VLBA Operations Memo 58

# Plan for VLBA RFI monitoring

Walter Brisken

8 Mar. 2021

**Abstract** This memo describes a new series of radio frequency interference (RFI) monitoring observations for the VLBA. While the primary goal is amassing data to explore long term trends, benefits to users on a shorter timescale are anticipated. This memo describes thoughts on the obervations to be taken, their cadence, data processing and curation.

## 1 Background

On 2020 July 07 a VLBA Test meeting was devoted to the subject of a new systematic approach to RFI monitoring that is well optimized for current VLBA observing capabilities (RDBE and Mark6 recorders) and that can be easily extended when new capabilities are rolled out. The plan presented here is heavily based on the discussion. It should be noted that RFI is a multi-dimentional problem (source location, frequency, timing, . . . ) and no single efficient observing program can provide a comprehensive data product. Additionally resources are thin and VLBA staff are already over-committed. The plan presented here aims to maximize impact while minimizing cost. Doing this requires exploiting the strengths of the VLBA in its current form. It should be pointed out that the group that met had high enthusiasm for the topic and a lot of ideas, some of which were at odds with others. In this effort I have tried to develop a plan that captures much of the discussion but remains practical to implement in the current conditions.

## 2 Goals and requirements

The primary objective of this effort is to develop a data set that is capable of quantitatively assessing the evolution of the RFI environment within VLBA observing bands over timescales of months to years. Such a data set may inform VLBA development priorities and may be useful in negotiations with regulatory agencies and establishing causality between equipment deployment and/or spectrum policy changes and the RFI environment faced by VLBA users. A strong secondary objective will be to provide timely guidance to VLBA users a means of minimizing likelihood of impact to scientific observations.

The RFI observing program is designed to meet the following requirements to maximize its utility:

- The observations should be flexibly schedulable by VLBA Operations to make use of gaps in observing time. Scheduling blocks as short as 1 hour can be supported.
- Observations should be made at a range of local clock times.
- A wide range of antenna pointings (elevations and azimuths) consistent with typical pointing directions of the antennas should be exercised.
- The entire spectrum tunable by the VLBA shall be monitored. Since nearly half of the tunable bandwidth is at W-band (80-96 GHz), which is rather infrequently used, an initial exhaustive frequency scan will be performed at this band, but only key observing frequencies (e.g., near the SiO line frequencies) will be monitored on a regular basis. Perhaps annually a full W-band sweep will be performed. Both native (circular) polarizations should be observed.
- Frequently one or more of the 88 VLBA receivers will not be fully functional. Tests on a subset of bands and/or antennas should be schedulable.

- Tests should be performed in interferometric mode. This maximizes utility of the data. For practical purposes (e.g., allowing observing at wide range of directions), the VLBA could be divided into eastern (SC, HN, NL, FD, PT) array and western (MK, BR, OV, KP, LA) subarrays.
- All data processing should be derived from VLBA correlator products that end up in the VLBA archive. This strategy will allow observations that were not intended for RFI monitoring to be added to the database.

While this document aims to guide development of observations that will be useful for long term monitoring of the VLBA RFI environment, it is expected (and encourated) that the person or group implementing these observations make sensible decisions which could be contrary to some of the thoughts presented here.

# 3 Implementation

This section describes the recommended observing and data reduction strategies.

## 3.1 Observing files

Observations will be scheduled using `sched` based on `.key` files that are either automatically generated or readily hand-derived from templates. Now that the VME systems are retired, there is no obligation to conform to experiment codes that are only 7 characters in length; here it is suggested to name the files RFI<*date*><*epoch*>. For example, `RFI20210122A` or `RFI20220902B`. The epoch code would be used to allow for the possibility of scheduling multiple observations in one day. Since it is likely that the entire span of a band will be covered in one observation, and data from different observing bands can be easily separated, it may make sense to label each output file with a band designation.

Depending on the observing time available in each slot, one more observing bands would be observed in one scheduling block. Table 1 enumerates the approximate number of tunings required to span the entire VLBA tuning range.

| Band | Tunings | Bandwidth (MHz) | Pointings | Duration (min) |
|---|---|---|---|---|
| P | 1 | $2 \times 64$ | 6 | 9 |
| L | 2 | $4 \times 128$ | 6 | 18 |
| S | 1 | $4 \times 128$ | 6 | 9 |
| C | 8 | $4 \times 128$ | 4 | 48 |
| X | 2 | $4 \times 128$ | 4 | 12 |
| U | 8 | $4 \times 128$ | 4 | 48 |
| K | 10 | $4 \times 128$ | 4 | 60 |
| Q | 12 | $4 \times 128$ | 4 | 72 |
| W | 32 | $4 \times 128$ | 4 | 192 |

Table 1: It is assumed that each scan (one source with one tuning) is allocated 90 seconds (about 60 seconds on source, plus some time for setup and slew). The 50cm band can be observed simultaneously with P-band. The bandwidth specified is for one polarization; double the number of bands for full polarization. With the exception of the P-band observing all observations will be performed at 4 Gbps record rate (8 channels of 128 MHz), always in dual polarization. Note that P-band includes both the 330 MHz and 610 MHz bands which can be observed simultaneously.

A full schedule of observations would require nearly 8 hours. In cases where only two setups were observed at W-band (the "abbreviated schedule") the schedule could be executed in about 5 hours. It is recommended here that the the abbreviated schedule be executed two or three times per month at different times of day and that a full schedule be executed once or twice per year. While it is possible to break the schedule down into smaller blocks, it is not recommended to perform partial-band observations; that is, all 8 C-band tunings should be executed at the same time. Weather should not be considered a strong constraint in observing.

## 3.2 Source considerations

The data reduction plan (see Sec. 3.5.1) specifies the option to make use of interferometrically-determined bandpass calibration. To this end, at least one of the sources observed for each tuning should be a strong source, ideally one of the identified "fringe finder" sources. It is recommended that at least three addtional sources, widely separated on the sky, are also observed.

To minimize slew time overhead, and to maximize time-diversity of observing a specific setup, it is suggested that all tunings within a scheduling block observe one source before moving on to the next one. This strategy also suggests that combining the P-, L-, S-, and X-band observations into one atomic observing block make sense. That observing block can be completed in about an hour.

## 3.3 Frequency setups

An effort to develop observing script templates to aid in receiver performce checkout is proceeding in parallel with this RFI monitoring effort. This "wide-band checkout" project also aims to observe the entire tuning range of a recever that may have been recently serviced. It is recommended that the two projects make use of the same sets of setup files. While the wide-band checkout files may be able to contribute to the RFI data sets, the goals of that effort differ sufficiently to treat its observatations and those from RFI monitoring as separate.

## 3.4 Correlation

A standard correlation will be performed on collected data. Full polarization products will be produced Nominal time averaging will be 1 second. Raw correlation spectral resolution of 10 kHz or finer will be used, with averaging to result in 100 kHz final spectral resolution with reasonably sharp spectral response.

## 3.5 Data processing

An AIPS pipeline will read the data and produce required data products based on the autocorrelation spectra. The pipeline will run automatically without human intervention. Standard amplitude calibration will be performed to yield noise spectra calibrated to approx. 10% accuracy. The output of the AIPS pipeline will be one text file for each observing setup containing various spectral statistics. A small set of Linux command-line programs will be developed to combine data from multiple tunings into full-band spectra and generate a web-viewable presentation of the most current data.

### 3.5.1 AIPS pipeline and data products

This section describes a possible concept for an AIPS pipeline. It is expected that during the development of the pipeline, and after some initial use, variations in function will be adopted.

The AIPS pipeline will run separately on each tuning configuration observed at a specific epoch. This one configuration will in general have multiple scans, each on a different source representing a different part of the sky. The output will be a text file with various statistics from each spectral point, plottable as a spectrum. The pipeline will perform the following steps:

1. Load the data.

2. Perform amplitude and bandpass calibration. This should follow the prescription established in VLBA Scientific memo 37[1]. This strategy will require at least one of the observed sources be suitably strong to serve as a bandpass calibrator. The interferometically-determined bandpass calibration will largely

[1] Available at: https://library.nrao.edu/public/memos/vlba/sci/VLBAS_37.pdf .

(but maybe not completely) be unaffected by moderate RFI; true RFI (either local to a site or far from the pointing direction) will dominate over any bandpass supression that is incurred. Interferometric band-pass calibration strategy is not likely effective at W-band, and it will likely fail at K-band and Q-band in bad weather conditions. Uncorrected bandpasses, or perhaps historically-determined default bandpasses, should be used in these cases with some resultant penalty in data interpretation.

3. Execute a RFI statistics task (to be developed) to generate summary statistics on a per spectral point basis.

The `VLBARUN` AIPS procedure implements the required calibration steps and may be an excellent starting point for an AIPS pipeline.

The spectral summary statistics envisioned at this time are to be formed as follows. First compute two summary statistics spectra for each scan. The first is the mean power (or spectral power density) as a function of frequency, or in other words, a scan-averaged spectrum. E.g., for antenna $a$, scan $s$, and polarization $p$, the average power is

$$P^p_{a,s}(\nu) = \frac{1}{n}\sum_{t=1}^{n} V^p_{aa}(t,\nu), \tag{1}$$

where $t$ is the time index, of which there are $n$, and $V^p_{aa}(t,\nu)$ is the calibrated autocorrelation value at that time at frequency $\nu$. The value should be calibrated and represent the system equivalent flux density (SEFD) at that frequency. Note that in general this will exceed the catalogued zenith SEFD because these will be observed at elevations below 90°, incurring increased spillover power and airmass. The second is the modulation index, defined as the RMS power divided by average power as a function of frequency, over one scan. This is computed as:

$$M^p_{a,s}(\nu) = \frac{\sqrt{Q^p_{a,s}(\nu) - P^p_{a,s}(\nu)^2}}{P^p_{a,s}(\nu)}, \tag{2}$$

where the quantity $Q^p_a(\nu)$ represents the mean squared power

$$Q^p_{a,s}(\nu) = \frac{1}{n}\sum_{t=1}^{n} V^p_{aa}(t,\nu)^2. \tag{3}$$

All sums are performed over the $n$ visibility records for scan $s$. The final summary spectra come from combining the statistics from the multiple scans that were observed with the same tuning. From these two quantities determined for each scan, compute the mean, min, and max values:

$$\begin{aligned}
P_{a,\mathrm{mean}}(\nu) &= \frac{1}{n_s n_p}\sum_{s,p} P^p_{a,s}(\nu), && (4)\\
P_{a,\mathrm{min}}(\nu) &= \min_{\{s,p\}} P^p_{a,s}(\nu), && (5)\\
P_{a,\mathrm{max}}(\nu) &= \max_{\{s,p\}} P^p_{a,s}(\nu), && (6)\\
M_{a,\mathrm{mean}}(\nu) &= \frac{1}{n_s n_p}\sum_{s,p} M^p_{a,s}(\nu), && (7)\\
M_{a,\mathrm{min}}(\nu) &= \min_{\{s,p\}} M^p_{a,s}(\nu), && (8)\\
M_{a,\mathrm{max}}(\nu) &= \max_{\{s,p\}} M^p_{a,s}(\nu). && (9)
\end{aligned}$$

The power spectrum statistics will provide some indication of how variable interference is on timescales of minutes and/or across different pointing directions. The modulation index statistics will provide some indication of variable interference is on timescales of seconds. It is unclear as of this writing how effective any of these summary statistics will be in predicting occurrence of RFI. Should improved statistics be developed, the pipeline can be rerun on existing data.

Note that "composite statistics" from multiple separate scheduling blocks can be computed from multiple sets of single-scheduling-block statistics. It might be effective to combine 3+ months of recent data when producing statistics to present to users.

### 3.6 Curation

No attempt will be made to preserve the pre-correlation baseband data. The correlator products will be-stored in the VLBA archive and made available to the public with no proprietary period. All primary data products will be stored on a filesystem that is automatically backed up. A good option for this might be in `/home/vlbiobs/rfi/`, with a separate sudirectory for each month, as is done currently for Open Skies observing files and for pointing data. The data files output from the AIPS pipeline should be named systematically so as to allow for easy determination of date and tuning. All source code used to manipulate the data (including the AIPS pipeline) will be stored within NRAO's source code repository (either in Subversion or Git).

## 4 Acknowledgements

Many thanks to those contributing to the July 8, 2020 discussion. Eric Greisen provided some information on how AIPS may be used in this effort. Anthony Sowinski, Emmanuel Momjian and Claire Chandler provided useful feedback on the first circulated draft.