RFI Mitigation in the ngVLA System Architecture
ngVLA Memo #71

Rob Selina, Urvashi Rau, Rafael Hiriart, Alan Erickson

February 6, 2020

Abstract

The RFI environment is expected to be increasingly challenging in the ngVLA era, with more interference in the scientifically important 10-50 GHz spectrum. However, strategies exist to mitigate the impact of RFI at multiple points in an observation and in the signal chain. Existing efforts, focused on mitigation in the post-processing stage, are effective but would discard significant portions of the observed data, and could render some parts of the spectrum inaccessible. Real-time mitigation approaches have had limited success to date primarily due to their experimental scope and lack of broader integration into existing facilities. Here we describe a systematic approach to mitigating RFI, building upon these previous efforts while employing a system architecture that can fully leverage the available mitigation strategies, to ensure the system is resilient and that the scientific productivity of the array is maintained in the projected environment.

1 Introduction

A key system design requirement is to ensure that the array design is resilient to changes in its environment, with sufficient flexibility and robustness in its configuration so that the Observatory can assure operational continuity throughout the design life of the instrument.

RFI is presently a manageable problem at cm wavelengths. The 1-5 GHz band is increasingly crowded with local RFI sources, along with navigation satellite constellations and satellite down-links, whereas observations above 5 GHz at the VLA generally have fewer interferers. [9] Forward projections of the RFI environment, though uncertain, suggest comparable RFI extending up to 50 GHz in the worst case scenarios. [2]

Existing mitigation approaches enable high scientific productivity with the VLA in the 1 - 5 GHz band (with maximum losses of order 30% of bandwidth in continuum modes), and we might expect comparable productivity for the ngVLA extending up to 50 GHz, but there are additional concerns to consider given the evolving environment and the expected science use cases.

Expected changes in the environment include a proliferation of Low Earth Orbit (LEO) satellite constellations. [23, 22, 21] This presents a new risk to all ground-based radio observatories, including interferometers, as a number of these satellites are likely to be visible to multiple antennas at any given point in time, negating some of the natural interference protection afforded by aperture synthesis techniques. High imaging dynamic range observations may therefore be more challenging in this environment.

Furthermore, the ngVLA science case requires a high spectral dynamic range in support of KSG2 (blind searches for pre-biotic molecules) where a priori information about the spectral lines is not available. [8] Disentangling the local and orbital RFI from astronomical signals will be important for system scientific productivity. To mitigate these new environmental risks, a systematic approach to RFI mitigation is required.

2 RFI Environment

The RFI environment at the VLA site is periodically surveyed, and the most recent survey results can be found in the VLA Observer's Guide [9]. The anticipated changes in the RFI environment on ngVLA timescales are discussed in ngVLA Memo #48 [2].

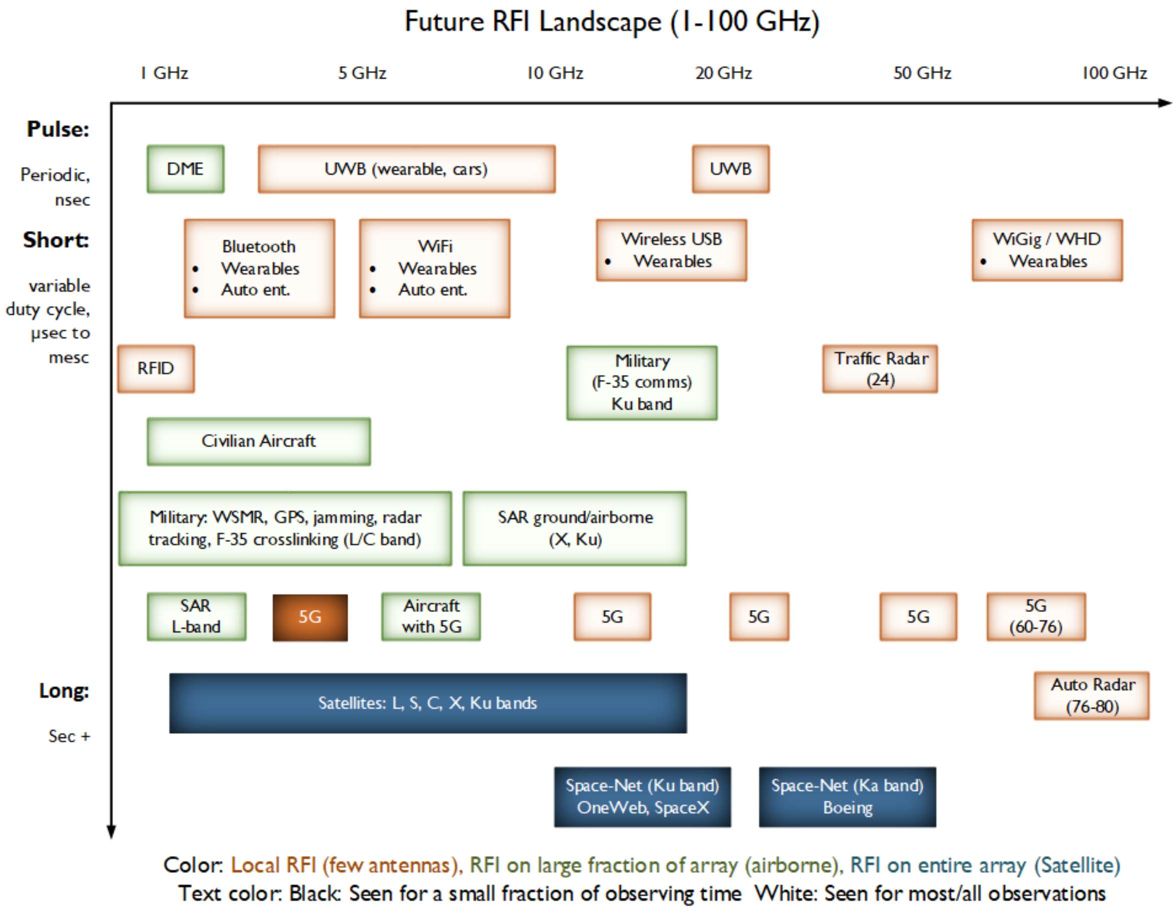

Figure 1: Diagram of projected RFI sources by frequency and transmission duration. Frequency spans are approximate. Note that this chart shows spectrum *allocations*, not use, which may vary at any given site.

Figure 1 summarizes the expected sources of interference on ngVLA timescales. Interferers are roughly sorted by frequency and the duration or duty cycle of transmission. Sources visible to large portions of the array (shown in white text) are most troubling, as these will correlate across many baselines and can corrupt the imaging process. Note that this figure approximately spans the frequency *allocations* provided to these services, but these allocations may not be fully occupied. Only a subset of these sources would be expected at a given point in time at any site.

Many of these sources are present (though not necessarily observed) today. The most notable expected changes are:

- *Satellites and Space Networks:* SpaceX, OneWeb, Amazon, Boeing, and others are planning large constellations of Low Earth Orbit (LEO) broad-bandwidth emitters that will be visible to many antennas in the array. [23, 22, 21]
- *5G Cell Bands:* 5G network allocations are wideband and appreciably higher in frequency than the present 3G and 4G network, extending above the crowded 1-5 GHz frequency space and spanning up into the O2 line at 60-70 GHz. Fortunately, the higher frequency allocations are not expected to be common at rural sites, but there is risk of the technology evolving to occupy these allocations. [18]
- *Automotive Radar:* Automotive radar is becoming a standard feature. It is high power, up to 50 dBm, and operates in the 76-80 GHz band. Care in site selection is required to avoid direct beaming, and in amplifier and feed design to ensure that the LNAs can withstand intermittent off-axis sources and reflections without damage.

3 Mitigation Approaches

A quantitative summary of the RFI mitigation strategies, their projected effectiveness at recovering time or bandwidth, and the associated cost-benefit analysis is described in ngVLA Memo #70 [19]. In this memo we will instead focus on the functional application of these strategies and their impact on the system design.

RFI mitigation strategies that are presently deployed or under development for ngVLA may be aggregated into three categories:

- *Avoidance:* Intelligent RFI avoidance strategies that adapt observing schedules to predictable and current RFI patterns.
- *Detection and Flagging:* Outlier detection algorithms that generate flags to mask RFI-affected data from further processing. [1, 3, 10]
- *Modeling and Subtraction:* Algorithms that attempt to recover sky signals from beneath the RFI. [6, 7]

We will summarize the approaches proposed at each relevant block in the proposed ngVLA technical concept. Implementation of these mitigation strategies aims to maintain the imaging dynamic range of the instrument and provide access to the full observable spectrum. However, even post mitigation, RFI will add to the system noise temperature in impacted channels. Process and policy based avoidance strategies are therefore still preferable to the signal processing techniques, and should be central to site selection, spectrum coordination, and facility operation and maintenance.

4 VLA & ALMA Implementation

We present the VLA and ALMA implementation of RFI mitigation as a baseline for comparison. This mitigation approach does not presently implement *Avoidance* or *Modeling and Subtraction* techniques as described in Section 3. Rather, the VLA and ALMA presently employ *Detection and Flagging* algorithms, applied to the visibilities at the recorded time and frequency resolution available to the

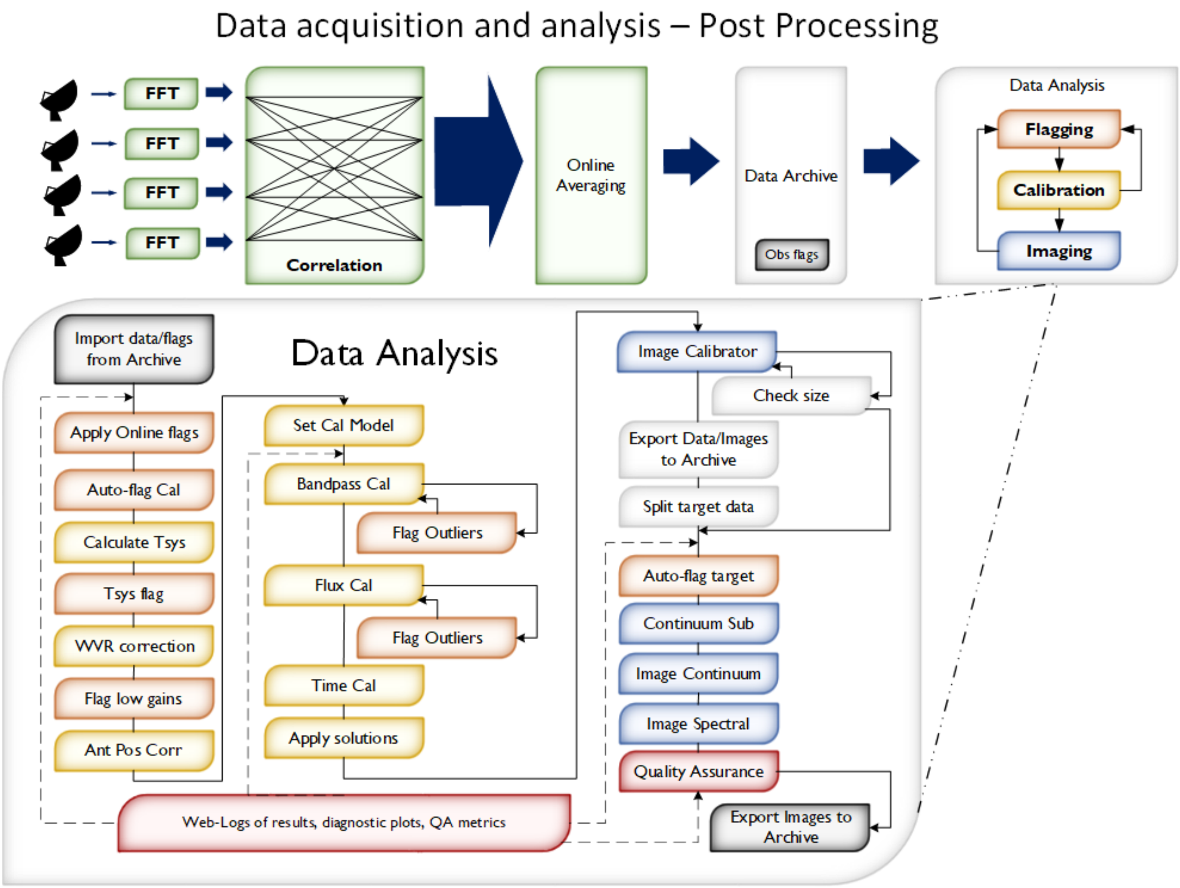

Figure 2: Implementation of detection and flagging algorithms in the ALMA data analysis system, employing automated analysis pipelines. The signal path through the system is shown at top, with an expanded view of the data analysis pipeline steps below. Flagging, calibration, and imaging steps are shown in red, yellow, and blue respectively.

post-processing system (as retrieved from the data archive). Figure 2 shows a graphical implementation of the detection and flagging steps that form part of the automated pipeline processing.

In this approach, no online flags are generated for RFI events. *Autoflag* seeks outliers, with no a priori knowledge of interferers previously identified and expected in the spectra. Automated pipelines have to strike a balance between under-flagging (leaving corrupting RFI in the imaged data set) or over-flagging, discarding useful data and reducing system sensitivity. The latter is generally a safer approach in automation.

Typical visibility archive rates may have a frequency resolution of $\sim$1 MHz and a time resolution of $\sim$1 second. Flagging of outlier visibilities post integration can result in higher rates of data loss (when signals are intermittent on shorter time scales or span narrower bandwidths), but such losses are not yet a significant detriment to observations at most frequencies.

This approach is adequate in the present VLA environment and operating model. If additional interference sources are present, the losses in sensitivity from discarding otherwise useful data may rise to a threshold that warrants more effort in processing.

The VLA operating model largely relies on expert users applying or tuning the flagging tools and iterating in the imaging process. This model makes up for the lack of awareness of the interferers in the system design, instead relying on the knowledge of the user in the application and tuning of the flagging tools. Such an approach is very effective with an expert user, but is less effective when automated for the production of higher level data products, since these flagging algorithms benefit from tuning to the specific sources of RFI present in the data set. The present approach does not provide for feedback and learning to improve automation. So, while this approach is effective today, developing robust flagging for an automated post-processing pipeline has intrinsic challenges and limits to its effectiveness. These challenges must be overcome in an operational model that aims to deliver *Science Ready Data Products* (i.e. high-level data products like image cubes) directly to users, in a more challenging RFI environment.

Finally, the extensibility of the current post-processing approach to new RFI sources may be limited. RFI from transiting orbiting sources (i.e., including the new Low Earth Orbit constellations) is partially decorrelated at these archived time and frequency resolutions. This makes it no longer feasible to model the signals sufficiently accurately to subtract them, limiting the application of these techniques.

5 ngVLA Technical Concept

The overall ngVLA system architecture and present design concept is described in [16] and the Reference Design volumes [14]. The ngVLA technical concept for RFI mitigation aims to improve upon the existing post-processing detection and flagging approaches, and aid in their automation, by storing a managed archive of known and identified emitters. Stored and retreived information about likely interference sources applicable to a specific observation (given the observation time, antenna pointing, and band selection) can be used to tune the detection and flagging algorithms to improve the detection ratio.

Knowledge of likely interferers can also reduce the risk of overflagging when implementing automated flagging and excision in the online system itself, excising offending visibilities at high temporal or frequency resolution, and adjusting the weight of the averaged visibilities accordingly.

Extending further upstream, detection algorithms can be applied to the antenna voltage streams directly pre-correlation. These voltage streams also provide high-temporal resolution information on the interfering signal properties, which can be an essential source of detection and characterization information to catalog these interference sources.

Once catalogued, interference sources can also be avoided by the observation scheduler, leveraging the same database of known interferers that tunes the flagging and excision algorithms, but supplemented by knowledge of source positions and their trajectories (for orbital RFI).

The most experimental solution may be interference modeling and subtraction, which could be applied after automatic flagging and excision in the correlator back end at ms-scale resolution, before any final averaging for data archiving.

Each of these approaches is best suited to a particular source or type of RFI, and in total they provide a robust mitigation system for most expected sources of interference. [19] The resulting signal path and associated RFI mitigation blocks can be seen in Figure 3.

The new RFI mitigation blocks for the ngVLA, shown in Figure 3, are as follows:

A *Outlier detection on the time series data:* Detect and mask nanosecond pulses seen by individual antennas. This mitigation block will be in the Digital Back End (DBE).

B *Outlier detection on the real-time spectrum:* Detect and mask or label corrupted channels per antenna. This mitigation block will be in the Correlator-Beamformer (CBF).

C *Automatic flagging on the visibilities at high time/frequency resolution:* Detect and mask correlated RFI in the baseline/time/frequency space. This mitigation block will be in the Correlator-Beamformer (CBF).

D *Interference modeling and subtraction:* Advanced interference excision or nulling using matrix

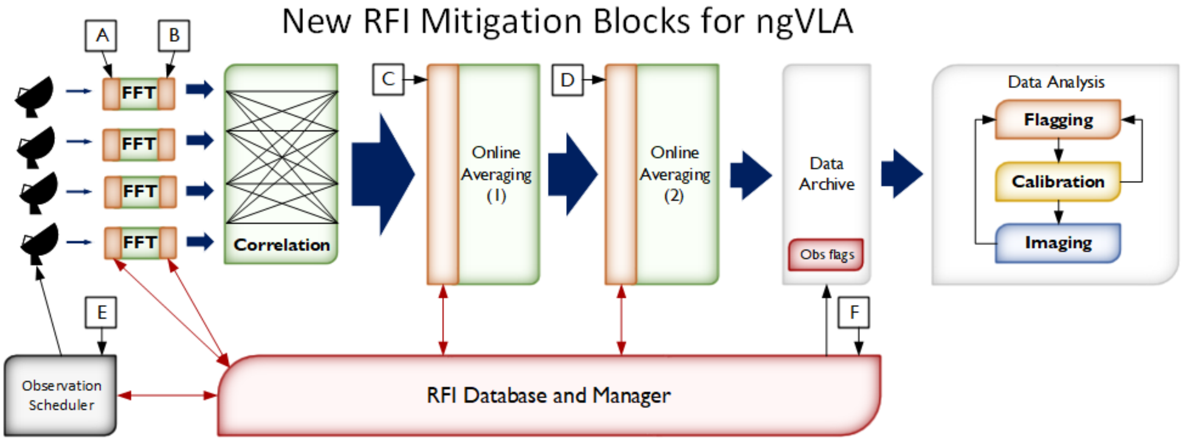

Figure 3: Implementation of new RFI mitigation blocks in the ngVLA technical concept. Blocks A through F are described in section 5. The detection and flagging in the data analysis post-processing system would largely match the existing VLA and ALMA implementation (Figure 2.)

subspace projection, real-time imaging, or source location. This mitigation block will be in the Correlator Back End (CBE).

E *RFI avoidance in the observation scheduler:* Analyze RFI metadata to inform observation scheduling as a function of frequency or sky position.

F *RFI database and manager:* Store RFI characteristics and metadata for reuse/retrieval. Analyze RFI metadata to inform the current observation. This will be part of the online software subsystem.

The detection and flagging in the data analysis post-processing system would largely match the existing VLA and ALMA implementation (see Figure 2.) However, *autoflag* could query the RFI manager for tuning parameters. The details of the ngVLA implementation are discussed in the following subsections, highlighting the type of RFI mitigated at each step along with design considerations relevant to each block and its interfaces. The broader ngVLA system architecture can be found in [16].

5.1 Scheduler

5.1.1 RFI Avoidance in the Observation Scheduler

The dynamic scheduler for the VLA accounts for conditions such as the system configuration (e.g., number of functional antennas at a given band), the weather and atmospheric phase fluctuations across the array (as measured by the Atmospheric Phase Interferometer), as well as project priority ranking. The ngVLA will also employ a dynamic scheduler that monitors equivalent information, but this implementation will account for additional RFI environment constraints, such as the positions and frequencies of known interferers, in determining the order of observations. Publicly accessible ephemerides, and previously observed transmission frequencies and signal properties, would be stored in the RFI Database for orbital RFI so that the array can be pointed preferentially at quiet locations in the sky.

Implementing this functionality does add complexity to the scheduler, and requires an interface to the RFI manager to retrieve the expected metadata and orbiting RFI ephemerides from the RFI database. However, similar concepts have been demonstrated to be effective at other facilities such as the GMRT. [11] Algorithmically, the RFI condition assessment could be treated similarly to system configuration and weather, where thresholds for acceptable RFI (e.g allowable proximity of a transiting LEO to the main beam during the observation) are established in the project observation metadata. This may be a more practical implementation than an attempt at minimization of RFI in an observation, but alternatives can be explored in the detailed design phase.

Another relevant design consideration is limiting the default size of a scheduling block, as this may facilitate scheduling around the transits of satellites in the planned LEO constellations. The existing ngVLA scheduling and data processing design allows for the calibration of visibilities using observations from other scheduling blocks (permitting efficiency enhancements like service observations), so a smaller default size for a scheduling block should not inherently reduce total observing efficiency, so long as continuing the active project has a high priority in the scheduler.

One final design consideration is an external interface to the scheduled observation stack, providing key observation metadata such as observation time, frequency band, and approximate pointing position. This information could be used to coordinate operations with external entities that transmit in the ngVLA-observed spectrum. While implementing such coordination is a future decision for Array Operations, providing the requisite interface or data stream in the system design is advisable.

5.2 Antenna Digital Back End

5.2.1 Outlier Detection on the Time Series Data

The Antenna Digital Back End (DBE) is the only place in the digital signal path where the full bandwidth of a given pair of I-Q digitizers is available as a concurrent stream. Therefore, it is the natural point in the signal chain to detect and flag broadband (wider than a 200 MHz sub-band) signals.

At the first filter bank in the DBE, the digitized voltage streams are split into subbands that span a frequency range of ∼200 MHz. Outlier detection algorithms such as spectral kurtosis filters can be run in real time, over the full bandwidth or in parallel on each sub-band (i.e. before or after the filter bank), to detect and generate flags that will accompany the real-time voltage streams.

The types of RFI expected to be caught at this stage include:

1 Broadband intermittent interference such as sparks and lightning.

2 Intermittent ultra-wideband (UWB) communication devices, such as wearable sensors.

3 Local persistent or intermittent RFI, such as nearby transmitters or cell phones.

4 RFI that might partially decorrelate, increasing the system noise, and then be undetectable on some baselines.

As noted in [19], the direct benefit of flagging and excision executed at this step is expected to be small in terms of the recovered time or frequency space, but detection at this stage serves two other key purposes:

1 To determine the duty cycle or natural frequency of an interferer, to populate the RFI Database with detected signal properties, and to tune associated downstream algorithms.

2 To inform any local excision algorithm prior to requantization for data transmission (applicable to antennas more than ∼300 km from the core, which rely on commercial networks for data transmission.)

Incorporating this functionality does require additional resources in the DBE as well as ancillary data streams that add complexity to the overall software and firmware implementation. Known impacts to the design to incorporate this functionality include:

1 Additional FPGA-based compute resources per antenna to perform the outlier detection. Depending on the implementation of these algorithms, they may also require data buffers (TBC).

2 Instantiation and transport of a parallel bit stream carrying flags from each antenna through to the multiply-accumulate (MAC) step of the correlator.

3 Excising outlier samples that would otherwise saturate a sub-band after requantization, possibly substituting the samples with synthetic noise. It must be possible to turn this feature on/off, and the record of these actions must be added to the online flags stream.

4 On demand, capturing a snapshot (∼1 sec) of the time series data and transmitting it over the M&C network to the RFI Manager for analysis. This function has ancillary value for maintenance diagnostics, and could be based on the "Oscilloscope function" described in the M&C requirements documents.

5 Enabling changes to the FPGA personality to be applied over the M&C network interface. Also being able to query the active FPGA personality over the same interface. The remote management of the FPGA personalities is desirable from a configuration management and maintenance perspective, but is especially prudent in the RFI context to enable updates to the detection, flagging, and excision blocks of the FPGA logic.

The development of the flagging algorithms for this stage of the signal processing path can benefit from work at other observatories. [20, 17] A group at GMRT have already demonstrated real-time flagging and excision of high power RFI by replacing RFI filled samples with synthetic noise, thereby replacing high power RFI with lower-power RFI, while maintaining the timing of the data streams. [5]

5.3 Correlator Beamformer

5.3.1 Outlier Detection on the Real-Time Spectrum

The next available processing step in the ngVLA architecture, employing an FX correlator, is after the F-engine. Some signals are more easily detected after the Fourier transform to frequency space, detecting outliers in the real-time spectrum from each antenna, prior to cross correlation. In particular, any strong narrow-band signal will be detectable in the spectrum.

The overall system design impact is very similar to the outlier detection on the time series data implemented in the DBE. The system would detect and mask or label corrupted channels on a per antenna basis, provide a record of these flags through the online flag stream to the multiply-accumulate stage of the correlator, and also transmit the associated metadata to the RFI manager. Flags would be applied downstream by leaving out affected values during the MAC step and/or generating weights along with each visibility.

5.3.2 Automatic Flagging on Visibilities at High Time/Frequency Resolution

Soon after the multiply-accumulate (MAC) step within the correlator, visibilities are available at a very high time and frequency resolution. These visibilities are typically immediately averaged down to the resolutions required for data transmission to the correlator back end (CBE). RFI excision and flagging prior to this averaging has the potential of handling intermittent RFI that is sparse on a μsec or $msec$ timescale but appears continuous when viewed at sec-scale time resolution (i.e. in post-processing flagging). This is also the stage where real-time antenna-based RFI information can be compared across antennas to derive useful information about RFI characteristics that may be used in order to automatically tune the algorithms at this and later processing stages.

In the ngVLA reference design, correlator read-out rates are limited to approximately 100 msec, to limit the data rate (over Ethernet) to the correlator back end (CBE). Therefore, any flagging at rates faster than ∼100 msec must take place within the correlator-beamformer. This boundary is a variable design constraint and can be adjusted, but is unlikely to be improved by more than 10^2 (at lower spectral resolution), so detection and flagging on time scales of μs to a few ms must architecturally reside within the correlator-beamformer. It is necessary to properly analyse RFI characteristics at these timescales to determine if there are enough gaps in between RFI signals to warrant the extra cost of such a real-time solution, as well as to determine whether there is a preferred time resolution at which data recovery is maximized.

Consequences for the correlator and system design include:

1 Additional compute (FPGA) resources within the correlator to run the detection and flagging (autoflag) algorithms. Algorithmic intensity estimates are available in [19].

2 Provide an interface to the RFI manager to receive autoflag tuning parameters.

3 Provide an interface to the flag data stream generated by the preceding processing steps in the antenna voltage stream (DBE) and antenna spectrum (Correlator F-Engine).

4 The flags, from all three processing steps, can be applied and new data weights evaluated while performing subsequent averaging to the post-processing time and frequency resolutions. An intermediate approach would be to use these flags to only re-adjust weights, but not actually discard data. It is desirable to support both modes for commissioning purposes.

5 The averaging step must be adaptive to the timescale required for optimal treatment of RFI, i.e. to match the cadence of the RFI. As RFI signals can vary in duration from μs to several $10ms$, an optional intermediate averaging stage is recommended in order to allow the choice of a long-enough sample window to cover RFI as well as clear samples.

6 Passing the real-time generated data weights to the CBE along with the averaged visibilities.

7 Enabling changes to the FPGA personalities to be applied over the M&C network interface. Also being able to query the active FPGA personality over the same interface. The remote management of the FPGA personalities is already a feature of the correlator to accommodate the various functional modes (e.g. beamforming and cross-correlation) on common hardware, but we note that this feature is also required in the RFI mitigation context.

As these real-time RFI mitigation approaches are currently experimental, it may be useful to design in the option of transmitting raw visibilities (i.e. no RFI excision; only averaging) concurrent with RFI-processed visibilities. This will double the data rate output from the correlator backend, but it may be required for telescope and algorithm commissioning, especially since the actual RFI environment seen by the ngVLA is not going to be 100% predictable beforehand. This dual recording mode would only be available within the constraints of the CBF to CBE data rate limit. I.e., the observation would have to run at less than half the full time and frequency resolution of the correlator to accommodate the dual streams within the data rate limit.

The types of RFI expected to be caught at this stage include communication signals that typically consist of few μs bursts spaced by a few μs, within ∼100 kHz channels, and frequency hopping within a known larger frequency range. Based on peripheral results from the RealFast and Fast-VLITE projects, there is considerable promise in being able to recover data in between these transmissions (perhaps even 50%) for situations where post-processing flagging would result in 100% data loss. The exact fraction of data recovered will depend on the fraction of the space filled with RFI at any given time.

5.4 Correlator Back End

5.4.1 Interference Modeling and Subtraction

As noted in the preceding section, the ngVLA reference design for the correlator has fast read-out rates of order 100 msec. The next RFI mitigation block would reside in the correlator back end (CBE), which is expected to be an off-the-shelf compute cluster.

This cluster could host the most experimental RFI mitigation approaches, emphasizing interference modeling and subtraction.[6] For these approaches, visibilities from all baselines need to be available together. Given the data aggregation requirements and the experimental nature, this is best handled in the CBE. Calculations with the VLA [7] suggest that for RFI correlated across the main array core, a time and frequency resolution of order 10-100 ms and 10-100 kHz will be required, which can likely be accommodated within the limitations in the CBF to CBE interface. A higher time resolution is not required to incorporate outlier antennas, as this approach is less suitable once RFI is attenuated by fringe washing on long baselines (which is expected to provide 30-45dB of attenuation, based on [12]). The less uniform distribution of antennas outside the Plains also limits the effectiveness of modeling and subtraction on these larger scales, so we focus on the application of these techniques within the core and plains subarrays only.

Consequences for the CBE and system design include:

1 Sizing the CBE cluster to allow for modeling/subtraction algorithms as well as algorithms to mine real-time antenna-based information for the RFI manager. Algorithmic intensity estimates are available in [19].

2 For modeling/subtraction approaches, there is no extra flag data to be carried along, although weights may be adjusted to account for potentially higher error on parts of the data where the most RFI has been excised.

3 Inclusion of an averaging (over time and frequency) block in the correlator back end to match the desired time and frequency resolutions for archiving and post-processing.

As with the real-time flagging, these approaches are currently experimental and it is desirable to have the option of recording raw visibilities (i.e. no RFI excision, but only averaging) as well as the RFI-processed visibilities. These needn't be archived long-term, but must be retained for a period to permit mode commissioning, and data quality checks in operations. The project can avoid long-term storage costs with such an approach, but the larger data buffer would still be required.

Satellite data transfers and cell phone data transfers (from airborne or local fixed towers) are likely to be continuous at this time resolution, but their phase coherence may be preserved. Modeling/subtraction algorithms may be the only viable option to prevent 100% data loss within channels exposed to such RFI.

5.4.2 Relation to Fast-Imaging Science Cases

A science use case that is currently not part of the KSGs but is to be designed into the system as a future option is real-time imaging to detect transients.

We note that this feature will have equivalent resource and data rate requirements as algorithms that model and subtract RFI in real time (Section 5.4.1) and can be designed in conjunction with similar RFI solutions.

RealFast, VLITE-Fast, EPIC (with LWA/HERA) have all demonstrated the ability to make all-sky or full-beam images of the sky at millisecond time resolution, and image the RFI either as satellites (or meteors) streaking across the sky, or terrestrial RFI located on the horizon. From this, one can make the argument that all such RFI is fundamentally separable from the sky signal.

Real-time RFI excision carries with it the risk of flagging astronomical transients (when sufficiently bright to be visible in single channels). This is where imaging-based methods can offer useful disambiguation.

5.5 Post-Processing System

5.5.1 Time-averaged Visibilities Saved to the Archive

At the post-processing stage, flagging algorithms will likely operate in a similar mode to current practice, with compute scaling dependent on the total number of visibilities presented to the algorithms, and with optimal parallelization achieved by strategic data partitioning.

Such an approach is still required, even with upstream high time/frequency resolution flagging, as this step can identify and mitigate different forms of RFI. Types of RFI likely to be caught only at this stage are:

1 Weak RFI that is not visible above the noise either in the antenna voltage streams or even in the visibilites at the CBE time and frequency resolutions, but visible only after combining several minutes or hours of data to build up robust sky statistics.

2 Signals that are continuous in time and that will not benefit from phase-signature modeling and subtraction.

The implementation of these flagging algorithms at this step are well understood, but we note that an interface should be provided to the RFI manager to aid in the tuning of these algorithms. E.g., a priori knowledge of weak RFI can help differentiate such signals from astronomical spectral lines.

5.6 RFI Manager

5.6.1 Monitoring, Mining, Avoidance, Removal

The RFI Manager serves a support function, aggregating information from the various detection and flagging steps, querying external sources of information (e.g., LEO satellite ephemerides), and storing this information in the RFI database. It also searches the database, based on matching parameter constraints, to implement avoidance strategies, and to tune real-time and post-processing algorithms.

High-level mitigation strategies enabled by the RFI manager include:

1 Identifying RFI characteristics by mining RFI detection results (flags) across antennas and baselines and use this information to tune *autoflag* and other algorithms in the data stream.

2 RFI avoidance via smart scheduling between known intermittent RFI emissions (in direction, frequency, and time), or in response to a real-time RFI detection.

3 Possible RFI avoidance via negotiated sharing of time/frequency space with key RFI generators. Given the nature of most RFI and the commercial implications of such an approach, this option may be limited.

4 The use of a reference array or reference antennas/receivers specifically for RFI monitoring and localization, removing the source of RFI when possible.

The last feature would require local dedicated site monitor stations and/or antennas, as is presently done with the VLA. This is a prudent feature for surveying the quiescent RFI background and detecting changes that may require a response from Array Operations or Spectrum Management.

In order to implement these mitigation strategies, the RFI manager subsystem will perform several functions:

1 Query the RFI database to predict RFI occurrences and provide this information to requesting systems, such as the scheduler.

2 Query external data sources for known ephemerides for orbiting RFI sources, updating the RFI database when appropriate.

3 Determine expected transits from identified orbital RFI sources that are relevant (e.g., due to sky position or frequency) to the next scheduling block.

4 Aggregate data from the flag data stream and other online data streams to populate the RFI database for future reuse.

5 Query the RFI database to predict RFI in a scheduled observation and to tune each detector, flagging, and excision module in the signal path.

6 Possibly instantiate and monitor RFI characterization modules in response to requests from the online system supervisor.

7 Possibly instantiate RFI detection, flagging and excision modules in response to requests from the online system supervisor.

The process of monitoring and aggregating the flag data streams to populate the RFI database for future reuse will require a degree of analysis. Local modules will be required, that are specialized to characterize RFI with a given set of properties, such as a pulse-emitter detector for detecting radars or aviation DME transmitters. Characterizers provide the RFI manager with an awareness of the RFI present in the current observation. The RFI manager may then take action, for example to adjust the tuning parameters for data flaggers or excisors, or otherwise alert the online system supervisor to identified problems with the observation.

These RFI characterizer modules must be extensible so that the system can grow to identify new sources of RFI over time. High certainty hits from the RFI characterizers would be logged to the RFI database for further analysis by Array Operations, Spectrum Management, and Engineering.

The RFI manager will record the tuning parameters applied to each observation in the system configuration database, ensuring that the impact on the data can be determined and data provenance be preserved.

5.7 RFI Database

The RFI database stores characteristics of both expected and observed RFI for reuse as well as to provide a long-term record of emitters visible to the array.

Expected sources are those whose impact on observations can be fully characterized in salient dimensions such as clock time, frequency, modulation, and emitter location. In most such cases, some expected source information is publicly or commercially available, e.g. satellite ephemerides or FCC transmit licenses. When trusted third-party information is available, that information will be incorporated by reference into the RFI database or by interfaces in the RFI manager. Expected sources have undergone a degree of verification by Array Operations.

Observed events are those that are not anticipated, but have been seen in the telescope data stream(s). These will be added to the database by the RFI manager through aggregation of the flags and detections generated from the voltage, spectral, or visibility domains (i.e. via the characterization modules described in Section 5.6). Periodically, these detections will be examined by scientists and engineers in Array Operations, and possibly elevated into the *expected sources* tables that are used for scheduling or tuning flagging algorithms.

5.8 Data Flows

The preceding sections focused on the functional requirements applicable to each block in the digital signal processing architecture and required interfaces. Here we summarize the data flows through the system necessary to support this functionality. The data flows and interfaces between these sub-systems are summarized in Figure 4.

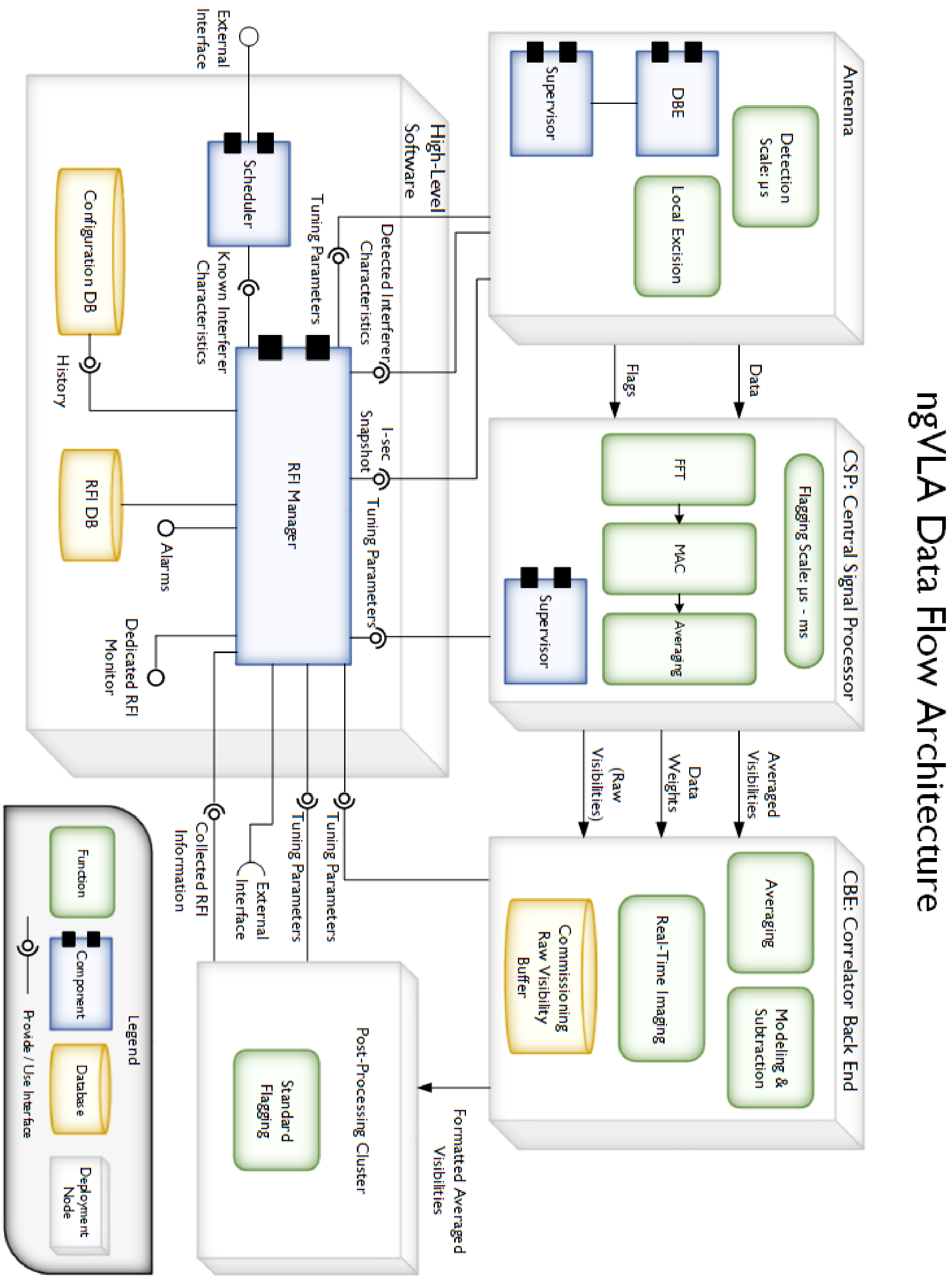

Figure 4: Detailed view of the data flows within the proposed ngVLA system architecture.

6 Other Considerations

While not signal processing blocks, we note a few additional system design and programmatic considerations that can aid in RFI mitigation for ngVLA:

- *Analog and Digital Dynamic Range:* The ngVLA architecture has adopted 8-bit sampling at all frequencies below 50.5 GHz. Ensuring adequate analog and digital dynamic range (i.e., not saturating the system) is a necessary precursor to implementing any of the detection and mitigation strategies discussed above. 20dB to 30dB of analog step attenuation is provided before the digitizers, depending on frequency. The 8-bit quantization, step attenuator range, and LNA dynamic range requirements are largely derived from VLA experience at low frequencies, but a more quantitative analysis should be performed by the system CoDR to ensure that sufficient analog and digital dynamic range is provided for each receiver band, given the expected RFI environment. A block diagram of the present implementation can be found in [4].
- *Avoidance of Self Interference:* As is best practice, ngVLA should avoid polluting its own environment with self-interference sources. The present approach to downconversion and digitization provides a degree of protection against this effect, using harmonically related local oscillator signals to avoid spurs and intermodulation products in each baseband, but care will continue to be required in system packaging and emission testing practices. The shielding specification can be found in [15].
- *Regulation:* The most sensitive area of the array to RFI is the main cluster of antennas at the VLA site on the Plains of San Agustin. The concentration of antennas at this site, with associated short baselines, means that many antennas could see correlated RFI from a single source. Legal protection of the site could provide a valuable additional tool for RFI mitigation. While such exclusion zones are desirable at all ngVLA antenna sites, a 50 km protection zone at the VLA could protect 80% of the ngVLA collecting area, and most short baselines, significantly mitigating the RFI risk to array operations.
- *Cooperation:* Cooperation with the LEO constellation developers, through enforcement of the astronomy protected bands, has proven effective to date and should continue. In particular, it would be desirable for satellite constellations to avoid forming beams pointed at ngVLA antennas to reduce the magnitude of the received RFI and possible system saturation.
- *Array Configuration:* The known locations of fixed emitters (courtesy of the FCC) are already a consideration in antenna site selection. Given the expected spot beam diameters for the LEO satellite constellations [13], placing ngVLA antennas at least 15 km line-of-site away from towns should enable spot beams on the town while placing the ngVLA antenna near a first null, limiting the magnitude of the RFI as described above, while reducing the likelihood of conflicts amongst stakeholders.

7 Conclusion

The proposed approach is expected to provide robust RFI protection for the ngVLA with a flexible and extensible architecture. The technical implementation is expected to evolve with its environment over the life of the instrument, with detection algorithms tuned to new emitters as they are identified and characterized.

A summary of the proposed implementation, the RFI mitigated at each step, the preferred detection methods, and associated mitigation actions is shown in Figure 5. Adopting such a strategy should enable the ngVLA to recover observational time and bandwidth across the spectrum, while ensuring that imaging artifacts are minimized, given the expected changes to the RFI environment of the ngVLA and other ground-based facilities.

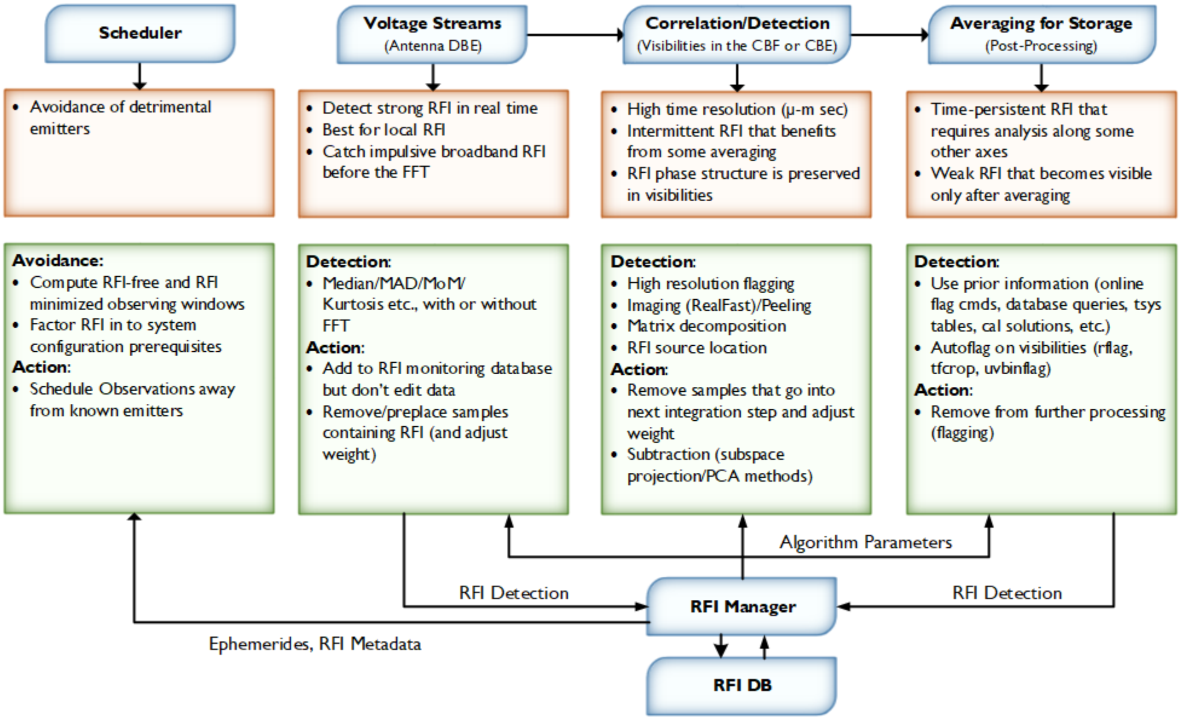

Figure 5: Summary of RFI mitigation blocks in the ngVLA system architecture. The various blocks and signals of the ngVLA signal chain are shown, along with the RFI they are best suited to mitigate, the associated detection strategy, and action at each step.

References

[1] CASAdocs. Flagging visibilities and calibration tables. https://casa.nrao.edu/casadocs/casa-5.6.0/calibration-and-visibility-data/data-examination-and-editing/flagging-measurement-sets-and-calibration-tables, 2019.

[2] A. Erickson. ngVLA radio frequency interference forecast. *ngVLA Memo Series, Memo #48*, 2018.

[3] E. Greisen. RFLAG. http://www.aips.nrao.edu/cgi-bin/ZXHLP2.PL?RFLAG, 2019.

[4] J. Jackson and R. Selina. ngVLA: antenna electronics front end enclosure block diagram. *ngVLA Doc. # 020.30.00.00.00-0002-BLK*, 2019.

[5] K. D. Buch, K. Naik, S. Nalawade, S. Bhatporia, Y. Gupta, B. Ajithkumar. Real-time implementation of MAD-based rfi excision on FPGAs. *JAI Special Issue : Interference Mitigation Techniques in Radio Astronomy - Volume 08, Issue 01*, 2019.

[6] M. C. Burnett, B. D. Jeffs, R. A. Black, K. F. Warnick. Subarray processing for projection-based RFI mitigation in radio astronomical interferometers. *ngVLA Memo Series, Memo #38*, 2018.

[7] M. Lambert, B. D. Jeffs. Radio frequency interference cancellation, appraisal, detection, and correction. *ngVLA Memo Series, Memo #56*, 2019.

[8] ngVLA Science Advisory Council. Key science goals for the next generation very large array (ngVLA). *ngVLA Memo Series, Memo #19*, 2017.

[9] NRAO. VLA observer's guide. https://science.nrao.edu/facilities/vla/docs/manuals/obsguide/rfi, 2019.

[10] A. Offringa. AOFlagger. https://ui.adsabs.harvard.edu/abs/2010ascl.soft10017O/abstract, 2010.

[11] P. Raybole, S. Sureshkumar, S. Katore, S. Rai. Real time prediction, detection and co-existing with satellite interference at GMRT. *IEEE RFI Conference Proceedings*, 2016.

[12] R. Perley. Attenuation of radio frequency interference by interferometric fringe rotation. *EVLA Memo #49*, 2002.

[13] I. Portillo. A technical comparison of three low earth orbit satellite constellation systems to provide global broadband. http://www.mit.edu/~portillo/files/Comparison-LEO-IAC-2018-slides.pdf, 2018.

[14] R. Selina et. al. ngVLA: system reference design, volumes i, ii, iii. https://ngvla.nrao.edu/page/refdesign, 2019.

[15] R. Selina. ngVLA: system L1 EMC compatibility and RFI mitigation requirements. *ngVLA Doc. # 020.10.15.10.00-0002-REQ*, 2019.

[16] R. Selina and R. Hiriart. ngVLA: preliminary system architecture. *ngVLA Doc. # 020.10.20.00.00-0002-DWG*, 2019.

[17] T. Gunaratne, B. Carlson and G. Comoretto. Novel RFI mitigation methods in the square kilometre array 1 mid correlator beamformer. *JAI Special Issue : Interference Mitigation Techniques in Radio Astronomy - Volume 08, Issue 01*, 2019.

[18] Telefonaktiebolaget LM Ericsson. 5G spectrum: strategies to maximize all bands. https://www.ericsson.com/en/networks/trending/hot-topics/5g-spectrum-strategies-to-maximize-all-bands, 2019.

[19] U. Rau, R. Selina, A. Erickson. RFI mitigation options for the ngVLA - a cost benefit analysis. *ngVLA Memo Series, Memo #70*, 2019.

[20] W. Baan, P. Fridman, R. Millenaar. Radio frequency interference mitigation with real-time FPGA digital signal processing. *IEEE Proceedings, 12th European Signal Processing Conference*, 2004.

[21] Wikipedia. Amazon - kuiper systems. https://en.wikipedia.org/wiki/Amazon_(company)#Kuiper_Systems, 2019.

[22] Wikipedia. Oneweb (satellite constellation). https://en.wikipedia.org/wiki/OneWeb_satellite_constellation, 2019.

[23] Wikipedia. Starlink (satellite constellation). https://en.wikipedia.org/wiki/Starlink_(satellite_constellation), 2019.