit sampling (all but a handful of experiments use 2-bit sampling).
- Mark III, Mark IV, VLBA, and Mark 5(A) recordings.
- sustained 512 Mb/s tape recordings or 1 Gb/s for Mark 5 recordings.
- parallel- and cross-polarization products as desired in dual-polarization observations.
- up to 2048 frequency points per baseline (see the discussion following equation 1 below).
- full-correlator integration times down to 0.25 s (half-correlator t_{int} down to 0.125 s).
- oversampling at 2 or 4 times the Nyquist frequency in order to provide subband bandwidths down to 500 kHz (the maximum Nyquist-sampled BW_{sb} is 16 MHz).
- multi-pass correlation (*e.g.,* for observations having >16 stations at any given time).

Capabilities whose development is still underway or not yet fully tested include pulsar gating, speed-up (playback at a bit-rate faster than that used in recording), and phase-cal extraction. Capabilities that are yet to come include sub-netting and recirculation (achieving greater equivalent correlator capacity for observations that don't use the maximum bandwidth per subband).

2.2. Correlator Capacity

The total correlator capacity can be expressed as:

$$N_{\mathrm{sta}}^2 \cdot N_{\mathrm{sb}} \cdot N_{\mathrm{pol}} \cdot N_{\mathrm{frq}} \leq 131072 \tag{1}$$

Here, N_{frq} is the number of frequency points per baseline/subband/polarization. N_{pol} is the number of polarizations wanted in the correlation (1, 2, or 4). N_{sb} represents the number of different subbands, counting lower- and upper-sidebands from the same BBC as distinct subbands. The value to use for $N_{\rm sta}$ is "granular" in multiples of 4: *e.g.,* if you have 5–8 stations, use "8". Independent of this equation, the maximum $(N_{\rm sb} \cdot N_{{\rm pol}_{\parallel}})$ is 16 (a station-unit limitation), and the maximum $N_{\rm frq}$ is 2048 (a single baseline/SB/pol must fit onto a single correlator board). All capabilities discussed in this report assume the use of local validity, which avoids problems ensuing from the Mark IV-format data-replacement headers correlating with each other in certain baseline-source geometries, but at the expense of a factor of two in $N_{\rm frq}$.

Table 1. Example configurations that would use the full correlator capacity (local validity)

$N_{\rm sta}$	$N_{\rm sb}$	$N_{\rm pol}$	$N_{\rm frq}$	comment
8	1	1	2048	EVN spectral-line
16	8	4	16	global cross-polarization mapping
16	14	1	32	Mark III/modeC-like (87.5% of the correlator)

2.3. Output Capacity

The minimum $t_{\rm int}$ for a configuration using the whole correlator is now 1/4 s; configurations that use no more than one-half of the correlator can achieve minimum $t_{\rm int}$ of 1/8 s. In the future, the Post-Correlator Integrator (PCI) will provide a minimum $t_{\rm int}$ for the whole correlator of 1/64 s.

These low integration times, together with the fine spectral resolution afforded by large $N_{\rm frq}$, will provide the possibility to map considerably wider fields of view through reduced bandwidth- and time-smearing effects in the *u-v* plane (see, *e.g.*, [5] §21.7.5). For example, the fields of view having $\leq 10\%$ decrease in the response to a point source arising from each of these two effects are:

$$FoV_{\rm BW} \lesssim 49.''5 \, \frac{1}{B} \, \frac{N_{\rm frq}}{BW_{\rm sb}}; \qquad FoV_{\rm time} \lesssim 18.''56 \, \frac{\lambda}{B} \, \frac{1}{t_{\rm int}} \tag{2}$$

Here, B is the longest baseline length in units of 1000 km, λ is in cm, and $BW_{\rm sb}$ is in MHz. A primary goal of such wide-field correlations would be to map the entire primary beam of each antenna composing the array with only a single correlation pass. With the existing $N_{\rm frq}$ and $t_{\rm int}$ capabilities, we can already achieve this for a variety of observing configurations. More details can be found in `www.evlbi.org/user_guide/limit.html`. Of course, one drawback to such wide-field correlations is the rapid growth of the size of the FITS file seen by the user — at our current maximum, ~7 GB per hour of observation.

2.4. Recent Improvements

The following points represent some of the data-quality enhancements since the previous General Meeting not discussed elsewhere in this report:

- Positions of EVN stations that don't usually participate in geodetic campaigns have been significantly improved, in some cases with adjustments on the order of 5m [6]. In addition, a better tie between the Wb single-dish and array positions was determined, and an unrelated effect that introduced discrete phase jumps in $\rm Wb_{arr}$ (typically once per several hours) was detected and repaired during this process. The EVN performance in phase reference experiments has significantly improved following these improvements [7].

- We implemented a mechanism in the logic of the responsible station-unit boards to detect and repair TRM byte slips and to flag associated data, thus cutting down on the number of recorrelations required and streamlining the review process.
- The van Vleck correction for 2-bit data depends on the fraction of high- and low-magnitude bits within an integration, but this dependence was not being taken into account correctly. We now compensate for the observed "sampler statistics" in a post-correlation program.

3. Operational Overview

3.1. Correlation and Logistics

We operate the correlator 80 hours per week, from which time system testing and development must also come; typically 50–60 hours per week are production. After solving a few growing pains, we can include Mark 5 stations seamlessly in jobs. However, we will achieve the full efficiency gains from Mark 5 only when the last tape disappears, since a job will start only when the last station is ready (job-preparation time for tapes continues to be ~9 minutes). We currently have 6 Mark 5A units connected to station units for production correlation, with another 3 in house. The status of Mark 5 acquisition by EVN stations can be viewed via the Mark 5 link from the TOG section of the EVN web page. In the longer term, when stations upgrade to Mark 5B we will be able to move away from using local validity, effectively doubling the correlator capacity as described in equation 1 (the impact on the minimum t_{int} would be more complicated, depending on the stage of PCI development). Use of Mark 5B without the station units could also raise the possibility of correlating more than 16 stations at once.

3.2. Post-Correlation Data Review

Our main priority is always the quality of the data we provide to the EVN users. Our internal data review process, as illustrated in Figure 2, begins by transforming the lag-based correlator output into an AIPS++ Measurement Set (MS). From the MS, we can investigate slices of the correlation functions in both time and frequency/lag, allowing us to detect and diagnose various problems with the recorded data or the correlation itself, and to find any scans for which recorrelation would be profitable. We can also make various plots more suited to providing feedback to the stations rather than to the PI (*e.g.,* parity-error rates, sampler statistics). We apply various corrections to the correlated data at this stage (*e.g.,* the 2-bit van Vleck compensation), and flag subsets of the data for low weights and other known problems resulting in (uncorrectable) spurious correlation amplitudes and/or phases.

The last step converts the final MS into FITS format, usually written to a DAT tape. We send this to the PI, along with a summary of the correlation itself and various diagnostic plots. The FITS files can can be read into AIPS directly using `FITLD`. The EVN pipeline operates on the FITS data to create the first few AIPS `CL` tables (*e.g.*, T_{sys}-based amplitude calibration, off-source flagging, *etc.*), as well as to make make preliminary images of sources authorized by the PI. Plots, summaries, and pipeline results also go to the EVN archive, `www.jive.nl/archive/scripts/listarch.php`, followed after a 1-year proprietary period by the FITS files themselves.

Unless contacted by the PI to the contrary, we aim to release an experiment's tapes/disks four weeks after the PI is notified of the experiment's completion. To supplement the review products

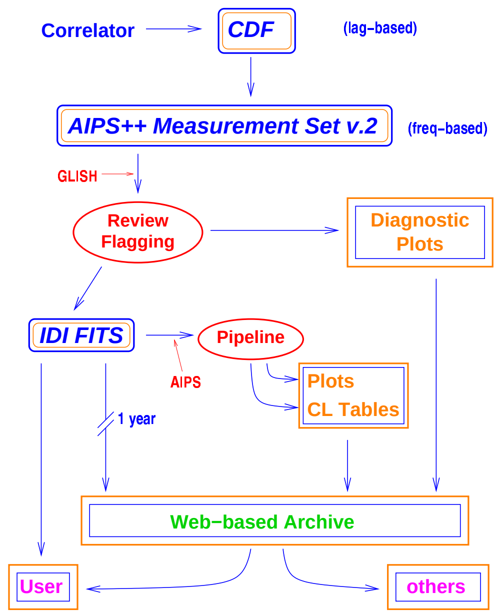

Figure 2. Post-correlation review process for an experiment.

mentioned above, the PI of course may also discuss the experiment/correlation with the responsible JIVE support scientist and/or arrange to visit JIVE for help in data reduction if desired. Specific programs exist to provide financial support for PIs from non-EVN European institutes and their collaborators in order to encourage visits to EVN institutes, including JIVE — see the "Access to the EVN" portion of the EVN web page: `www.evlbi.org/access/access.html`.

References

[1] Casse, J.L., "The European VLBI Network Mark IV Data Processor," in *Proceedings of the 4th EVN/JIVE Symposium*, eds. M.A. Garrett, R.M. Campbell, & L.I. Gurvits, *New Astron. Rev.*, **43**, 503–508, 1999.

[2] Schilizzi, R.T., Aldrich, W., Anderson, B., *et al.* "The EVN-Mark IV VLBI Data Processor," *Exper. Astron.*, **12**, 49–67, 2001.

[3] Van Langevelde, H.J., Pihlström, Y.M., Conway, J.E., *et al.* "A Thin HI Circumnuclear Disk in NGC4261", *Astron. & Ap.* **454**, L45–48, 2000.

[4] Klöckner, H.-R., Baan, W.A., & Garrett, M.A., "Investigation of the Obscuring Circumnuclear Torus in the Active Galaxy Mrk 231," *Nature*, **421**, 821–823, 2003.

[5] Wrobel, J.M., "VLBI Observing Strategies," in *VLBI and the VLBA*, eds. J.A. Zensus, P.J. Diamond, & P.J. Napier (ASP: San Francisco), 411–427, 1995.

[6] Charlot, P., Campbell, R.M., Alef, W., *et al.* "Improved Positions of Non-geodetic EVN Telescopes," in *Proceedings of the 6th EVN Symposium*, eds. E. Ros, R.W. Porcas, A.P. Lobanov, & J.A. Zensus, (MPIfR: Bonn), 9–12, 2002.

[7] Campbell, R.M. "Operational Issues of the Improved EVN Station Positions," in *Proceedings of the 6th EVN Symposium*, eds. E. Ros, R.W. Porcas, A.P. Lobanov, & J.A. Zensus, (MPIfR: Bonn), 37–40, 2002.