PROCEEDINGS OF SCIENCE

Developments at the EVN Mk IV data processor at JIVE

Robert M. Campbell[*†]
Joint Institute for VLBI in Europe
E-mail: campbell@jive.nl

If you haven't had an EVN experiment recently, you might well be pleasantly surprised by what you now encounter. Many people at several institutions have worked actively to improve the reliability of all aspects of your experiments, while the capabilities of the correlator continue to advance. You would find that there is much more pro-active liaison prior to your observations, that station problems are often pre-empted before they affect your data thanks to ftp fringe-tests, and that pipelining and the EVN Archive make using your correlated data more straightforward. You would see that some correlator features that were fledgling at the time of the previous EVN symposium in 2004 (e.g., Gbps playback) have now largely matured and that some new features currently in development are nearing fruition (e.g., recirculation, incorporating more than 2 MERLIN stations into an EVN correlation). Here, we will review recent operational and technical developments that pertain directly to the kinds of experiments you can do.

e-VLBI and real-time correlation are not explicitly addressed in this paper (see Szomoru, these proceedings). Operationally, the practical differences between e-VLBI and traditional EVN experiments lie in the pre-observation stages, since the time frame from proposal to observation is compressed significantly (currently only 2 weeks). Once projects are approved by the EVN PC, JIVE handles the scheduling. Thus, it is critical that proposals are thorough enough to allow this, and an extra premium is placed upon consulting with the Science Operations & Support group at the proposal-creation stage. There are also numerous initiatives being investigated under the aegis of the RadioNet ALBUS programme in the areas of user software development and the extension of the data product distributed following correlation (see van Langevelde, these proceedings).

8th European VLBI Network Symposium
September 26–29, 2006
Toruń, Poland

[*]Speaker.
[†]on behalf of the EVN Data Processor Group at JIVE.

PoS(8thEVN)054

1. Operational overview

The Science Operations & Support Group at JIVE provides vertically integrated assistance for all segments of your experiment—from proposing/scheduling, through correlation, to analysis of the resulting FITS data. We continue to manage actively the JIVE (`www.jive.nl`) and the EVN (`www.evlbi.org`) web pages to improve the ease of navigation and the mutual cross-linking. The EVN Users' Guide on the EVN web page (`www.evlbi.org/user_guide/`) remains the best "first stop" for on-line help. It has direct links to more explicit help geared towards the specific tasks encountered in conducting EVN experiments (proposing, scheduling, correlating, analyzing), to the EVN Data Archive (see § 1.2), to travel support information, to handy tools & documents (the EVN calculator, the EVN data analysis guide), and to EVN facts & figures (frequency availability, *u*-*v* coverage, resolution, baseline/image sensitivity, imaging limitations). Discussions with various people at the symposium have raised further suggestions for improving the web pages, both in terms of capabilities and user-friendliness, which we are now pursuing.

1.1 Pre-observation, pre-correlation

The left-hand panel of figure 1 summarizes operational and communication flow among the PI, JIVE, and EVN assets during an EVN experiment. The first step is creation of the experiment schedule. We actively encourage the PI to consult with Science Operations & Support group at JIVE during scheduling, in order to help side-step the myriad little pitfalls that may lead to unpleasant surprises when (and after) the observations are carried out. Following the observation but prior to the correlation itself beginning, we confer with the PI to make sure the correlation parameters (*cf.* § 2.1 & 2.2) are appropriate, to discover any other desired changes (*e.g.*, improved source coordinates), and to ascertain which sources should receive the one-year proprietary protection in the EVN Archive (*cf.* § 1.2). Each experiment is assigned a support scientist, who shepherds it through the correlation and post-correlation analysis stages discussed below.

1.2 Post-correlation data review

Our internal data review process, as illustrated in the right-hand panel of figure 1, begins by transforming the lag-based correlator output into an AIPS++ Measurement Set (MS). From the MS, the support scientist can investigate slices of the correlation functions in both time and frequency, allowing us to detect and diagnose various problems with the recorded data or the correlation itself, and to find any scans for which re-correlation would be profitable. We can also make various plots more suited to providing feedback to the stations rather than to the PI (*e.g.*, sampler statistics). We apply various corrections to the correlated data at this stage, such as the 2-bit van Vleck correction, and flag subsets of the data for low weights and other known problems resulting in spurious correlation amplitudes and/or phases. Finally, we convert the final MS into IDI FITS format, which can be read into (classic) AIPS directly using `FITLD`. At this stage, the support scientist sends e-mail to the PI describing the correlation and any points of interest noticed during our data review. Distribution of the FITS data can occur via physical media (DAT, DVD) or via the EVN Archive, as per the PI's preference.

During the course of the post-correlation review, we also begin populating the EVN Archive (`www.jive.nl/archive/scripts/portal.php`) for the experiment. Feedback from the

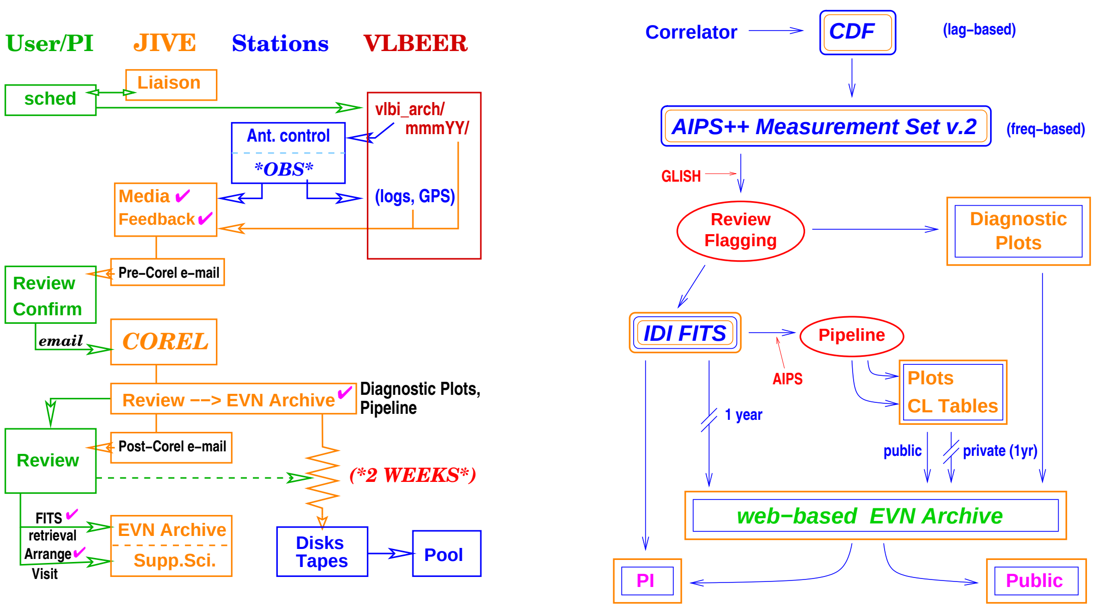

Figure 1: Operational flow within an experiment. Left-hand panel: observations and correlation (check = item available via JIVE or EVN web pages); Right-hand panel: post-correlation review.

stations and the diagnostic plots from the MS-based data review go into the archive immediately to allow the PI to get an idea about the success of the correlation even before receiving the data. The standard plots typically comprise automatically generated plots of weight(t) for all stations throughout the experiment, station amplitude(ν) and baseline amplitude/phase(ν) for a couple calibrator scans, and baseline amplitude/phase(t) for $\sim$90 min in the vicinity of one of the calibrator scans. More extensive plots may be also be made if the situation warrants. The FITS files also go to the archive, but are kept private for a one-year proprietary period (see the *EVN Data Access Policy* in the EVN Users' Guide for more details). The PI can arrange through the support scientist for a password to download the FITS data directly from the EVN Archive, if desired.

Once we receive the `ANTAB` files from the stations, the pipelining of the experiment can begin. The EVN pipeline (`www.evlbi.org/pipeline/user_expts.html`) flags data known to be invalid (*e.g.*, antenna off-source), applies an *a priori* amplitude calibration using the $T_{\rm sys}$ and gain curves from the stations, fringe-fits sources authorized by the PI, makes preliminary CLEAN images using a fixed scheme for phase and amplitude self-calibration, and creates a set of AIPS tables from various stages of the calibration/fringe-fitting process that the PI can later apply directly to the raw data, if desired. A variety of plots are saved during the above process (*e.g.*, `POSSM`s, `VPLOT`s, dirty maps), which can provide more information with which to assess antenna performance. Plots specifically associated with sources identified as private in the pre-correlation liaison are also password protected in the EVN Archive for the one-year proprietary period. The resulting AIPS tables help to simplify the initial stages of the analysis. The quality of the preliminary images may be affected by the lack of interactive data editing inherent in the pipeline concept. More details about the EVN pipeline can be found in the original pipeline paper (Reynolds *et al*. [1]).

To supplement the review products mentioned above, we encourage the PI to discuss the experiment/correlation with the responsible JIVE support scientist and/or to arrange a visit JIVE for help in data reduction, if desired. In order to facilitate such visits, the eligibility of European PIs for financial support has been broadened (the bar against EVN-institute affiliation has been dropped); see the "Access to the EVN" portion of the EVN web page (`www.evlbi.org/access/`) for more details.

2. Current capabilities

The EVN Mk IV data processor can correlate simultaneously up to 16 stations with 16 channels per station, each having a maximum sampling rate of 32 Msample/s (thus a total of 1 Gbps per station for 2-bit recordings). The EVN Mk IV data processor can currently correlate/provide:

- 1- and 2-bit sampling (all but a handful of experiments use 2-bit sampling).
- sustained 1 Gbps for Mark 5A disk recordings.
- parallel- and cross-hand polarization products as desired in dual-polarization observations.
- up to 2048 frequency points per baseline/subband/polarization (*cf.* § 2.1).
- full-correlator integration times down to 0.25 s (*cf.* § 2.2).
- 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).
- an improved 2-bit van Vleck correction to account for the statistics of high/low bits for each BBC/sideband's data stream from each station.

Capabilities whose development are still underway or not yet fully tested include phase-cal extraction and recirculation (achieving greater equivalent correlator capacity, through a time-sharing scheme for observations that do not use the maximum bandwidth per subband).

2.1 Correlator capacity

The total correlator capacity can be expressed as:

$$N_{\mathrm{sta}}^2 \cdot N_{\mathrm{sb}} \cdot N_{\mathrm{pol}} \cdot N_{\mathrm{frq}} \leq 131072 \tag{2.1}$$

Here, N_{frq} is the number of frequency points per baseline/subband/polarization. N_{pol} is the number of polarizations 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_{sta} is "granular" in multiples of 4: for example, if you have 5–8 stations, use "8". Independent of this equation, the maximum number of input channels ($N_{\mathrm{sb}} \cdot N_{\mathrm{pol}_{\|}}$) is 16, and the maximum N_{frq} is 2048 (a single interferometer must fit onto a single correlator board). The minimum N_{frq} is 16. Table 1 shows some configurations that would use the full correlator capacity. You can now evaluate interactively whether your experiment adheres to equation (2.1) via the EVN calculator.

$N_{\rm sta}$	$N_{\rm sb}$	$N_{\rm pol}$	$N_{\rm frq}$	comment
8	1	1	2048	EVN spectral-line
9	1	1	512	$9^{\rm th}$ sta: $N_{\rm frq} \rightarrow N_{\rm frq}/4$
16	8	4	16	global cross-polarization
16	2	2	128	re-arrange $\{N_{\rm sb}, N_{\rm pol}, N_{\rm frq}\}$
8	16	1	128	How $N_{\rm sta}$ increase can be
12	7	1	128	absorbed by $N_{\rm sb}$
16	4	1	128	(not constrained to be 2^n)

Table 1: Examples of “maximal” correlator configurations.

There are two independent developments coming in the near future that will effectively increase the correlator capacity you would see as a user. Recirculation will increase the right-hand side of equation (2.1) by a factor $\mathscr{R} = 16\,{\rm MHz}/BW_{\rm sb}$ up to a likely maximum of 8, but the maximum $N_{\rm frq}$ would remain 2048, as discussed above. The most significant benefit of recirculation would be increased spectral resolution for narrow-band global spectral-line observations; these could get $N_{\rm frq} = 2048$ just as 8-station experiments can now achieve. The second gain comes when we can play back Mark 5B recordings. Thanks to a change in the recorded data format, the correlator will be able to double the amount of its hardware devoted to processing correlation functions, essentially doubling the spectral capacity. In this case, the maximum $N_{\rm frq}$ would increase to 4096.

2.2 Output capacity

The minimum $t_{\rm int}$ for a configuration using the whole correlator is now $1/4$ s. There have been promising tests of modes using the Post-Correlator Integrator (PCI) at $t_{\rm int} = 1/16$ s, but these are not yet operational. The final goal of the PCI remains $t_{\rm int} = 1/64$ s. However, should recirculation be used, the minimum integration time would be increased by a factor of $\mathscr{R}$ from its nominal value.

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. 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 (Wrobel [2], § 21.7.5):

$$FoV_{\rm BW} \sim 49.''5\,\frac{1}{B}\,\frac{N_{\rm frq}}{BW_{\rm sb}}; \quad FoV_{\rm time} \sim 18.''56\,\frac{\lambda}{B}\,\frac{1}{t_{\rm int}} \tag{2.2}$$

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. The EVN calculator also evaluates time- and bandwidth-smearing fields of view in accordance with equation (2.2). More details can be found in an on-line field-of-view guide (`www.evlbi.org/user_guide/fov/fov06.html`). Of course, one potential drawback to such wide-field correlations (with the short $t_{\rm int}$ and large $N_{\rm frq}$ they require) is the rapid growth of the size of the FITS file seen by the user—reaching about 7 GB per hour of observation at our current maximum output rate. The current record for output FITS files for a single experiment in one session stands at 674.9 GB.

BW_{sb}	Modes		
4 MHz	256–16–2		
8 MHz	256– 8–2	512–16–2	
16 MHz	256– 4–2	512– 8–2	1024–16–2

Table 2: Recording modes with "unused" subbands available for a third MERLIN station.

There is also the perhaps non-intuitive conflict between Gbps recordings and FoV_{BW} arising from the current maximum BW_{sb} of 16 MHz: in order to achieve Gbps data rates, $N_{sb} \cdot N_{pol}$ in equation (2.1) will be 16 (no cross-pols) or 32 (with cross-pols). Thus an 8-station Gbps experiment can get at most $N_{frq} = 128$ in a single correlator pass, and $FoV_{BW} \sim 6.'6/B$, with B still in units of 1000 km (N_{frq}, FoV_{BW} halved if using cross-pols). This of course will have consequences for combining sensitivity and wide-field surveying in an optimal way.

3. A Third MERLIN station in EVN correlations

Because there are only two VLBI recorders at Jodrell Bank, to date only the Jb–Cm baseline has been in common in both the EVN and MERLIN correlations for EVN + MERLIN observations. This can sometimes lead to difficulty in tying the two *u*-*v* datasets together. Now, because of the 16 MHz-per-polarization limit in the MERLIN out-station micro-wave link, Cm has "unused" subbands in several recording modes. Earlier this year, Jodrell recorded signals from both Cm and Darnhall (Da) onto a single disk-pack, placing the subbands from Da containing actual data into subbands not used by Cm. Through a few unorthodox steps, we were able to correlate both Cm and Da as separate stations. Table 2 lists the recording modes for which "unused" subbands exist per BW_{sb}. For any of these you could get 3 MERLIN stations (Jb, Cm, another out-station) in common in both the EVN and MERLIN correlations, which should increase the robustness with which you can tie the resulting *u*-*v* datasets together. When proposing, you should still request EVN + MERLIN as you do now; if your experiment can incorporate a third MERLIN station in its EVN correlation, the pre-scheduling liaison will discuss this further with you.

Acknowledgments

The European VLBI Network is a joint facility of European, Chinese, South African and other radio astronomy institutes funded by their national research councils. This research was supported by the European Commission's I3 Programme "RadioNet", under contract No. 505818.

References

[1] C.R. Reynolds, Z. Paragi, & M.A. Garrett, *Pipeline Processing of VLBI Data*, in proceedings of *The 27th URSI General Assembly*, URSI, Ghent, J8.P.4, 2002.

[2] J.M. Wrobel, *VLBI Observing Strategies*, in *VLBI and the VLBA*, eds. J.A. Zensus, P.J. Diamond, & P.J. Napier, ASP, San Francisco, 411, 1995

PoS(8thEVN)054