ngVLA Memo No. 56

Radio Frequency Interference Cancellation,

Appraisal, Detection, and Correction

(*Draft thesis version submitted to the NRAO ngVLA Memo Series as final report:*

ngVLA Community Study, AST-1519126, BYU sub-award)

Michael J Lambert

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Master of Science

Brian D. Jeffs, Chair
Karl F. Warnick
Michael Rice

Department of Electrical and Computer Engineering

Brigham Young University

ACKNOWLEDGMENTS

I would like to thank my wife for her patience and support during my research. Without her none of this would have happened.

This work was made possible by a grant from the National Radio Astronomy Observatory. Furthermore, I wish to thank the NRAO staff in Socorro, New Mexico, for their support and assistance. Special thanks go to Dr. Urvashi Rao for mentoring and support during a stay in Socorro.

TABLE OF CONTENTS

CHAPTER 1. INTRODUCTION

Astronomy studies the distant cosmos using the electromagnetic spectrum. While optical astronomy is the most common form brought to mind, information can be found at all wavelengths, from the visible spectrum up in energy to ultra-violet, x-ray and gamma ray photons as well down into the infra-red, microwave, and radio spectra. Optical astronomy is hindered when urban light pollution increase sky-brightness, radio astronomy must contend with man-made terrestrial and Earth orbiting satellite broadcasts in the band of interest. In general, man-made broadcasts that interfere with radio astronomy are termed radio frequency interference (RFI).

The predominant methods for dealing with radio astronomical RFI are to locate observatories far away from sources of man-made radio signals and to discard any measurements corrupted by RFI. While in the early years of Radio Astronomy the first method achieved great success, Earth-based radio sources are spreading such that there are very few suitably uninhabited locations to site new observatories. Also, as new radio-astronomical instruments are built using long baselines (the distance between antennas), the required quiet areas are becoming large and thus hard to find. As for the second method, as sources of RFI proliferate, the uninfected bandwidth and observing time windows are decreasing. Even were a large, uninhabited location found for a radio observatory, it would still have to contend with radio interference from satellite down-links. Although many satellites are transitory interferers, geostationary satellites increasingly interfere with radio astronomy observations over the entire Earth, particularly of sources near the celestial equator.

To probe signals of ever decreasing strength (or even of constant strength) in the current regime of ever increasing RFI, new methods must be employed. Subspace projection, wherein observed data samples are digitally modified so that interference components are attenuated as possible, offers significant advantages.

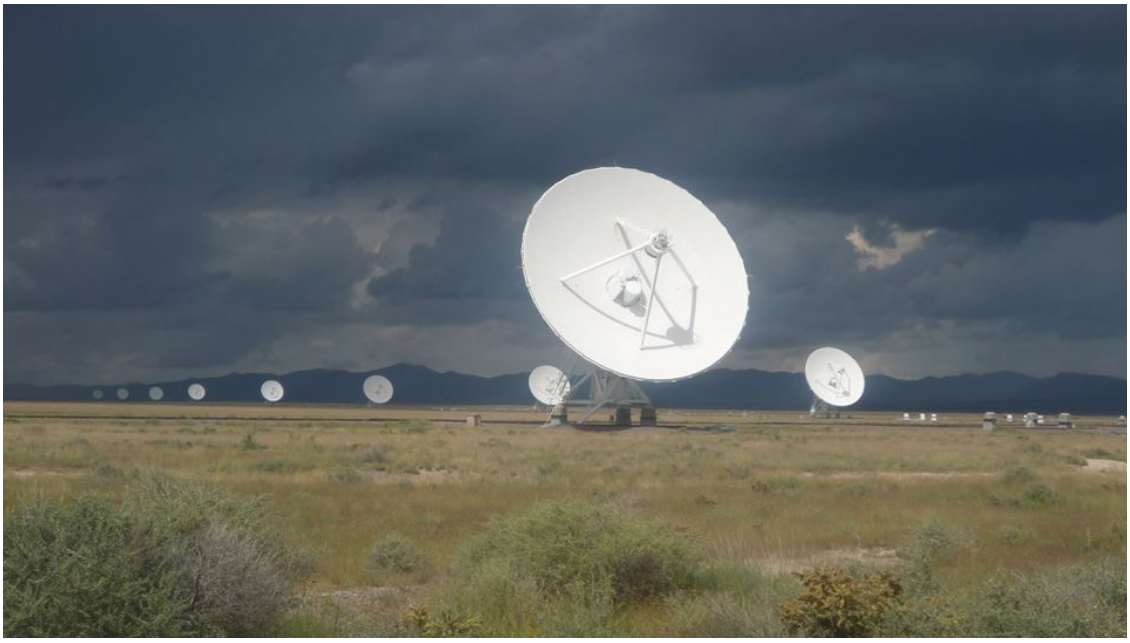

Figure 1.1: Image of the VLA north arm in C-configuration taken September 2013. Photo by Virginie Montes.

Of particular interest are locations where interference has become ubiquitous. For example, the celestial equator is completely unobservable in some bands due to the prevalence of geo-stationary satellites. With subspace projection, the prospect of ”peering behind” the satellite band would be an attractive possibility.

1.1 Radio Frequency Interference Mitigation

While where a radio telescope is built cannot be changed easily, eventually RFI enters the region. There are three standard methods used to remove corruption from RFI [2]. First is to either completely discard corrupted data, or to replace corrupted data with an estimate of what it should be based on the uncorrupted data [3] [4]. These estimates can be either a clipping filter or a scaled version of the median or mean. Second, if the RFI is consistent and easily localizable, an adaptive filter can be established to remove the RFI from the astronomical data. This requires either a low noise copy of the RFI or generating a low noise copy locally by detecting and estimating the RFI's parameters. [5] [6] Finally, there are spatial projection methods which manipulate the effective antenna beam pattern to remove the effects of RFI on the data. This last method is where we will spend all of our energy.

The seminal paper on RFI mitigation by subspace projection was written by [7], discuss their efforts to apply projection based methods to data from the Westerbork Synthesis-Array Radio Telescope (WSRT).

1.2 Thesis Statement

This thesis examines applying spatial projection based methods to data from the Very Large Array (VLA). While projection based RFI mitigation methods have been applied to a data from some arrays, they have never been applied to actual data from the VLA. This thesis therefore applies projection based RFI mitigation methods to data, both simulated and real, from the VLA. It examines the effects of time-windowing, channel bandwidth, and angular separation between interferer and center of focus.

1.3 Thesis outline

In chapter 2 we will discuss the mathematical and theoretical basis of radio astronomy. Chapter 3 explores simulations of our application of the method of subspace projection. Chapter 4 details application of our method to the VLA with real data. Primary concern is given to data collected when the array was in a compact configuration, but consideration is also given to data collected when the VLA was in its most expansive configuration. Chapter 5 features a conclusion and some proposals for future work.

CHAPTER 2. MATHEMATICAL BACKGROUND

To properly remove RFI corruption from data, we must first have an accurate signal model for both the RFI and the astronomical signal. We will assume that the corresponding received time-series data sequences are mutually statistically independent. We will further assume that these two signal sources add linearly (both in free space and on our receiving equipment). Finally, we will assume that the RFI originates from outside the receiving array. While the model here is applicable to any imaging array interferometer, we will be using the VLA geometry extensively.

We will assume that all electromagnetic waves we are dealing with are plane waves arriving from point-like far field sources, with time harmonic electric fields described by an equation of the form

$$E(x,y,z) = E_0 e^{j2\pi(k_x x + k_y y + k_z z)}. \tag{2.1}$$

Here E is an electric field, x, y, and z, are rectilinear axes in space, and k_x, k_y, and k_z are spatial frequencies associated with directions x, y, and z respectively. Spatial frequency is related to wave length λ by $k = \frac{2\pi}{\lambda}$. Spatial frequency is related to time frequency by $v = \frac{f}{\lambda} = \frac{fk}{2\pi}$, where f is the time frequency in Hz and v is the speed of propagation in the medium. For electromagnetic waves in free space, $v \approx 3.00 * 10^8$ m/s.

2.1 Signal Model

First, we shall define the voltages measured by the various ADCs as a vector, $\mathbf{x}[t]$, with elements $x_i[t]$, where i is the antenna whose ADC we are sampling. In the case of the VLA, this ADC is 8 bit sampling at 1 gigasample per second the voltages produced by the 27 antennas, each in 2 polarization (left and right circular). Included are the desired signals from deep space as well as the undesired signals from RFI and noise from atmospheric and warm ground blackbody radiation and thermal electronic noise from the various analog pieces before the ADC. Thus we

will say that

$$\mathbf{x}[t] = \mathbf{x}_s[t] + \mathbf{x}_{RFI}[t] + \mathbf{x}_{noise}[t], \tag{2.2}$$

where $\mathbf{x}_s$[t] is the voltage caused by the sky source, $\mathbf{x}_{RFI}[t]$is the voltage caused by RFI, and $\mathbf{x}_{noise}[t]$ is the voltage caused by all the various noise sources in the system. $\mathbf{x}_{noise}$ could reasonably be divided into noise caused by the sky (or ground) and noise introduced after the antenna, but we will leave it as a single entity. Furthermore, we will assume that while $\mathbf{x}_s[t]$ and $\mathbf{x}_{RFI}[t]$ are correlated across the array, we will assume that $\mathbf{x}_{noise}[t]$ is independent for each antenna.

Now we have a time varying signal for each antenna, but we would like a time varying relationship between each pair of antennas. We would also like to have an idea of how much of each frequency is involved in our signal. There are two, ultimately equivalent options to pursue. We will therefore define $p_{i,j}[t,s]$ as the correlation in the signal between the i^{th} antenna at time t to the signal in the j^{th} antenna at time s. A sample correlation can be mathematically defined as

$$p_{i,j}[t,s] = x_i[t]x_j^*[s], \tag{2.3}$$

where * denotes the complex conjugate. We can condense our math using vectors and matrices, leading to

$$P[t,s] = \mathbf{x}[t]\mathbf{x}^H[s]. \tag{2.4}$$

If we then perform a channelization algorithm, such as a fast fourier transform (FFT) or polyphase filter bank (PFB), we can examine spectral content instead of time behavior. Thus we arrive at $r_{i,j}[t,s,k,l]$, where i and j index the antennas, t and s index time indices, and k and l index spectral channels.

While it is right and proper to have defined our signal thus, there are several non-sensical trivialities that we can discard before proceeding. First, interferometers rely on the minute propagation delays between different paths for the same signal. Since signals from deep space are generated by wide sense stationary (WSS) random processes, their correlation between times is nil. As such, the correlator automatically aligns samples such that the time of travel from the center of field is the same for every antenna. Second, the processes that generate deep space radio signals are

independent at every frequency. Therefore, we should see no correlation between frequency channels. Thus we are left with $r_{i,j}[t,k]$, where i and j index antennas, t indexes time, and k indexes frequency.

While we could have our time samples at the same rate at this point as the sample rate at the antenna, it will typically be a much slower rate. Firstly, the channelization algorithm takes in a number of time samples to output a single set of channelized samples. The simplest algorithm (FFT) produces as many channels as it consumes time samples; better algorithms consume more time samples than they produce frequency samples. Secondly, to get better parameter estimation error, channelized samples are further time averaged.

2.2 Correlation

In the preceding section, we made the jump from voltages to correlations between voltages without explaining the mechanism by which we accomplished this transition. In general, we will do a sample correlation, wherein we multiply samples of the signals we want to correlate together at each sample step. Since we have N antennas, we get N^2 correlations at each time.

2.2.1 Parameter Estimation

In truth, we cannot actually measure the correlation between the two electric fields. However, we can estimate it. A good estimator produces an estimation of the correlation whose variance decreases as we add more samples to our estimation and whose expected value is the same as the true value.

If we were trying to measure the average electric field at a source, we could do so with an unbiased mean estimator. These are generally formulated as

$$\hat{E} = \frac{1}{N}\sum_{n=1}^{N} E_{measured}, \tag{2.5}$$

where $\hat{E}$ is our estimate of the electric field, N is the number of samples, and $E_{measured}$ is the electric field measured at our antenna. Unfortunately, because DC electric fields do not propagate, any non-zero mean we estimate will be from DC electric fields across our antenna.

Our other parameter of interest is the power (P), variance (σ^2), or standard deviation (σ) of the signal. An unbiased estimator for the variance is of the form

$$\hat{\sigma}^2 = \frac{1}{N+1} \sum_{n=1}^{N} E^2 - \hat{E}^2 \tag{2.6}$$

2.2.2 Beamforming

Before we continue, let us examine beamforming: a concept closely related to, but not necessarily foundational to our process. If a selection of antennas are linked, they can gain some of the benefits of a larger antennas. The system of linked antennas is often termed a phased array. By changing the weights assigned to the various antennas, we can change the shape of the array response. Conversely, by changing the relative phases assigned to each antenna output, we can change the direction to which the array is most sensitive.

In simple terms, the output of an array will be termed x_{out}, and will be defined as

$$/x_{out} = \sum w(n) * x_{in}(n) \tag{2.7}$$

where $x_{in}(n)$ is the voltage recorded on the n^{th} antenna and $w(n)$ is a weight (with both phase and magnitude) applied to the n^{th} antenna.

For a more comprehensive look at beamforming, we suggest [**?**].

2.3 Coherence

If our signals were perfectly narrow band (i.e. pure sinusoids), we would be able to know their relative phase and magnitude at all points by sampling them at a small number of points (or,

conversely, at a small number of times). However, since we are dealing with signals of finite bandwidth, we have more stringent requirements for determining how a signal is propagating across the array.

Let us take a single channel of bandwidth $2B$ and center frequency f_c. If we have two narrow band signals at $f_c + B$ and $f_c - B$, we cannot differentiate them within our channel. As we propagate across space, they transition from in-phase to out-of-phase and back. The spatial frequency associated with this transition is the same as the spatial frequency associated with $2B$.

Alternatively, if we have two signals with different propagation directions, we see a similar effect. The total spatial frequency is the vector sum of the individual spatial frequencies.

For our purposes, "coherence" shall be used to mean the space where varying signals have small impacts on how well we know the phase and amplitude relationships from a small set of samples.

2.3.1 Sky Distance vs Bandwidth

The aggregate effect of different propagation directions with varying frequencies in the channel causes interesting effects when combined with the time-alignment imposed by correlation. Small fields of view introduced by the individual responses of the antennas normally keep the effects of incoherence small. Interferometers work, to some extent, off the combination of the two previous effects. However, if a signal is powerful enough to enter the system through a side-lobe of the antenna pattern it causes all sorts of problems.

The end result is that an array experiences bands of coherence. The signals become incoherent in the direction of the vector difference between a vector pointing to the center of field and a vector pointing to the RFI. The rate of incoherence is also related to the length of the resulting vector.

To show this result, take two points on the celestial sphere. We will label one the signal of interest, and the other the RFI source. We will then draw a vector pointing at each from the array center, labeled S_{SOI} and S_{RFI} respectively.

2.4 Array Antenna Pattern and Interferometry

To use an array of antennas as a single antenna, the synthesized voltage can be computed by summing the antenna voltages with appropriate weights. If the antenna voltages are contained in a voltage vector $\mathbf{x}$, and the weights are contained in a weight vector $\mathbf{w}$, then the array output y can be computed by taking

$$y = \mathbf{x} \cdot \mathbf{w}. \tag{2.8}$$

The shape of the antenna pattern can be controlled by changing the magnitudes of the elements of $\mathbf{W}$, while the direction can be controlled by changing the phases of same. Figure 2.1 shows the change in array antenna patter than can be induced by fairly small changes in the weights given to the various antennas.

On the other hand, an interferometer attempts to use the difference in propagation distance to calculate characteristics that are indistinguishable to the individual instruments. To do so, it produces an interference pattern between the two signals. While an interferometer can be done in analog, radio waves can be sampled and reproduced, allowing digital interferometry. This works out to be a covariance between the antennas. Thus, instead of taking the inner product of the antenna voltages with a weight vector, we will instead take the outer product of the voltage vector with itself. Before we do so, we will need to align the voltage samples in time so that voltages from the appropriate times are correlated.

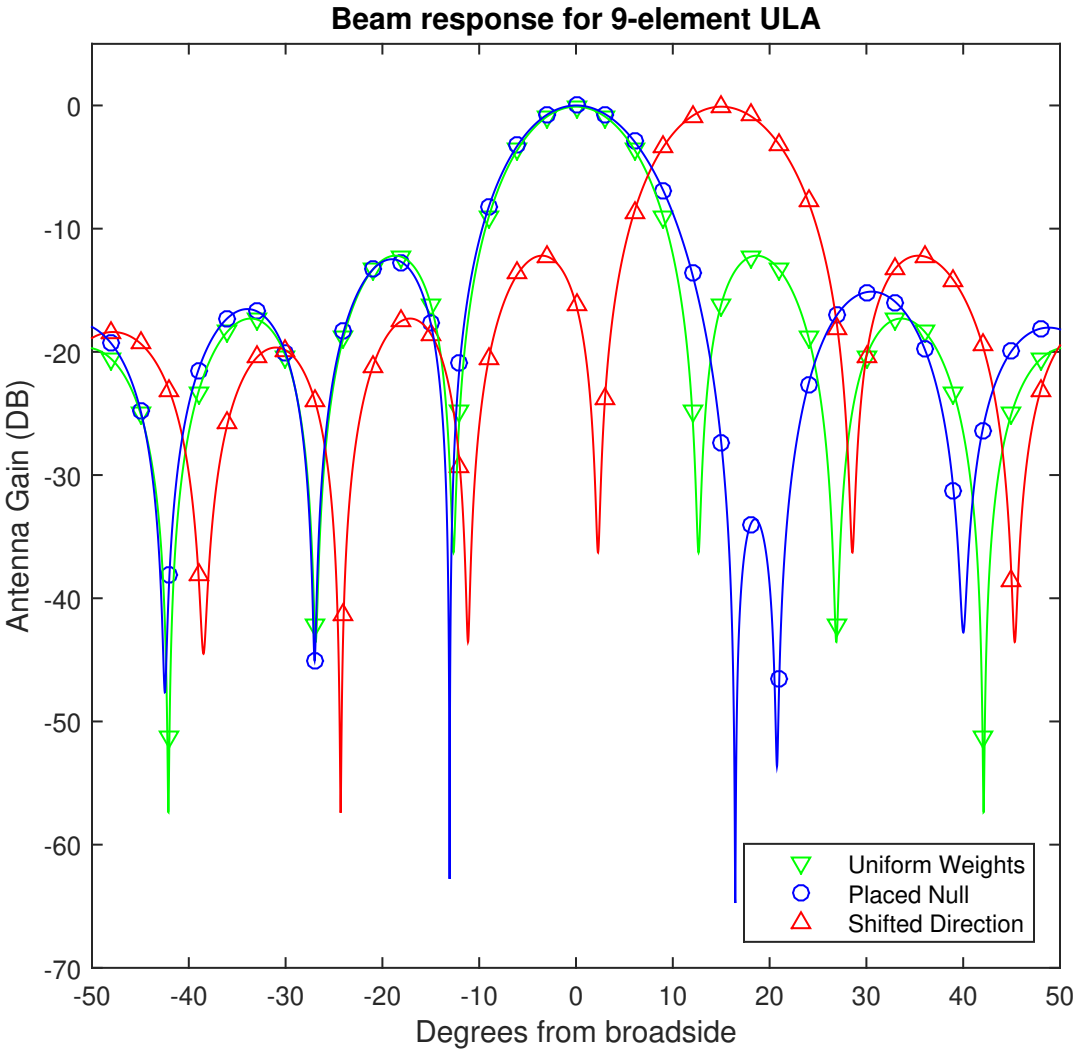

Figure 2.1: Three antenna patterns for a uniform linear array. The two triangle marked lines are different by a phase shift (and thus mostly direction). The circle marked line is placing a null.

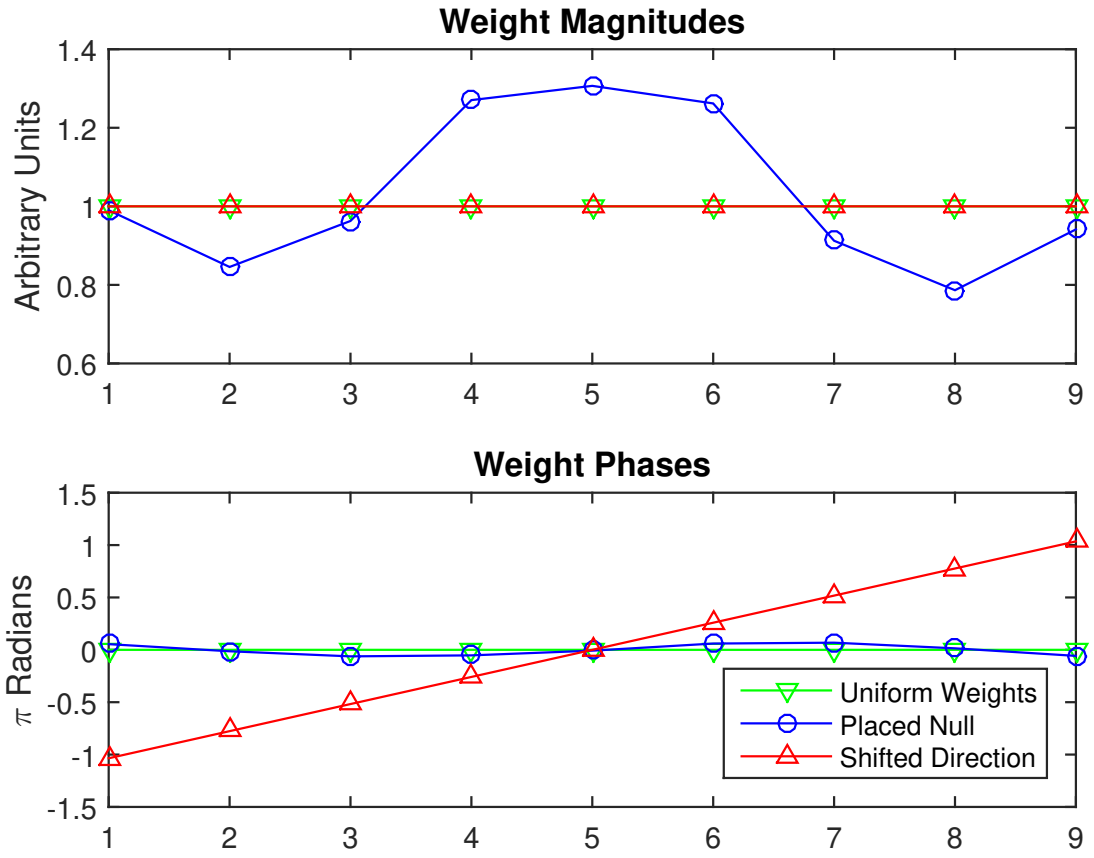

Figure 2.2: Magnitudes and phases for weights used to create antenna patterns shown in figure 2.1.

CHAPTER 3. VLA SIMULATION

As the VLA is a complicated and (as the name implies) very large system, it was useful for me to get a handle on how things worked before I attempted to control it during my student internship in Socorro during summer of 2017. Numerical simulations allowed access to data with strictly known properties, allowing the isolation of problems caused by the inherent properties of the instrument from problems caused by operator error.

First, we simulated VLA antennas as isotropic receivers and generated data as a string of random powers incident upon the array. Using this data, we calculated the approximate antenna pattern of the array and how the antenna pattern changed as a result of sub-space projection. The weakness of this first method was that it did not impose bulk time delays on the data, resulting in a somewhat more optimistic assessment of performance than was appropriate. Secondly, we generated data as the VLA would see it, including putting it in a VLA data format. This allowed us to put it through our RFI cancelling algorithm without changing the algorithm.

Finally, we consider the decorrelation that bulk time delays add to RFI. While showing that baselines are decorrelated is beyond the scope of this work (see Burnett et al. [**?**]), we will show simulations of how frequency channel processing bandwidth, sky separation, and RFI decorrelation interact.

3.1 Isotropic VLA

As described in section 2.4 when used as a single antenna, an array of identical antennas has a spatial pattern equal to the antenna pattern of each element multiplied by the array beamforming pattern of the whole array formed with isotropic antenna elements [**?**]. While the VLA is used as an interferometer rather than a classical beamforming antenna array, a similar principle applies. Thus, when we perform operations on the VLA, the effects of the element antenna patterns can be

ignored within the field of view established by the individual dis antennas' mainlobe beam. We assume all antennas are steered to point to the center of the desired field of view.

We developed a simulation in Matlab to apply cancellation on the VLA. We started with a simplified model for the antenna positions on the VLA, though the results continue to hold when actual VLA positions are used. Our simplifications consisted of making all arms the same length and aligning the axes such that the north arm points perfectly north. In actuality, the VLA north arm points slightly east of north and is slightly shorter than the other arms. Furthermore, we placed the VLA on a perfectly flat plain whereas the VLA has a vertical displacement in the tens of meters range. We placed our x-axis so it points east, our y-axis so it points north, and our z-axis pointing up from the center of the VLA.

With this model for the VLA geometry, we then modeled our signals. We started with the signal of interest (SOI). For our purposes, we modeled the SOI as a single point source at an azimuth and elevation of ϕ_{SOI} and θ_{SOI} respectively. We also modeled our RFI as a point source, but at a different azimuth and elevation of ϕ_{RFI} and θ_{RFI}. We also included a noise component. Although this noise represented a combination of sky and receiver noise, we modeled it as completely uncorrelated across the array, as if it were completely receiver noise. Positive azimuth was measured clockwise from north. Positive elevation was measured up from the horizon.

Then steering vectors were generated which pointed at the SOI and the RFI. A steering vector is a vector of the phase response of the array to a signal from a certain direction. For a plane wave incident on the array we can define its propagation with

$$\begin{aligned} \mathbf{k} &= k\hat{k}, \\ k &= \frac{f}{c}, \\ \hat{k} &= [\cos\theta\sin\phi, \cos\theta\cos\phi, \sin\phi], \end{aligned} \tag{3.1}$$

where $\mathbf{k}$ is the wave vector, f is the frequency of the incident plane wave, θ and ϕ are the elevation and azimuth and c is the speed of light. A signal in direction $[\phi, \theta]$ has $\mathbf{k}$ pointing at $[\pi+\phi, -\theta]$.

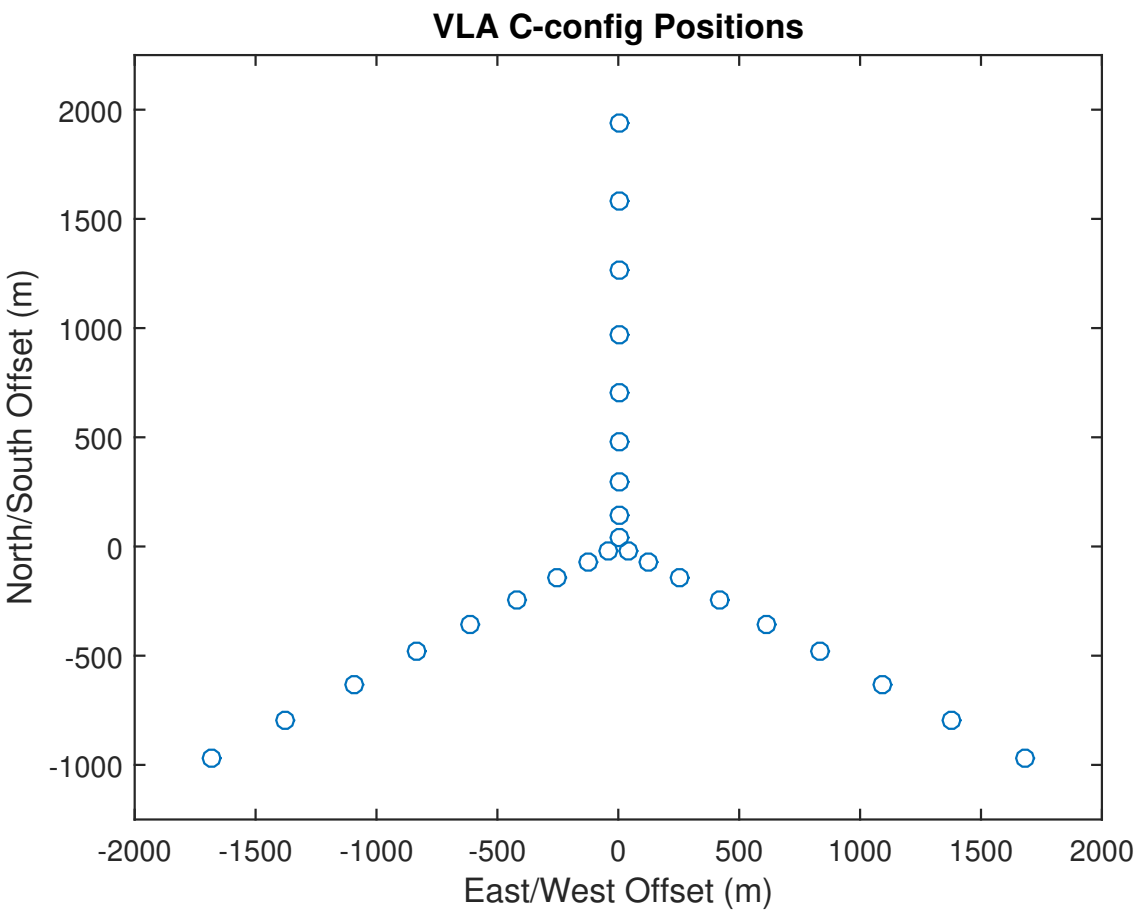

Figure 3.1: A simplified model of the VLA in C-config with the antenna locations marked with circles. All three arms are 10 km long. In reality, the north arm of the VLA is 10% shorter than the south-east and south-west arms.

For each antenna, we then found the phase response of the array by creating a vector $\mathbf{d}$ such that

$$\mathbf{d}_m = \mathbf{r_m} \cdot \mathbf{k} \tag{3.2}$$

with r_m being the vector pointing from a reference point to the location of the m^{th} antenna.

Having steering vectors pointing at the SOI and the RFI, we then generated a set of complex random numbers to represent the power incident on the array. We generated a vector $\mathbf{v}$ of circularly symmetric, unit-norm, complex random variables. Then we used the steering vectors to apply

structure to the string of random number as

$$\mathbf{V} = \mathbf{d}\mathbf{v}^T. \tag{3.3}$$

This produced an $M \times N$ matrix with the columns corresponding to samples and the rows corresponding to antennas. By producing a matrix for the RFI, the SOI, and the noise, we were able to construct a set of data for our system. We thus had

$$\begin{aligned} \mathbf{V}_{RFI} &= \mathbf{D}_{RFI}\mathbf{v}_{RFI}^T, \\ \mathbf{V}_{SOI} &= \mathbf{D}_{SOI}\mathbf{v}_{SOI}^T, \\ \mathbf{V}_{Array} &= \mathbf{V}_{RFI} + \mathbf{V}_{SOI} + \mathbf{V}_{noise}, \end{aligned} \tag{3.4}$$

Where $\mathbf{V}_{RFI}$ was the contribution to the array from the RFI, $\mathbf{V}_{SOI}$ was the contribution from the SOI, $\mathbf{V}_{noise}$ was an $M \times N$ matrix of random values for the noise, and $\mathbf{V}_{Array}$ was the total output of the array.

3.2 Robust Model

In order to assess the effectiveness of our algorithm, we first applied it to simulated data in the output format of the VLA. This data was created using a tool produced for analysis of RFI mitigation on the VLA.

The first pass over the image showed a marked decrease in the RFI's effects, as shown in figure 3.2. Further examination shows a decrease of 10% in the magnitude of the imaged source or an increase of 10% in the RMS background noise compared to an RFI free simulation.

The primary caveat about this data is that removing the self-power terms from the auto-correlation matrix hampers performance. When we attempted mitigation after deleting the diagonal terms in the covariance matrices, mitigation only worked halfway. While projecting out the eigen vectors associated with the largest eigen values removed much of RFI, the process stalled with significant

RFI still present. The RFI remained in the image until the final eigen vector was removed, at which point we were left with a featureless image of a blank sky: a particularly counter-productive result.

3.3 Unpacking Data

VLA Data is downloadable in a variety of formats. For our purposes, we used the measurement set (.ms) format. This format stores the data in a collection of tables. We used this format, rather than a more data compact format, because we we had tools available that could work with this data. Unfortunately, said tools were largely black boxes; our algorithms diverge from the standard usage such that we required substantial unpacking and reworking of the data before our algorithms could be applied.

Data in the .ms format is unpacked from the table with the `.getcol` command, which outputs the data in a 3 dimensional array. The 3 dimensions correspond to polarization, channel, and sample index. By using appropriate meta-data (also retrieved with the `.getcol` command) we are able to reshape the array into 5 dimensions with dimensions corresponding to first antenna, second antenna, polarization, channel, and time. Each slice with constant final 3 indices is thus a covariance matrix.

3.4 Analyzing Data

At this point, we have a few methods of analyzing the data. First, we could image the data. While this will give us some idea of how much RFI we are dealing with, it is qualitative and subjective. Alternatively, we could analyze the power in each channel compared to the average channel power; this would allow us to determine which channels were corrupted if we suspected that a small number of channels had been corrupted. Power in a channel is proportional to the trace of the covariance matrix. Finally, we can use a matrix decomposition to estimate how many independent sources are in our data.

3.4.1 Imaging

In an interferometric image, RFI will appear on the image as strong artifacts unrelated to the observed sky field of view, possibly masking desired signal. This masking ranges from incredibly subtle to blatantly obvious. In figure 3.2 we see the difference between a clean image and one with RFI. We can clearly see that, although some data is preserved, we lose a significant amount of data to masking by the RFI.

We can quantify how much an operation has changed the data by subtracting the images at a specific resolution and summing the squared differences. While this is straightforward, it does not handle changes in beam pattern caused by switching frequency channel. If our change in frequency is small, however, this is the best way to assess changes between images. We can both take a norm of the values before, after, and changed, and show the three images. Alternatively, we can compare to a ground truth image.

3.4.2 Power Levels

We can also analyze the power in any or all correlations. For instance, it is fairly easy to isolate channels with RFI by looking at the amplitude of their correlations: RFI infected channels tend to have amplitudes 10 - 1000 times as high as non-corrupted channels.

We can quantify change in this case by taking a frobenius norm of the correlation matrix. This runs into problems with changing frequency. Since a plane-wave of different frequency will propagate high and low amplitudes differently, different frequency channels will not maintain similar correlation matrix structure.

3.4.3 Matrix Decomposition

Another method of analyzing the data is to take the eigen-decomposition of the covariance matrices. Very large eigenvalues typically correspond to strong sources or RFI whereas low eigenvalues correspond to weaker, possibly diffuse, sources and noise.

One of our simulation tests was creating a data set with exactly two sources. We tested with two sources of equal strength and with two sources where one source was three times as strong as the other. In both cases, there were two non-noise eigen vectors. However, the image corresponding to the removal of a source were very different in the two cases.

In the first case, with two equal strength sources, the first eigen vector removed part of each source, leaving a bright ring of nearly uniform brightness behind.

In the second case, with two unequal sources, the first eigen vector removed almost all of the stronger source and only a little of the second source.

3.5 Cancellation

As explained in section **??**, when we have a covariance matrix R with unit-norm eigenvectors Λ_1 through Λ_M (where M is the dimensionality of the covariance matrix), we can partition the eigenvectors into those corresponding to the RFI and those corresponding to the stellar signals of interest. If we have p eigenvectors corresponding to RFI, we can form a $p \times M$ matrix as $U = [\Lambda_1, \Lambda_2, ..., \Lambda_p]$. Then we form a projection matrix by taking

$$P_{RFI} = U_{RFI} U_{RFI}^H. \tag{3.5}$$

This will project our covariance matrix in the direction of the RFI. Since we want to remove the RFI's impact on the data, we will form a projection into a perpendicular space by creating another projection matrix of

$$P_{\perp RFI} = I - P_{RFI}. \tag{3.6}$$

This allows us to achieve Q_{clean} by taking

$$R_{clean} = P_{\perp RFI} R P_{\perp RFI}^H. \tag{3.7}$$

3.6 Self Power Terms

When we move from isotropic antennas viewing an infinitely distant sky to actual antennas viewing objects in 3-dimensional space, we introduce significant complications. Some of these issues work in our favor, while some work against us. A highly directional antenna reduces the ability of RFI to enter the signal path of the interferometer. On the other hand, while antenna locations can be treated as perfectly interchangeable under an infinitely distant sky, when viewing objects in actual space they become less so.

Take, for instance, a VLA placed on a perfect spherical Earth. Set the center of the VLA at θ_{VLA} west longitude, ϕ_{VLA} north longitude, and an altitude of e_{VLA} meters above sea level. At the center of the VLA, these values are approximately, $\theta_{VLA} = 107.6\,\text{deg}$, $e_{VLA} = 34.08\,\text{deg}$, and $e_{VLA} = 2124$ meters.

With the position of the VLA thus defined in a spherical coordinate system, we can translate to a rectilinear system by assigning an x-axis through the point where the equator and prime meridian meet (i.e. 0 degrees latitude and longitude), a y-axis through 90 degrees east latitude 0 degrees longitude, and a z-axis through the north pole. We can find the

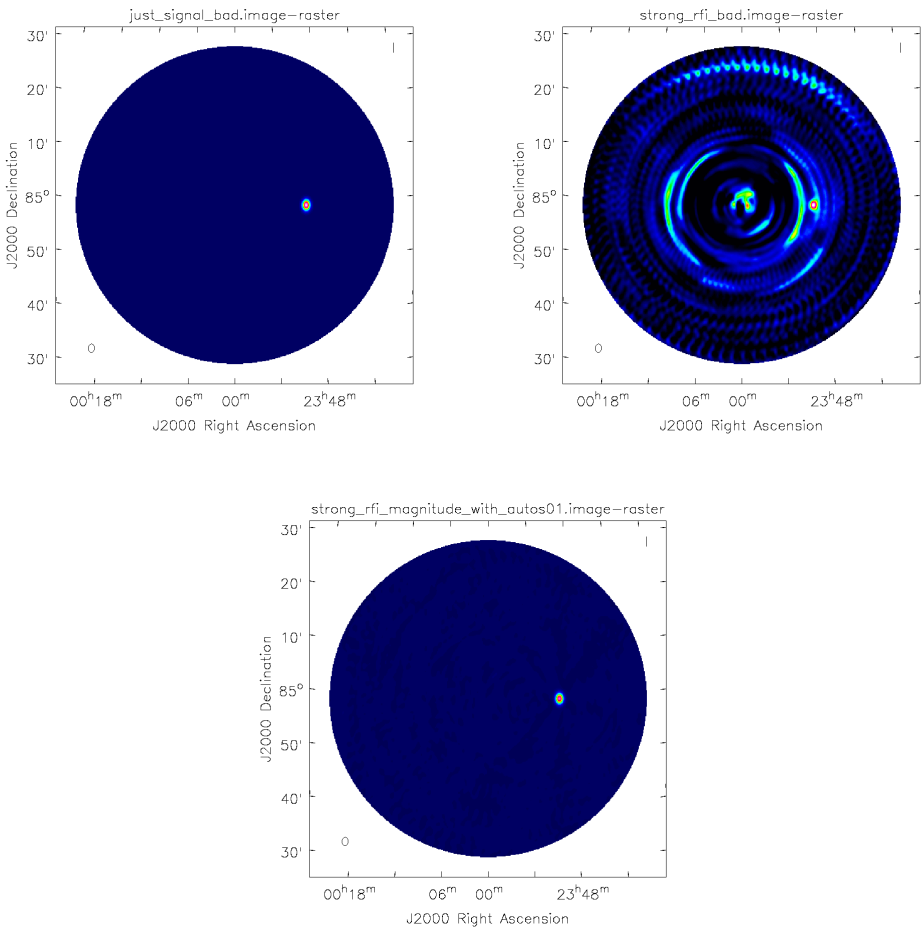

Figure 3.2: Images of a simulated stellar source without RFI (top left), with RFI added (top right), and with RFI added and then removed by subspace projection (bottom). The final case does not quite return the image to the pristine case, but nevertheless is much improved.

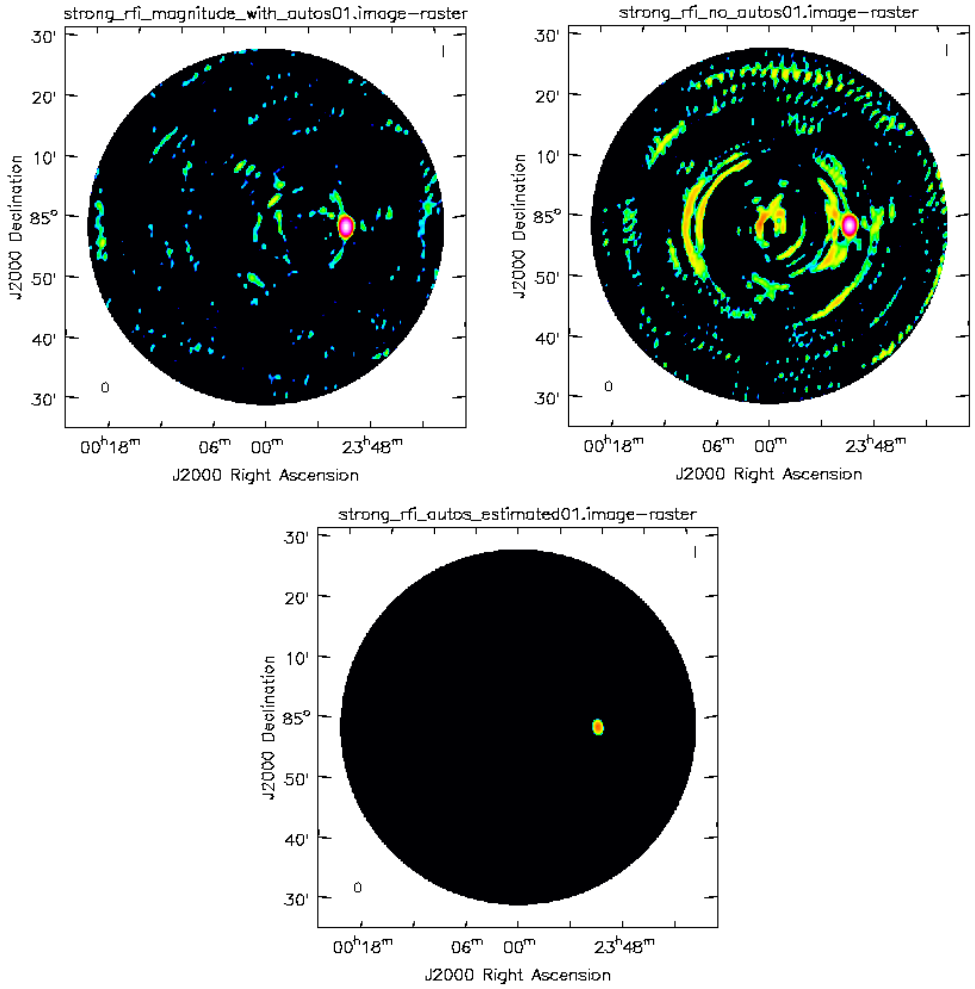

Figure 3.3: Images showing the difference between a subspace projection where the diagonal terms are known (top left), where they have simply been removed (top right), and where they have been deleted and only guessed at (bottom). The peaks of the first 2 are about 0.9 tall, while the third is only 0.04 tall. The low end of the range has been compressed to better show the difference between the outcomes.

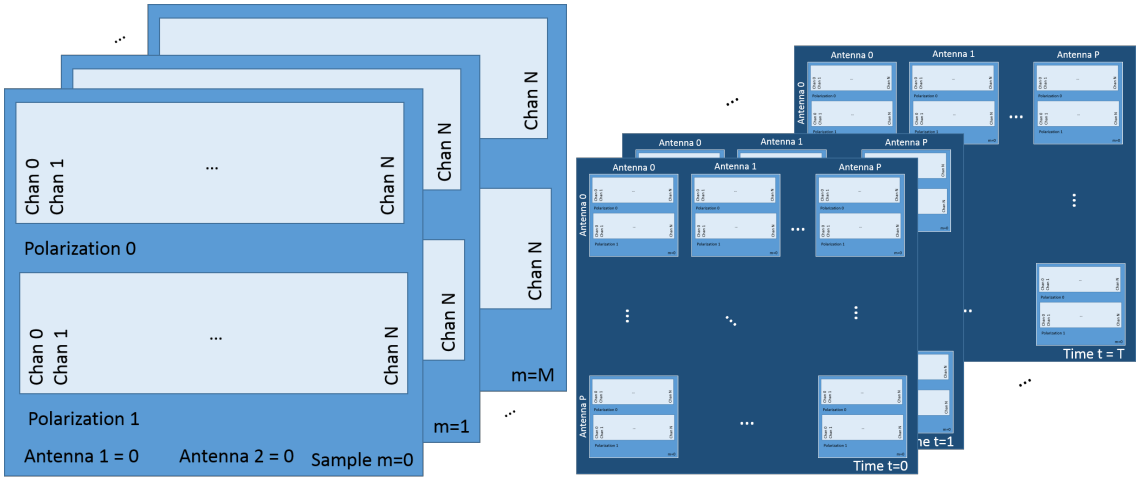

Figure 3.4: Cartoons showing the dimensionality of the arrays of data in the .ms format (left), and as covariance matrices (right). In the .ms format the data is stored by samples which each have an attendant first and second antenna. We have reshaped the array such that the samples are grouped by time into covariance matrices.

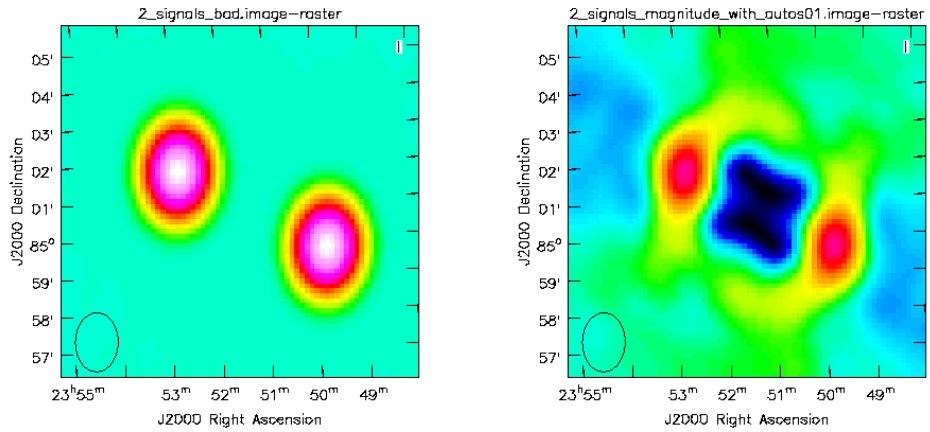

Figure 3.5: Two images of two unequal strength sources. Left: both sources with no mitigation. Right: image after removal of largest eigen vector.

Figure 3.6: Three images of two equal strength sources. Top left: both sources with no mitigation. Top right: image after removal of largest eigen vector. Bottom: image of only largest eigen vector.

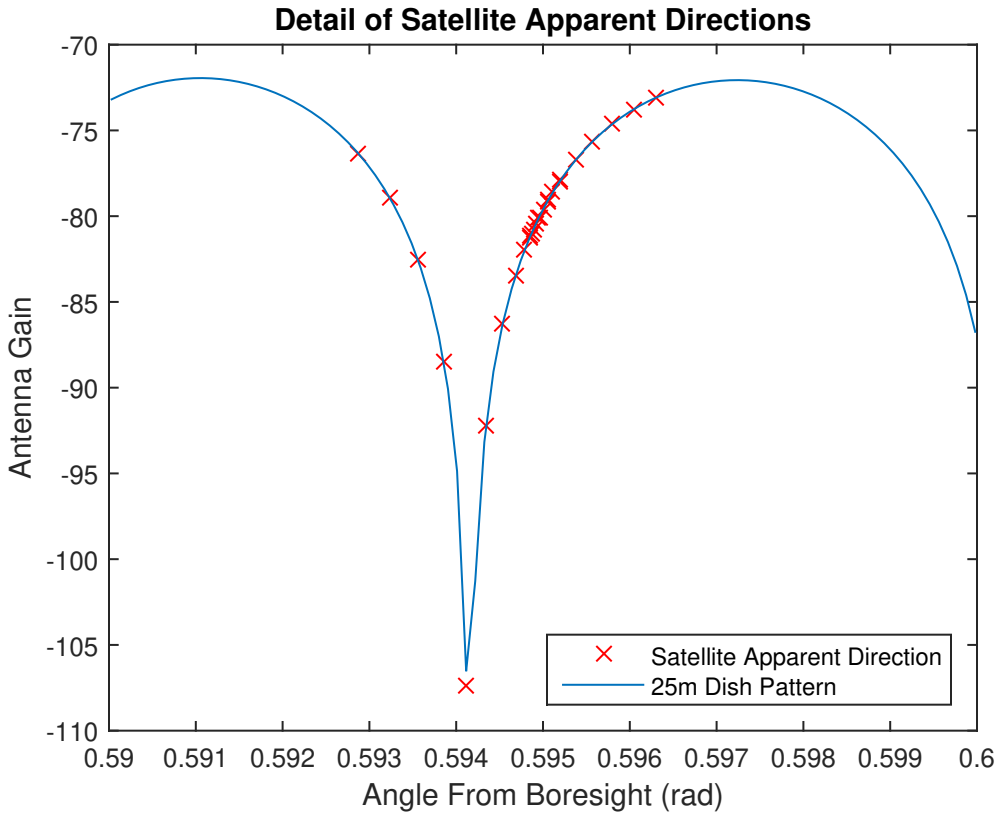

Figure 3.7: Detail of the antenna pattern of an ideal 25 meter aperture antenna. The direction of the satellite from the antennas of the VLA are marked with red 'x's. The VLA antennas are not idea 25 meter apertures, further complicating antenna patterns. Directivity varies between -73.06 dB at the far right point to -107.4 dB in the depth of the trough.

CHAPTER 4. IMPLEMENTATION OF SUBSPACE PROJECTION ON VLA DATA

After getting a handle on the mechanics of the RFI cancelling system using simulations, the time came to turn our attention to real data. Real data has the benefit of being very ”hard” on simplifying assumptions. On the other hand, real data collected on an instrument like the VLA has unknown levels of signal, RFI, and noise. We tried to collect frequency channelized covariance data located near an edge of the expected RFI spectral bandwidth so there would be both clean and corrupted channels for comparison. Although the conditions we operated under were artificially constrained, we hope that a working proof of concept on the hardware will encourage further exploration.

4.1 Data Description

A test data set was collected with the VLA in ”C” configuration. The VLA is composed of 27 antennas that are variously placed on 72 positions. There are four standard arrangements, labeled with letters from ”A” through ”D”. ”A” configuration has the longest baselines and the higher letters have successively shorter baselines. The VLA had to be put in a special mode to collect full autocorrelation data sets for each antenna. In the ordinary course of things, the VLA correlator only produces autocorrelations for odd numbered antennas.

The data set had 20ms averaging times and 256 62.5 kHz-wide channels for a total bandwidth of 16 MHz. The center of the spectral window was at 2.34 GHz. The channel bandwidth was chosen such that there was minimal decorrelation across the array at the sky separation we were planning to scan at. The center frequency was chosen to contain the edge of the band used by the Sirius XM satellite radio satellites based on data

The averaging time was chosen so that many samples could be taken before the RFI and celestial sources moved enough relative to eachother to cause smearing. We chose 20ms because the VLA

has a synthesized beam of about 7 arc seconds in S-band and with the VLA in C-configuration [8]. Our imaging will use a 2.0 arc second pixel. With the approximate sidereal rotation rate of

$$\frac{360^\circ}{\text{day}} = \frac{1^\circ}{4\text{minutes}} = \frac{15''}{\text{s}} = \frac{1''}{66.7\text{ms}} \tag{4.1}$$

we have about 6 integrations before motion can even theoretically cause smearing.

Of particular interest is where we pointed the VLA. We desired an RFI source operating on a known channel and in a known location. For known location, ground stations and geo-stationary satelites were optimal. We decided to focus on targets in geo-stationary orbit because data on band-usage by ground beacons was inconclusive. We also erroneously believed that shorter wavelength signals experienced more decorrelation across a constant spacing array, which caused us to select the signal with the lowest frequency that we could reliably locate. Ultimately, we decided that a Sirius XM satellite radio satellite in geostationary orbit was a good choice. From the VLA site, two are located within 40 degrees of due south.

From there we were able to select a cosmic continuum (boradband) source such that when we phased up on the source, the RFI would be within the decorrelation limit (see section 2.3.1) during the times we expected to have access to the VLA. We scanned a source called ”3C138” for approximately 10 minutes. Detailed information about data sets is tabulated in table 4.1. During our observation, the satellite was separated from the source by an average of 54.24° azimuth and 5.53° elevation, for a total average sky separation of 32.9°. Images of the source are in figures 4.1 and 4.2. One of the images is a synthesized from covariance data from our data set. The other is a product of the VLA Low-frequency Sky Survey (VLSS), showing a $1^\circ \times 1^\circ$ region of sky centered on 3C138. [9] As both images show, our source of interest is a bright point in a dark field, which makes it a good calibration source, but makes our analysis somewhat boring.

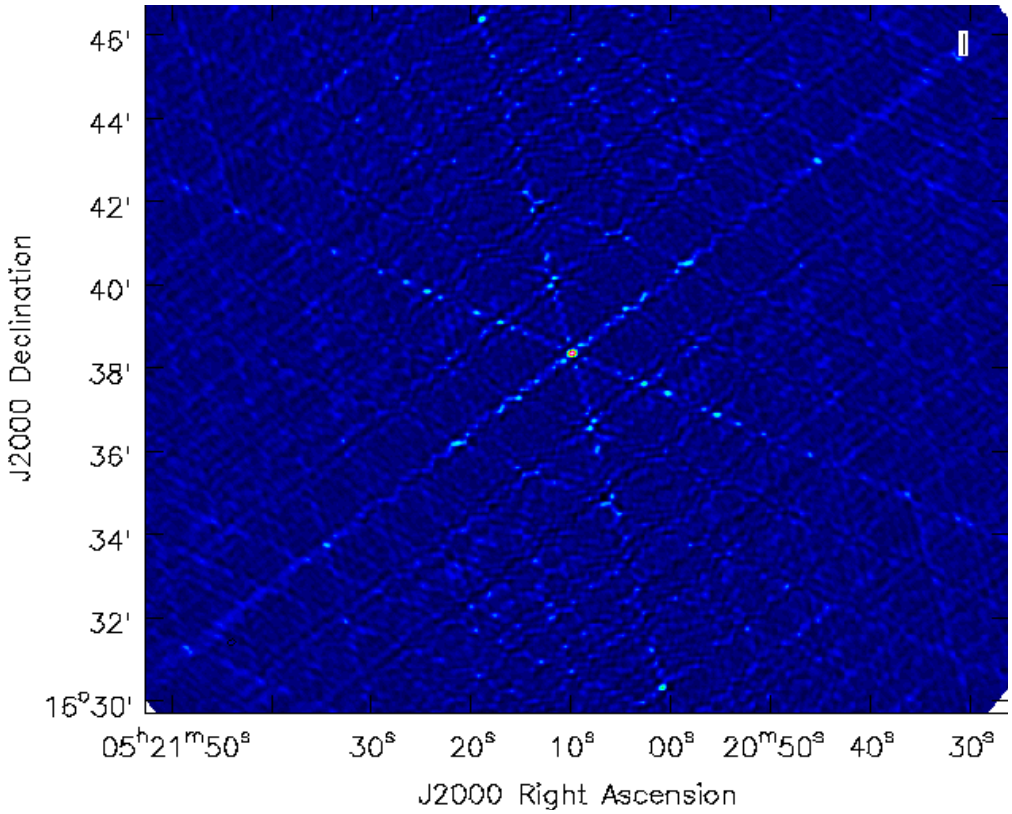

Figure 4.1: Image of 3C138 created from clean channels in the dataset. The source is the bright spot in the center. The six lines emanating from the center are imaging artifacts caused by the ”dirty beam pattern” of the VLA. [1]

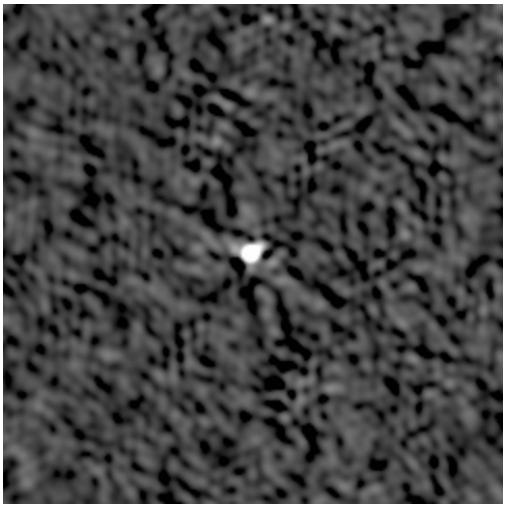

Figure 4.2: Image of 3C138 pulled from VLA Low-Frequency Sky Survey (VLSS) archive showing half a degree to either side of our source of interest.

Name	$TRFI0003_sb34043916$
Time	$2017-08-03-17:53:08.8$
Source	$J0318+1628(i.e.3C138)$
Integration Time	$20ms$
Time Steps	29872

Table 4.1: 3C138 Dataset Particulars

4.2 Data Processing

After unpacking the data, we were able to partition it by either channels or time segments. After a cursory examination of the power levels in each channel, we were able to divide the channels into 3 types: highly corrupted channels, moderately corrupted channels, and RFI free channels. As can be seen in figure 4.3, channels from about 30 to 85 are highly corrupted, channels below 20 and from 86 to 105 are moderately corrupted, and channels above 106 are not corrupted. The frequency bins were 62.5 kHz wide. The lowest channel was centered at 2.34 GHz and the highest channel was centered at 2.356 GHz for a total of 16 MHz.

For our purposes we will create a ground truth image using an average of channels outside the influence of the RFI. Specifically, we formed an image with the casa "tclean" command. The exact nature of the "tclean" command is beyond the scope of this work, though the straightforward explanation is that it applies an inverse 2-dimensional FFT to data that has been averaged in visibility space. Then it iteratively finds peaks and subtracts out the dirty beam pattern scaled to the height of the respective peak. The dirty beam pattern is replaced by a more concentrated shape. Among its outputs are an image and a model, both shown in figure 4.4. For this, we used channels 150 to 200. We used the model here created as the benchmark against which all errors will be calculated.

We used 4 methods of analyzing the channels. First, we look at the eigen values of the channels. We collected every eigen value across both time and channel, but as plotting a surface for all 25 eigen values is both too much information to be understood and not enough information for complete analysis, we will instead examine channel cuts in time and channel averages.

The second analysis method was to take what we will call "total normalized error". To find total normalized error, we took an image-under-test, scaled its central pixel to the same height as the

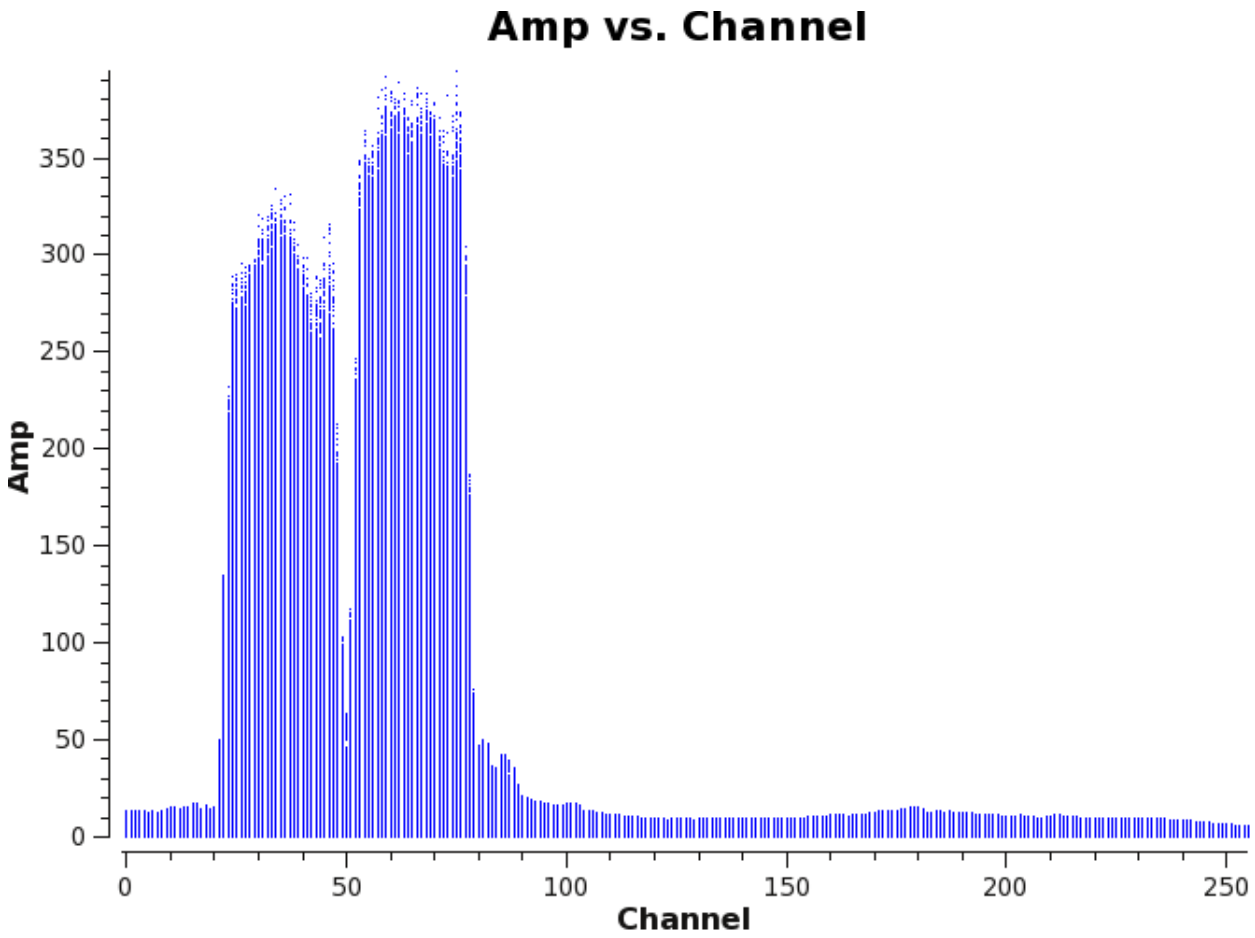

Figure 4.3: The power seen by every baseline and every time step over a single minute plotted by channel as a scatter plot. The range from 15 to 105 are corrupted by RFI.

ground truth model, and took the L_2 norm of the difference between the image-under-test and the ground truth model. This has the benefit of being constant for constant signal-to-noise ratio.

Our third analysis method is to examine the height of the central pixel in the image, which statistic we will call ”central peak height”. While this is a poor method for analyzing the quality of RFI supression, it can be used to draw conclusions about other results. For example, if we have rising total normalized error coupled with declining central peak height the simple conclusion is that we are projecting out the central peak faster than we are projecting out any remaining RFI or background noise.

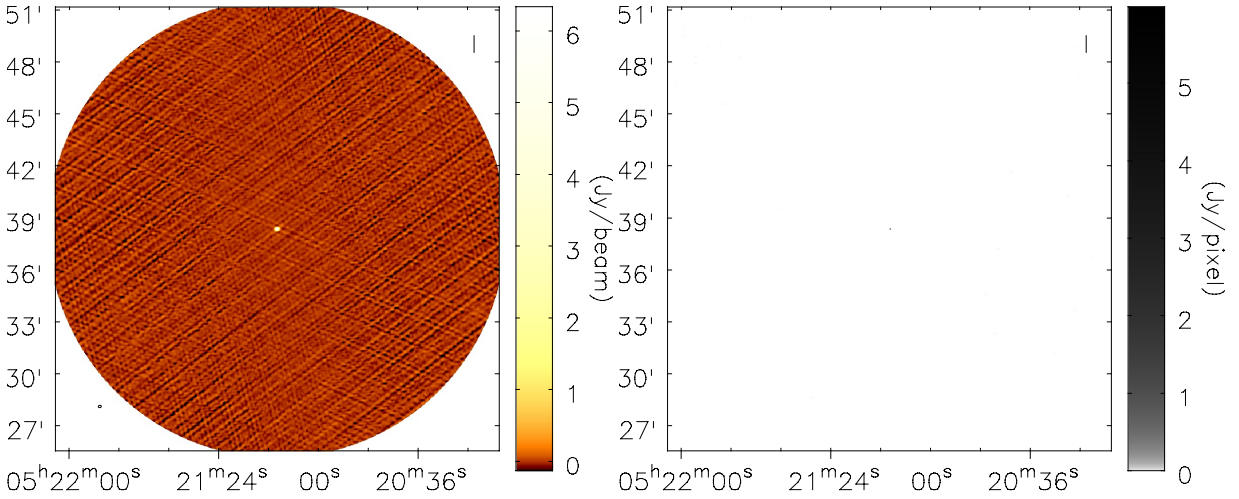

Figure 4.4: ”Cleaned” images formed from RFI free channels from 150 to 200. Left is the image formed by replacing the dirty beam pattern with a more Gaussian shape. Right is the model image consisting of a single non-zero pixel in the center.

Our fourth and final analysis technique is to look at the images we have formed. By looking closely at the images, we can make judgements about how much RFI we have removed and determine whether the removed eigenvectors have detrimentally impacted the astronomical data. All of our images are of a simple point source in a blank sky, so there is not very much astronomical data to judge that on, but if we had a diffuse source we could much more easily examine its shape in the image as a first-order check on its integrity.

We applied subspace projection as described in section **??** with one important caveat: instead of ordering the eigenvectors by eigen value, we instead ordered the eigenvectors by magnitude of eigen value. Why we did so and what effect the negative eigenvectors had will be explained below in section 4.5.

4.3 Data Analysis

We have done an in depth review of channel 30 quantifying the RFI that infects it and how it responds to mitigation. We have also done a cursory analysis of every channel from 0 to 149. First we examined a single channel, then we have done an overview of all the channels.

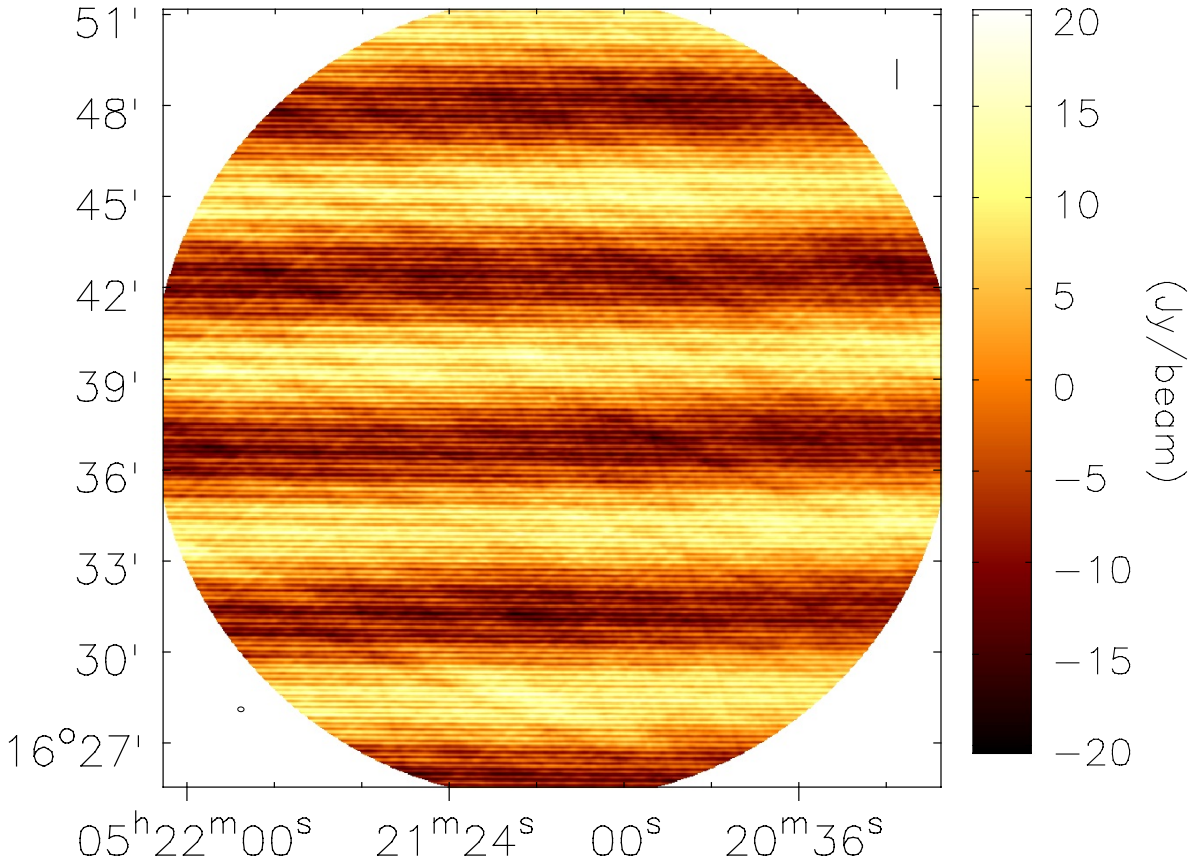

Figure 4.5: Image created with data from channel 30 with no mitigation. 3C138 is completely hidden by the RFI artifacts. There is a blip near the center of the image but its center does not align with 3C138. Horizontal axis is in right ascension, vertial axis is in declination.

4.3.1 Single Channel Case Study

We chose channel 30 for our case study because, as shown in figure 4.5, the image is dominated by the RFI. The RFI artifacts are approximately sinusoidal in declination with an amplitude of about 20 janskies per beam. 3C138 (the source of interest) is hidden in the side of the central artifact. This makes it a perfect candidate for RFI mitigation as all improvements are dramatic. A less corrupted channel, while equally valid mathematically, does not improve as much under the effects of mitigation. A completely non-corrupted channel gets progressively worse under mitigation.

As shown in figure 4.6, the minimum error in the corrupted channel is about 4 times the error in a cleaned, multi-channel image and is about 50% higher than the error in an unmitigated non-

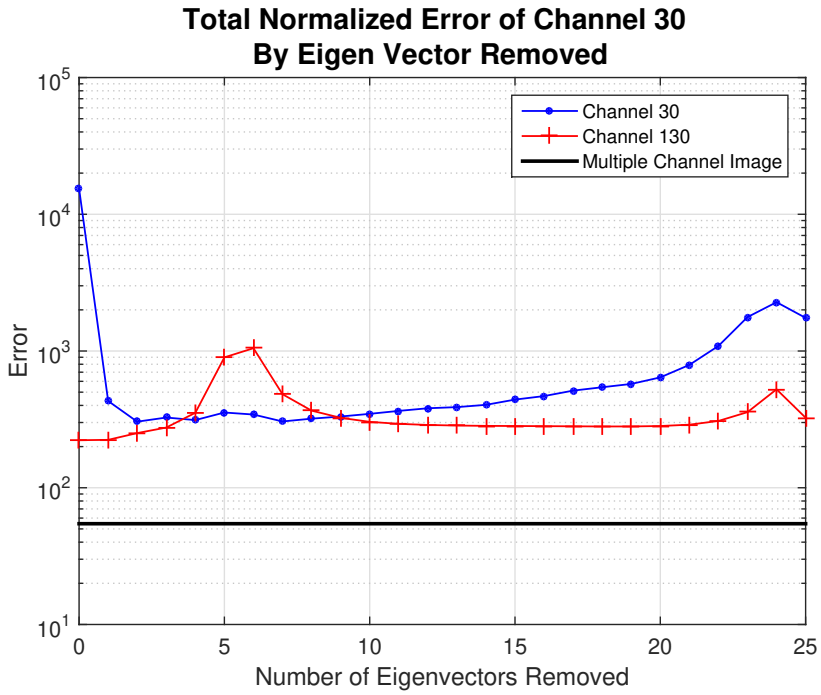

Figure 4.6: Total normalized error in an image of channel 30 by eigenvectors removed. Lines representing the total normalized error in channels 130 as a function of eigenvectors removed and the total normalized error in the ground truth image included for reference purposes. The minimum error in channel 30 happens at 2 eigenvectors removed, although 4 and 7 are also local minima.

corrupted channel. Channel 30 had a minimum error of 304 Jy/beam compared to its unmitigated total normalized error of 1539 Jy/beam. Channel 130 had an unmitigated total normalized error of 223 Jy/beam. The cleaned, multi-channel image had a total normalized error of 55 Jy/beam.

When looking at the total normalized error as eigenvectors are removed, we can see a rapid decrease in the error, followed by a fairly constant error, followed by a slowly rising error. At 2 eigenvectors removed, the error experiences its absolute minimum, though it has local minima at 4 and 7 eigenvectors removed. Error rises after 7 eigen vectors removed because the RFI has been removed and the process starts removing the source of interest.

If we compare to a non-corrupted channel, we see the error rise quickly as we project out the source of interest. However, starting with 7 eigenvectors removed the total normalized error falls again.

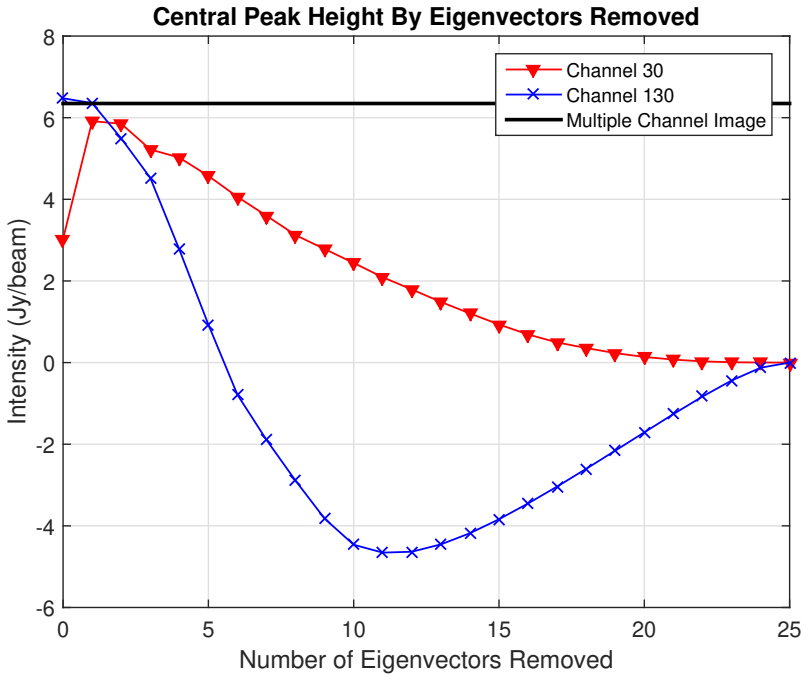

Figure 4.7: Plot of central peak height by eigenvectors projected out for channels 30 and 130. Channel 30 sees its peak height reduced in a fairly exponential way. Channel 130 falls off much faster and sees its central peak go negative after 5 eigenvectors removed.

Some of the questions about the non-corrupted channel are put in to perspective when we examine the central peak heights of channel 30 and 130 under varying level of mitigation, as plotted in figure 4.7. We see that the corrupted channel quickly sheds the RFI artifacts that were overwhelming its central peak. The central peak height slowly falls from 4 eigenvectors removed to 25 eigenvectors removed. Perhaps the greatest problem is that when the RFI is most removed, the central peak has lost about one third of its height.

On the other hand, the non-corrupted channel has its central peak quickly wiped out by mitigation. After 5 eigenvectors removed, however, its central peak becomes negative. This is a clearly non-physical result, but no explanation is readily apparent.

Turning our attention to the process of mitigation, we examined the eigenspectra of channels 30 and 130 across our sampled time. As shown in figure 4.8 the eigenspectra of the channels vary over time. Particularly, the largest eigenvalue has a fairly large range (going from 2×10^5 to 8×10^5

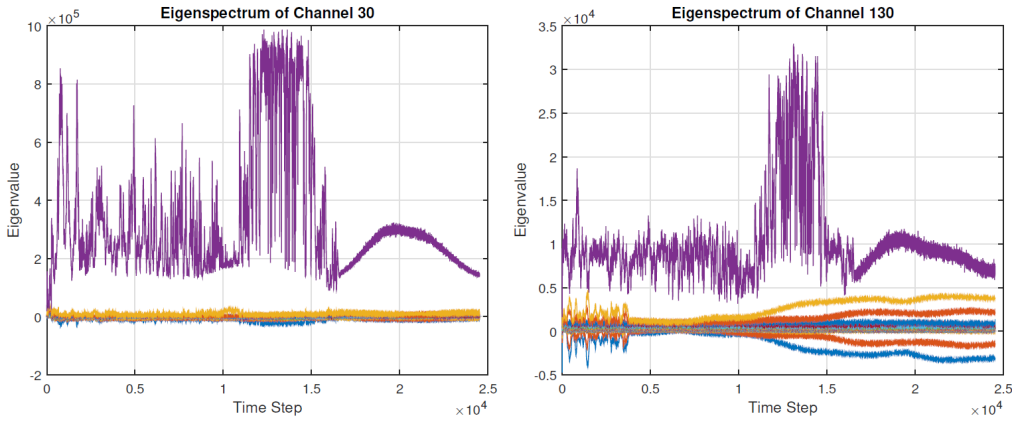

Figure 4.8: Eigenspectra of channels 30 (left) and 130 (right). While they have similar structure in the largest eigenvalue, channel 30 is has maximum eigenvalues about 20 times the maximum eigen value of channel 130.

in the corrupted channel) and has a high variance, but the other eigenvalues vary more slowly. On the other hand, as shown in table 4.2, there are two negative eigenvectors in channel 130 and fourteen negative eigenvalues in channel 30. The visibility matrix should be positive definite, and thus should not have negative eigenvalues. Further consideration will be given in section 4.5.

The ultimate fruit of our work is extracting astronomical data out of corrupted channels. As can be seen in figure 4.9 we can turn an image completely washed out by RFI in to an image that looks like a clean channel.

Channel 30		-8319	-4134	-2172	-1227	-961	-804	-719	-573
		-364	-322	-214	-160	-130	-99	0	40
	141	207	315	460	945	4723	8103	14628	143834
Channel 130		-3114	-1461	36	186	226	263	296	310
		334	343	360	393	406	430	442	474
	504	550	588	681	734	858	2095	3769	7049

Table 4.2: Eigenvalues of the final time step for channels 30 and 130 in ascending order. Channel 30 has 14 negative eigenvalues to channel 130's 2. The maximum eigenvalue is also about 20 times greater in channel 30 than in channel 130.

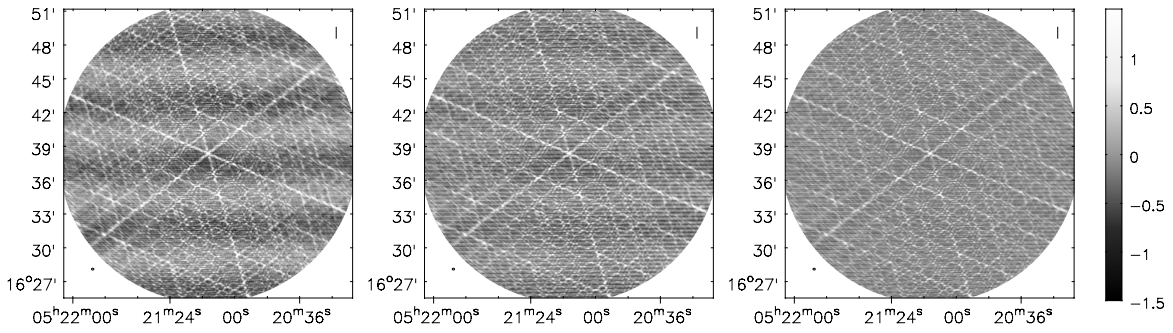

Figure 4.9: Images of channel 30. From left to right, images have 2, 4 and 7 eigenvectors removed. Horizontal axis is in right ascension, vertial axis is in declination. These three images correspond to the 3 lowest errors. Although the dirty beam pattern is clearly visible with 2 and 4 eigenvectors removed, there are still artifacts from RFI in the image. It is only with 7 eigenvectors removed that the artifacts disappear below the noise floor.

4.3.2 Multiple Channel Overview

While a thorough examination of every channel would doubtless be fruitful, there is too much data for comprehension and understanding. Instead, we reduced the data to make it more accessible. Good metrics were crucial for this purpose. Total normalized error, as shown in figure 4.6 and figure 4.9, tracks well with image quality. As such we began by examining total normalized error by channel.

The most obvious feature of the error by channel plot is that the corrupted channels are spread over several orders of magnitude. From channel 0 through channel 105 the total normalized error is raised above the almost constant error from channel 105 to 150.

In the non-corrupted channels removing eigenvectors increases error. As shown in figure 4.11, the first eigenvector has little effect but further eienvectors drive the error significantly higher.

On the other hand, removing eigenvectors from the corrupted channels creates a marked improvement. In order to show how much of an improvement, we have only shown the lower end of the unmitigated case in figure 4.12. The first eigenvector removed brings the images largely into line with the expected composition, further eigenvectors removed only makes incremental improvements.

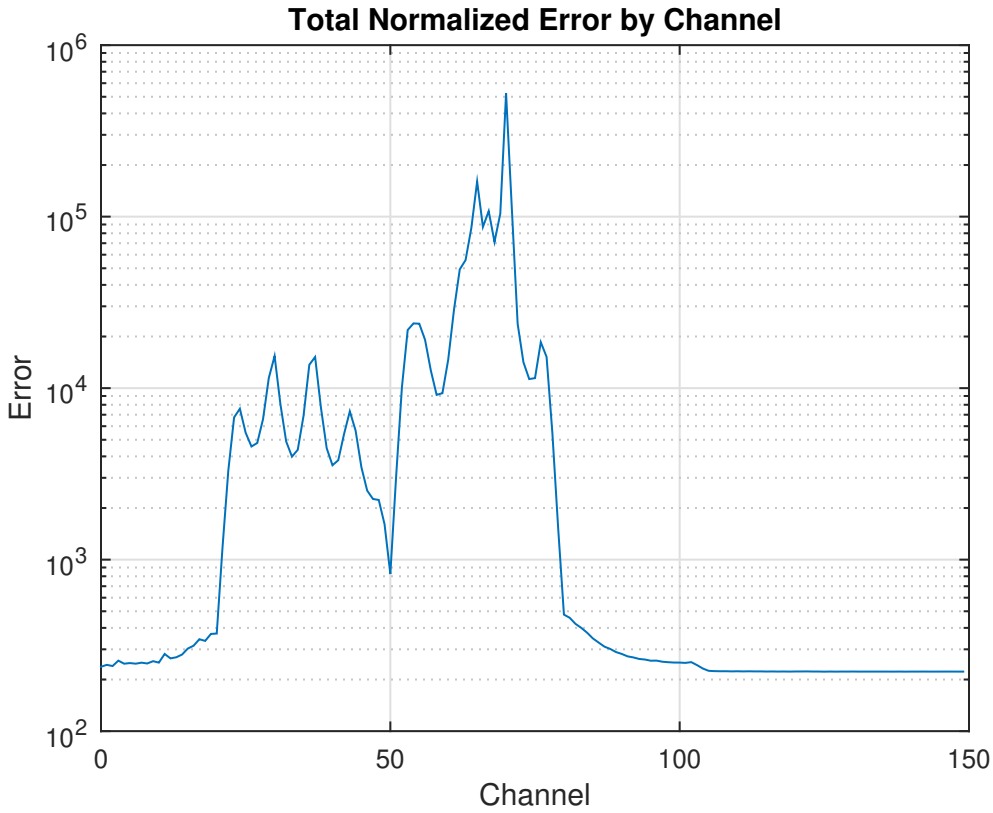

Figure 4.10: Total error of channels 0 to 149 with no mitigation applied. We can see the two highly corrupted bands between channel 20 and channel 80

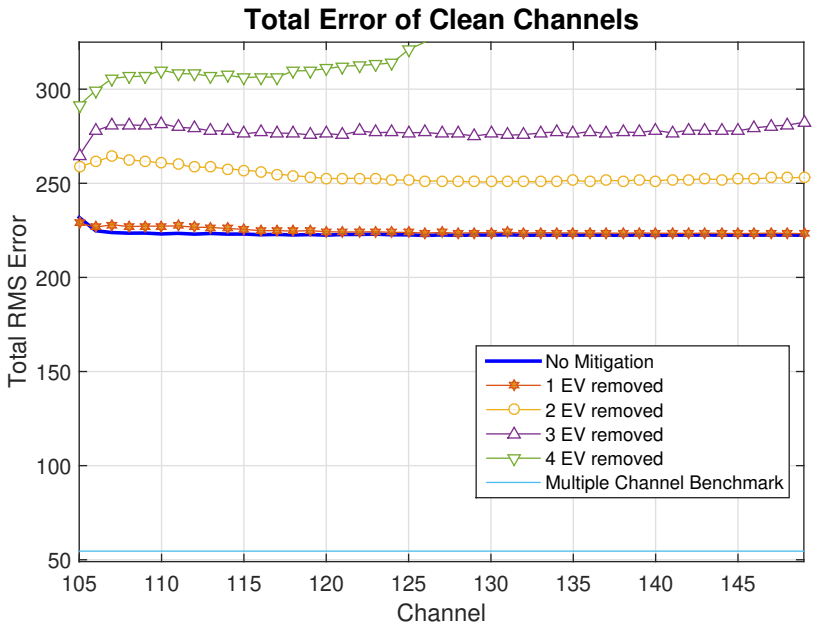

Figure 4.11: Plot of the total normalized error in non-corrupted channels. Removing 1 eigenvector has little effect on the total error in these channels, but further eigenvectors increase error.

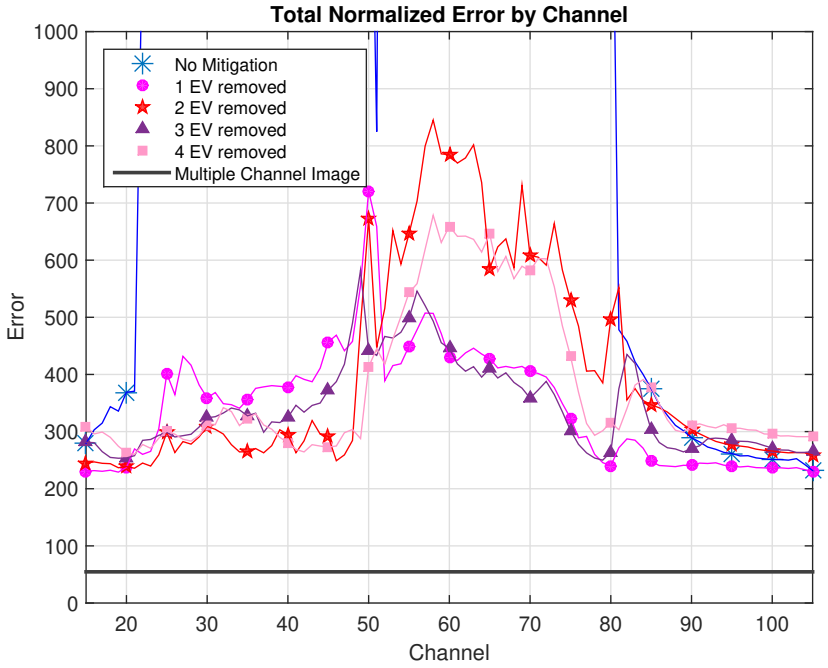

Figure 4.12: Plot of error in corrupted channels. Although single eigenvector removes the majority of the RFI, removing further eigenvectors eventually yields improvements.

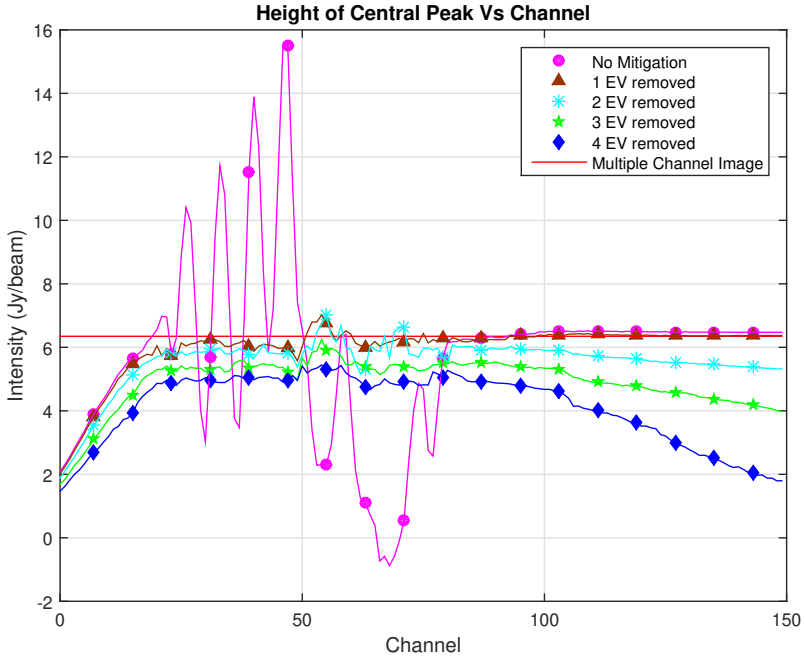

Figure 4.13: Plot of central pixel height be channel. The oscillations in peak height correspond to variations in error across the RFI band.

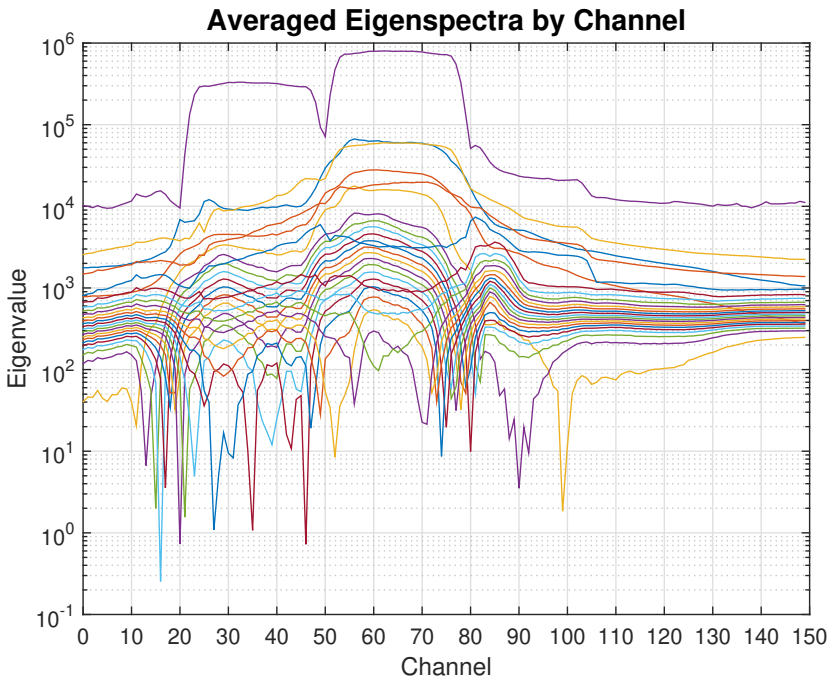

Figure 4.14: Plot of averaged eigenspectra by channel. The two RFI bands are readily evident in the eigenspectra.

Turning our attention to the height of the central peak, as shown in figure 4.13, some of the variability in error is explained by examining the central heights of the channels. Since total normalized error is calculated after normalizing the central pixel, extreme variations in central pixel height result in variations in total normalized error. The oscillations in central peak height from channel channel 25 to channel 50 correspond to variations in the total normalized error or the same range. The extreme deviation in channel 70's total normalized error is explained by channel 70 having an unmitigated peak height very close to zero.

When we examine the mitigated case, we see that corrupted channels have a peak height very close to the multiple channel image ground truth after one eigenvector has been removed, with central peak height dropping off as further eigenvectors are removed.

Examining the spectrum of eigenvalues (or eigenspectra) of all channels together reveals that the largest eigenvalues (shown in figure 4.14) nicely follow the maximum amplitudes of the channels (plotted figure 4.3). Furthermore, lower level eigenvalues rise to the level previously occupied by

the highest level eigenvalue. This implies that the highest eigenvalue corresponds to the RFI in corrupted channels and to either noise or the source of interst in RFI free channels.

4.4 Problems Introduced by Removing Autocorrelations

It is worthwhile to examine the behavior when we zero out the diagonal terms in our correlation matrices. This approximates the behavior of the VLA in its normal operating mode. Autocorrelations are not used in synthesis imaging. Thus, when we began operating on normally acquired data we were surprised by the apparently high floor that the algorithm experienced. Figure 4.15 shows the marked difference between data that has the autocorrelations and data without them.

We saw worse behavior with respect to both the signal of interest and the RFI. While the RFI is reduced in the autocorrelations removed case, it still dominates the image with 2 eigenvectors removed. On the other hand, the RFI is still noticeable in the autocorrelations retained case, but its impact is much more muted. The signal of interest, on the other hand, is reduced from an expected 6.35 Jy/beam in the displayed in the clean channels to 5.54 Jy/beam after 2 eigenvectors have been projected out and to 4.26 Jy/beam after 4 eigenvectors have been projected out.

To analyze this, we used two methods of image space error analysis and a visual, qualitative inspection of the image to see how much the artifacts introduced by the RFI were reduced.

The first image space error analysis method we used was to calculate the total error in the image against an idealized ”ground truth” image. We used the same ”ground truth” image for this analysis as in section 4.2. The second technique was to leverage knowledge of source locations in the image and compare source intensity to the ”ground truth” source intensity.

Error was calculated by taking the image values at each point in an image, normalizing the central peak to value of the central peak in the ”ground truth” image, and then taking the root-mean-square error between the ”ground truth” image and the channel image. Source intensity was

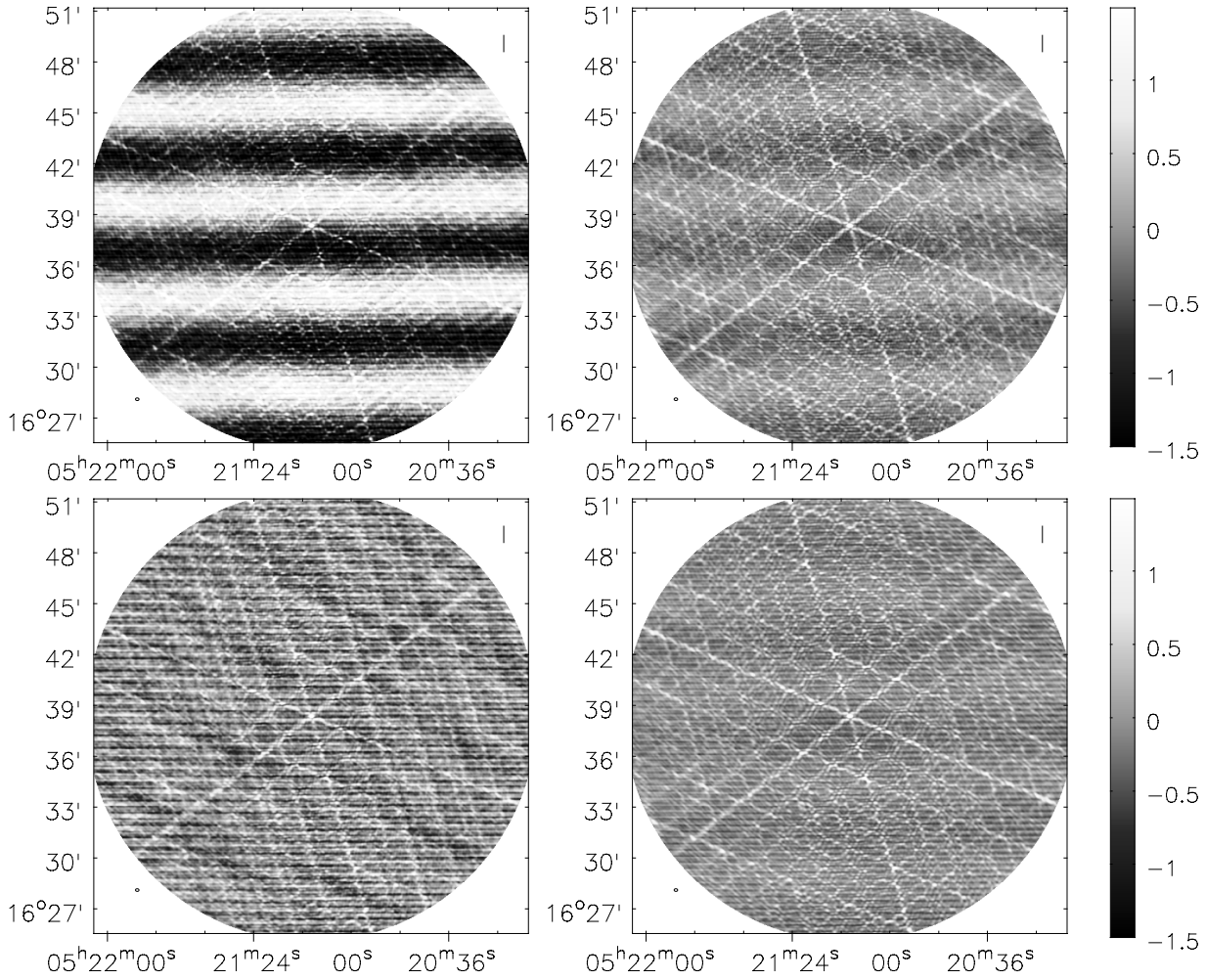

Figure 4.15: Images of channel 30. Top two have 2 eigenvectors projected out, bottom two have 4 eigenvectors projected out. Left was projected upon with the autocorrelations removed. Right was projected upon with the autocorrelations intact. Horizontal axis is in right ascension, vertial axis is in declination. All images have the same intensity scale ranging from -1.5 Jy/beam in black to +1.5 Jy/beam in white.
(Top)Note the significantly more pronounced RFI artifacts when the autocorrelations are missing as opposed to their almost complete disappearance when autocorrelations are preserved.
(Bottom) Although the dirty beam pattern is visible in both images, without autocorrelations, the noise is significantly higher.

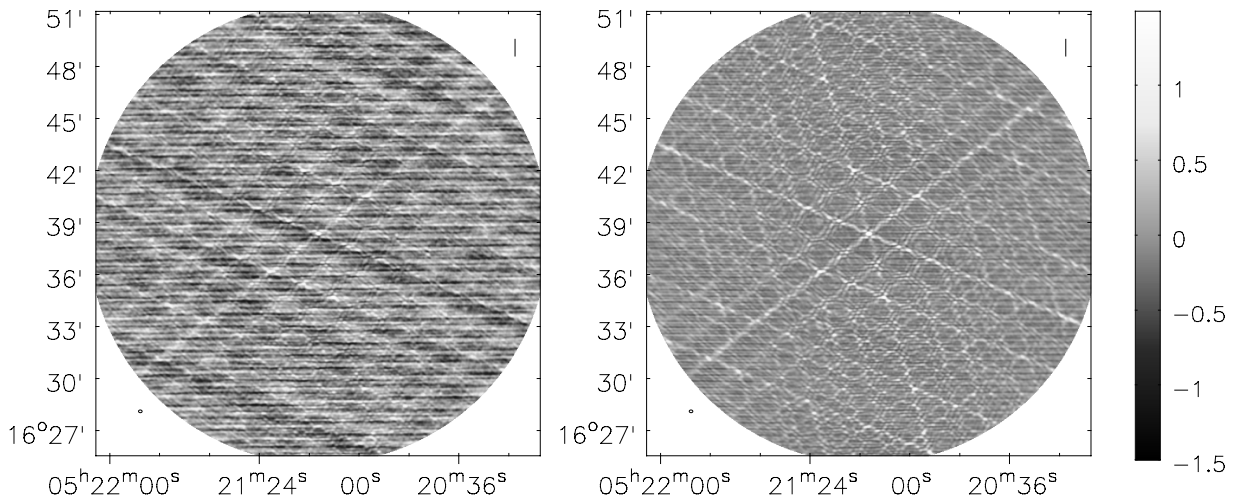

Figure 4.16: Images of channel 30. Both have 7 eigen-vectors projected out. Left was projected upon with the autocorrelations removed. Right was projected upon with the autocorrelations intact. Horizontal axis is in right ascension, vertial axis is in declination.
7 eigenvectors removed with autocorrelations retained is the first image where the artifacts introduced by the RFI (the 5 almost horizontal ripples across the image) are not visibly apparent. However, without autocorrelation the noise introduced by the lack of autocorrelations dominates the image.

found by recording the value of the central pixel. For very heavily corrupted images, this value is not noticeably different from the RFI introduced oscillations on the image.

Looking first at the total error, shown in figure 4.18, we see several interesting features in the case where the autocorrelations have been zeroed. The most obvious is that while the AC retained case has a peak at 0 eigenvectors removed and at 24 eigenvectors removed, the AC zeroed case has a third peak at 13 eigenvectors removed. We can also see that the minimum error of the AC zeroed case has about twice the error of the minimum seen by the AC retained case. Finally, while both have a few low error entries, the AC retained case has a much broader range, implying that it has more cancellation power.

We will divide the autocorrelation zeroed data in to 3 distinct zones. First, we have the RFI dominated zone that lasts from 0 eigenvectors removed to 3 eigenvectors removed. In this zone, RFI contributions to apparent noise are easily discernable in the images. Then there is a maximum SNR region, going from 3 eigenvectors removed to 11 eigenvectors removed. In this second zone the

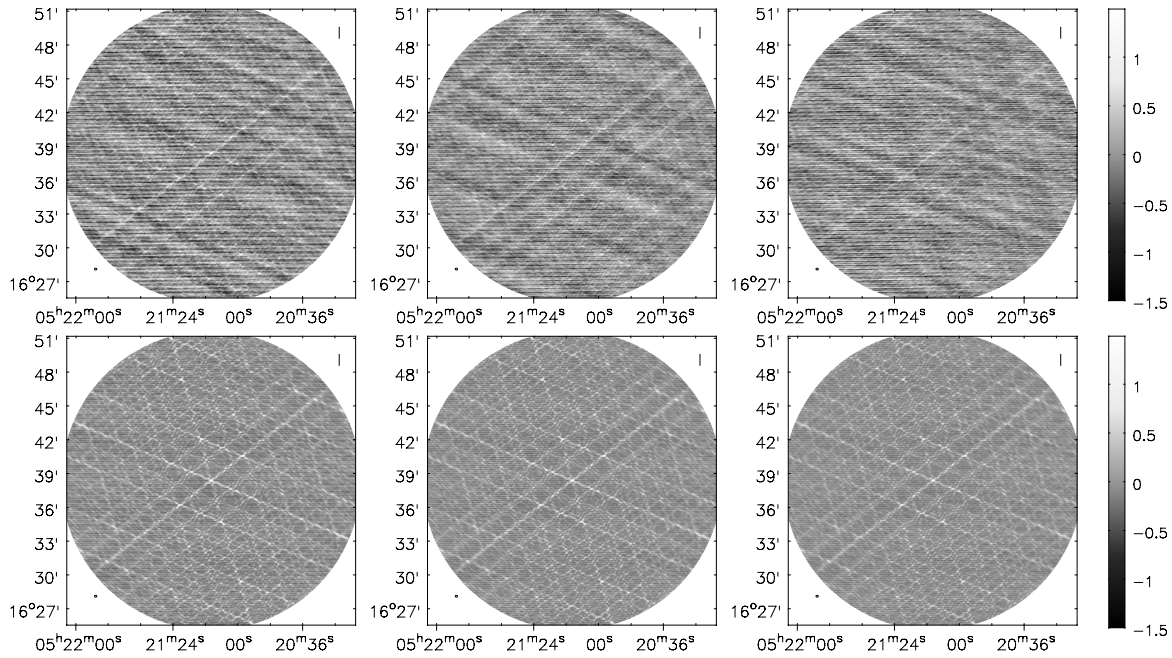

Figure 4.17: Images of channel 30. From left to right, images have 8, 9 and 10 eigen-vectors removed. Top row was projected upon with the autocorrelations removed, bottom row was projected upon with the autocorrelations intact. Horizontal axis is in right ascension, vertial axis is in declination.
These images correspond to a steady decrease in signal of interest power as the projected out eigenvectors begin to interact with the signal of interest.

peak height is dimished with every eigenvector, but the background noise is not similarly diminished. Finally, we have a minimum signal zone from 11 eigenvectors removed to 25 eigenvectors removed. In the minimum signal zone the the signal has been completely eliminated, leaving only noise to fill the image.

When we look at the central intensity, as plotted in figure 4.19, we gain some insight about what is driving the error levels in the autocorrelation zeroed case. e can see three distinct regimes with different marginal central pixel power per eigenvector removed. First, there is the region from 0 to 5 eigenvectors removed, then there is a region from 6 to 11 eigenvectors removed. Finally, there is a region from 11 to 25 eigenvectors removed. These regions correspond with the three regions observed in the error plot, albeit with slight differences.

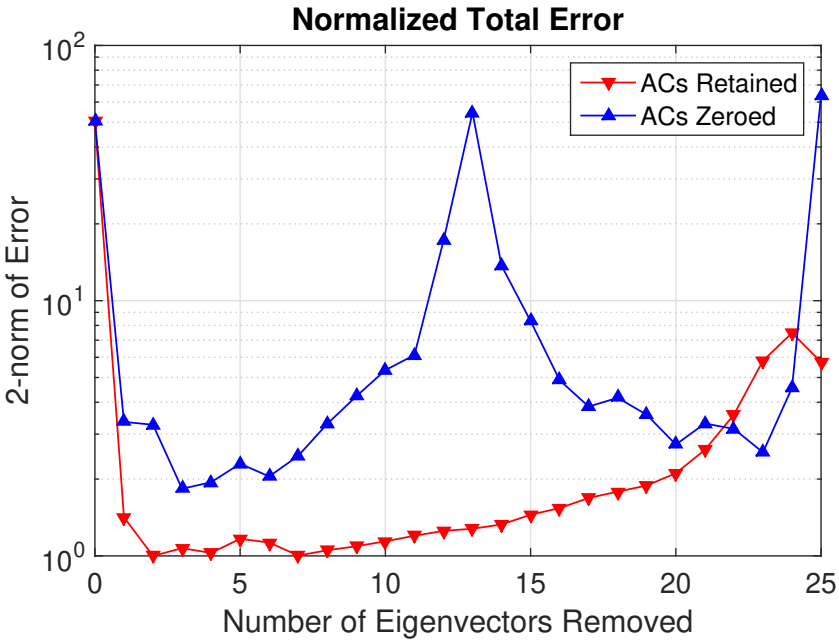

Figure 4.18: A plot of the normalized total image-space error at each amount of removed eigenvectors. The upper line is with the autocorrelations zeroed out for projection, the lower line retained all autocorrelations. The central pixel of the image was normalized before the error was computed. Without autocorrelation, the minimum error is significantly higher.

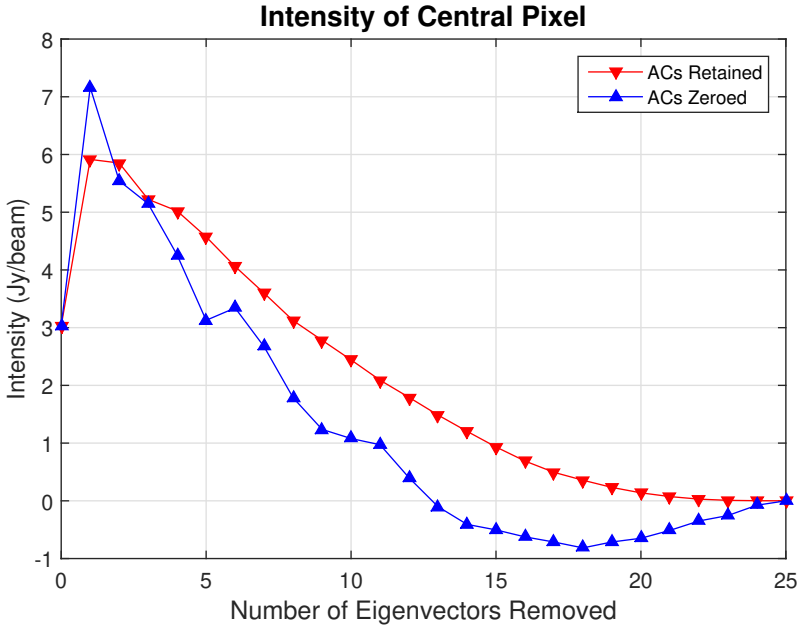

Figure 4.19: Plot of central pixel intensity as a function of eigenvectors removed. The autocorrelations retained case has a smooth, almost exponential relationship between 4 eigenvectors removed and 25 eigenvectors removed. The autocorrelations zeored case has a similar shape from 6 to 11 eigenvectors removed and from 11 to 18 eigenvectors removed, but they are offset by about +1 and -1 respectively.

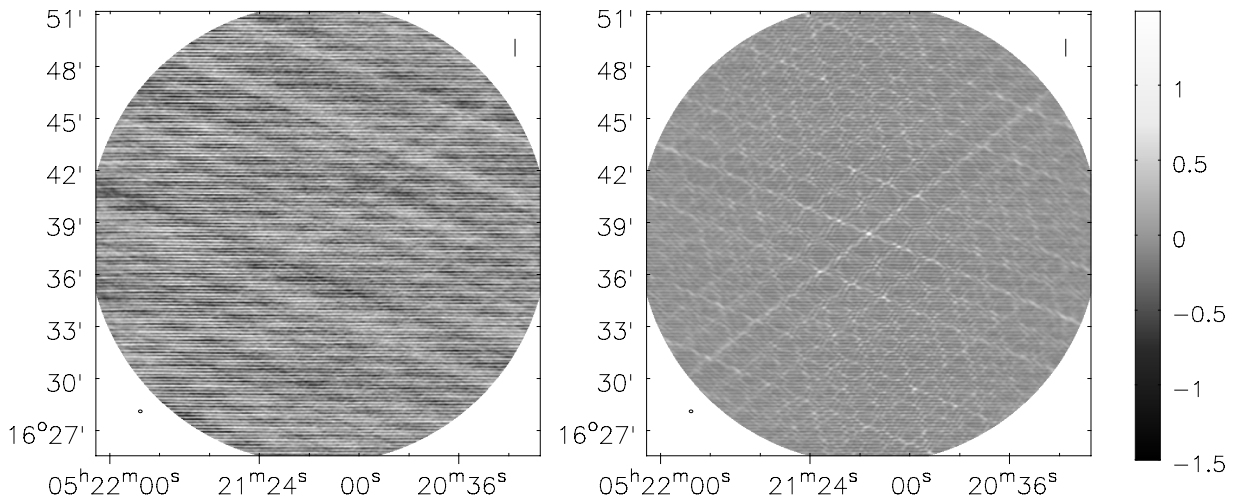

Figure 4.20: Images of channel 30. Both have 13 eigen-vectors projected out. Left was projected upon with the autocorrelations removed. Right was projected upon with the autocorrelations intact. Horizontal axis is in right ascension, vertial axis is in declination.
13 eigenvectors removed with autocorrelations retained was a low point for normalized noise, having noise level between 2 and 4 eigenvectors. However, without autocorrelation the noise introduced by the lack of autocorrelations dominates the image.

The central pixel intensity falls almost linearly until 5 eigenvectors removed, which corresponds fairly well with the RFI dominated zone mentioned above. Then the peak intensity falls in a somewhat exponential way from 6 eigenvectors removed to 11 eigenvectors removed. Finally, the peak intensity follows a different curve from 11 eigenvectors removed to 25 eigenvectors removed. The autocorrelation retained case, on the other hand, has an almost exponential decline from 4 eigenvectors removed to 25 eigenvectors removed.

One thing examining both figure 4.18 and figure 4.19 makes clear is why the error has a peak at 13 eigenvectors removed: the peak intensity is very close to zero. Figure 4.20 shows both the autocorrelation zeroed case and the autocorrelation retained case for 13 eigenvectors removed.

4.5 Sorting Eigenvalues for Effective RFI Removal

One of the the problems we encountered in applying subspace projection based mitigation to VLA data was that the eigenvalues of the visibility matrix are not exclusively non-negative. Fur-

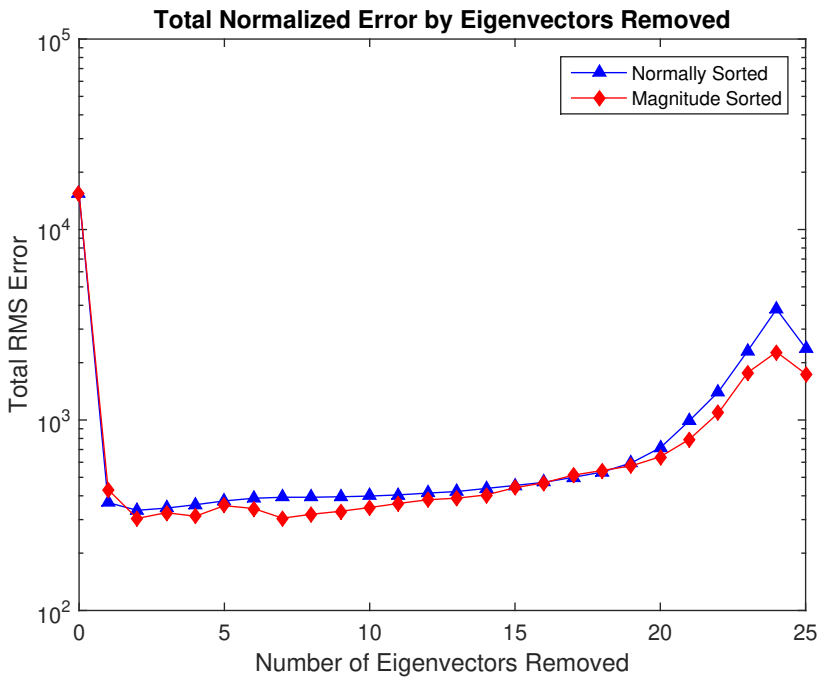

Figure 4.21: Plot of the total normalized error of channel 30 under successive eigenvector removals with normally sorted eigenvalues (triangles) and with magnitude sorted eigenvalues (diamonds). The error in the normally sorted case is greater than the error in the magnitude sorted case for all values of eigenvectors removed greater than 1.

thermore, simply ignoring the negative eigenvalues resulted in poor algorithm performance, as shown in figure 4.22.

To achieve effective cancellation we found that ordering the eigenvalues by magnitude allowed the removal of RFI artifacts with minimal signal corruption. When the eigenvalues were sorted by value, the RFI was greatly reduced for the first 2 eigenvectors removed, but saw no improvements after that. In fact, the total normalized error slowly increases from 3 eigenvectors removed to 18 eigenvectors removed. Above 18 eigenvectors removed, it increases quickly.

When we plot the total normalized error of normally sorted data versus magnitude sorted data, as in figure 4.21, we see that although the first eigenvector removed effectively reduces the error in the normally sorted case, more eigenvectors do not have as positive effect as under the magnitude sorted case.

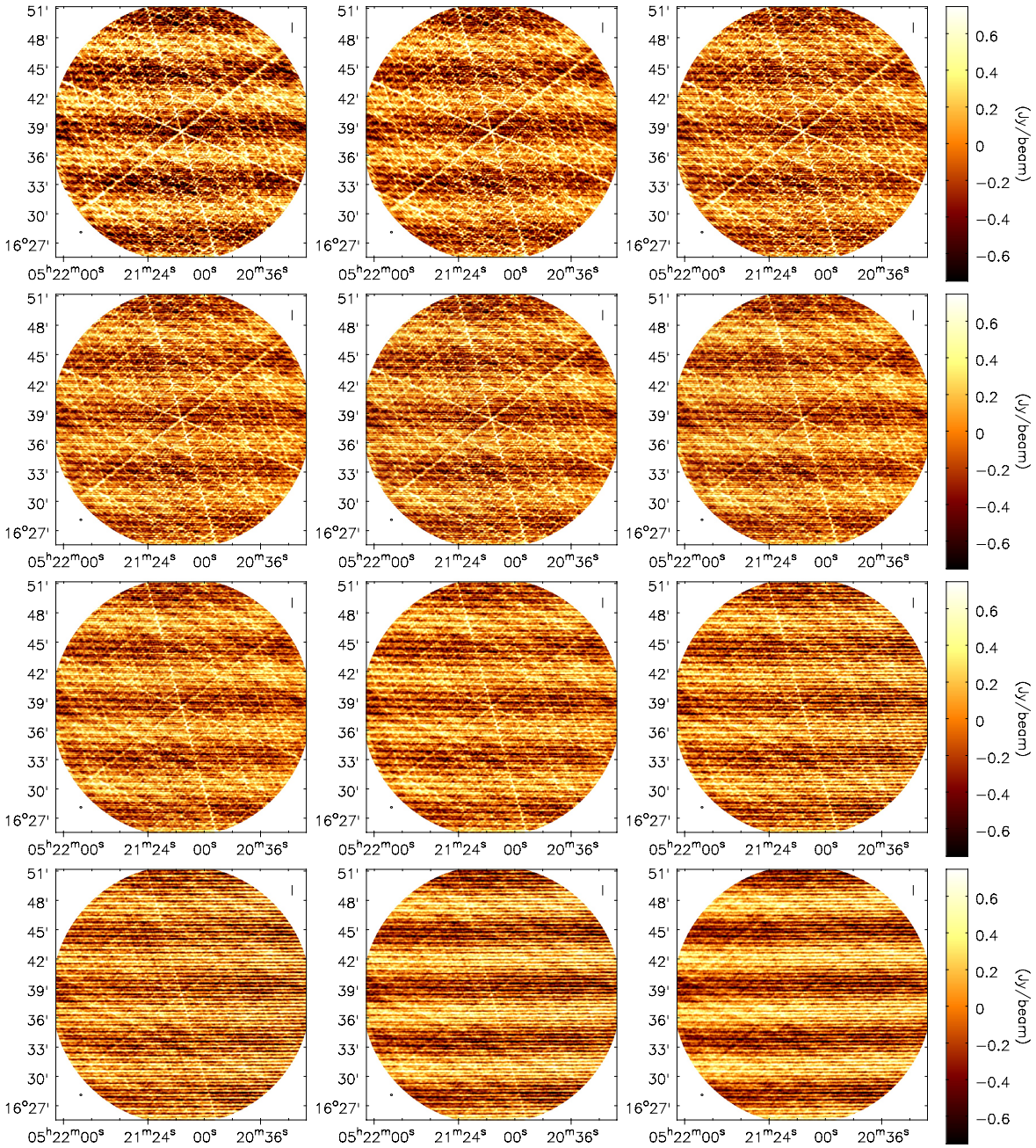

Figure 4.22: Images of channel 30 with 1, 2, 3, 7, 12, 18, 19, 20, 21, 22, 23, and 24 eigenvectors removed when the eigenvalues are sorted by value instead of magnitude. With 25 total eigenvectors, the 24 eigenvectors removed image being a good approximation of the RFI is indicative that the negative eigenvectors are not simply indicative of where the true zero point of the system is.

The reason for negative eigenvectors is unknown; the visibility matrix is formed from an outer product of the antenna values, and thus should be positive semi-definite. Negative eigenvalues are suspected to be partially due to numeric error in calculating the visibilities as some of the diagonal elements are complex, with imaginary parts about 2^{-8} times the real part. Another potential source of error are VLA correlator response to roll-over. More research would be required to determine the true source of this error.

4.6 Sub-space projection on widely dispersed array

In May of 2018 another data set was acquired with the VLA in A-configuration (which has baselines about 10-times the length of C-configuration baselines). In this data set, the same source was observed at the same frequencies with the same sample rate. However, the observation was done about 2 hours earlier in local sidereal time. The goal with this second observation was to introduce decorrelation into the RFI by increasing the baseline lengths.

As explained in [10], interferers which are too far removed in sky-separation from the desired field of view at too high a bandwidth experience decorrelation that reduces the effectiveness of subspace projection. By segregating the array into sub-arrays such that each sub-array can have sub-space projection applied to it individually we hoped to apply the power of sub-space projection with lesser limitations on sky-separation and bandwidth.

On the first point, introducing decorrelation by increasing baseline length, we succeeded spectacularly. Segregating the array and using subarray processing to remove RFI, on the other hand has been unsuccessful so far.

Initial attempts to apply subarray processing were stymied by confusion about an appropriate threshold for segregating subarrays. While [10] includes both metrics and a complete explanation of the algorithm involved, it does not mention what decorrelation threshold was used in its simulations. Nevertheless, after some trial-and-error, we were able to remove RFI almost as effectively with subarray processing as with full array subspace projection.

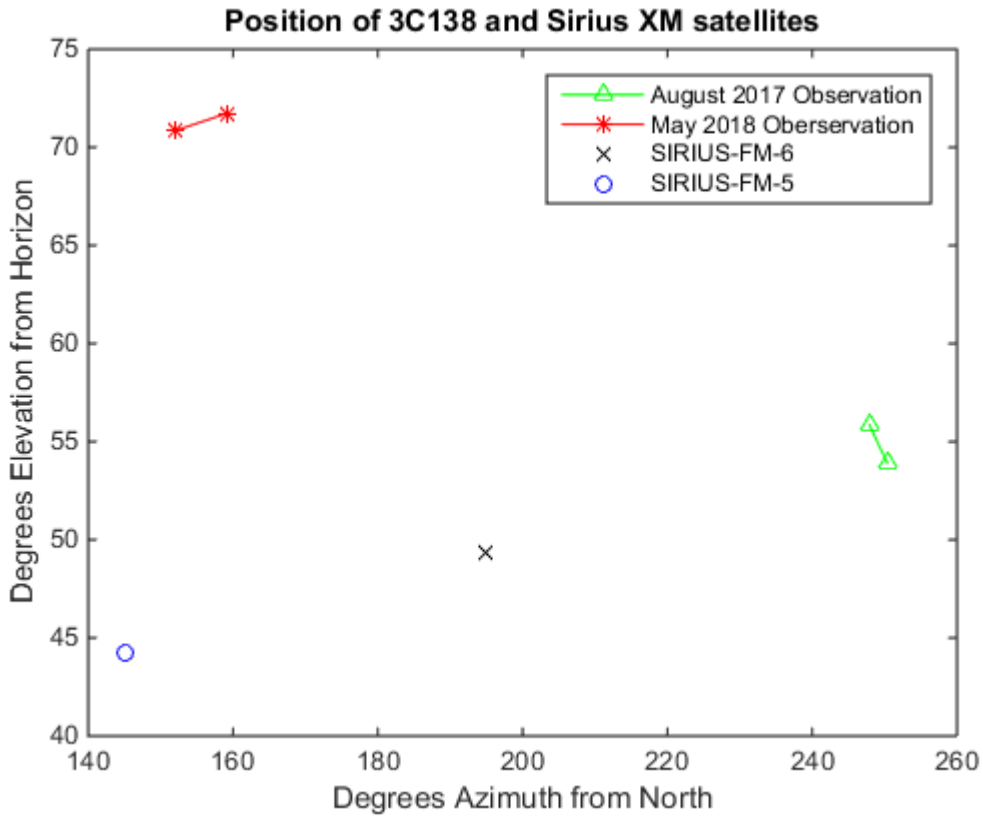

Figure 4.23: Sky Positions in azimuth and elevation of the source of interest and the two Sirius Satellites during observations in august 2017 and may 2018 from the center of the VLA. Positions from each antenna of the satellites vary slightly.

Turning first to unmodified subspace projection on this data, we see a marked decrease in image quality compared to C-configuration data. As shown in figure 4.24, with unmodified subspace projection, we can remove the majority of the RFI corruption, but not all of it.

When we attempted to apply subarray processing driven subspace projection to the data, it reduced RFI to central peak height by the same amount as unmodified subspace projection, as shown in figure 4.25. Unfortunately it also reduced the central peak (i.e. the source of interest intensity) significantly more than unmodified subspace projection. Using similar time averages between the two methods resulted in a 15% central peak decrease with unmodified subspace projection and a 67% decrease for subarray processing driven subspace projection; this corresponds to a central pixel height of 3.53 Jy/beam for an uncorrupted channel, 3.02 Jy/beam for unmodified subspace projection, and 1.16 Jy/beam for subarray processing driven subspace projection.

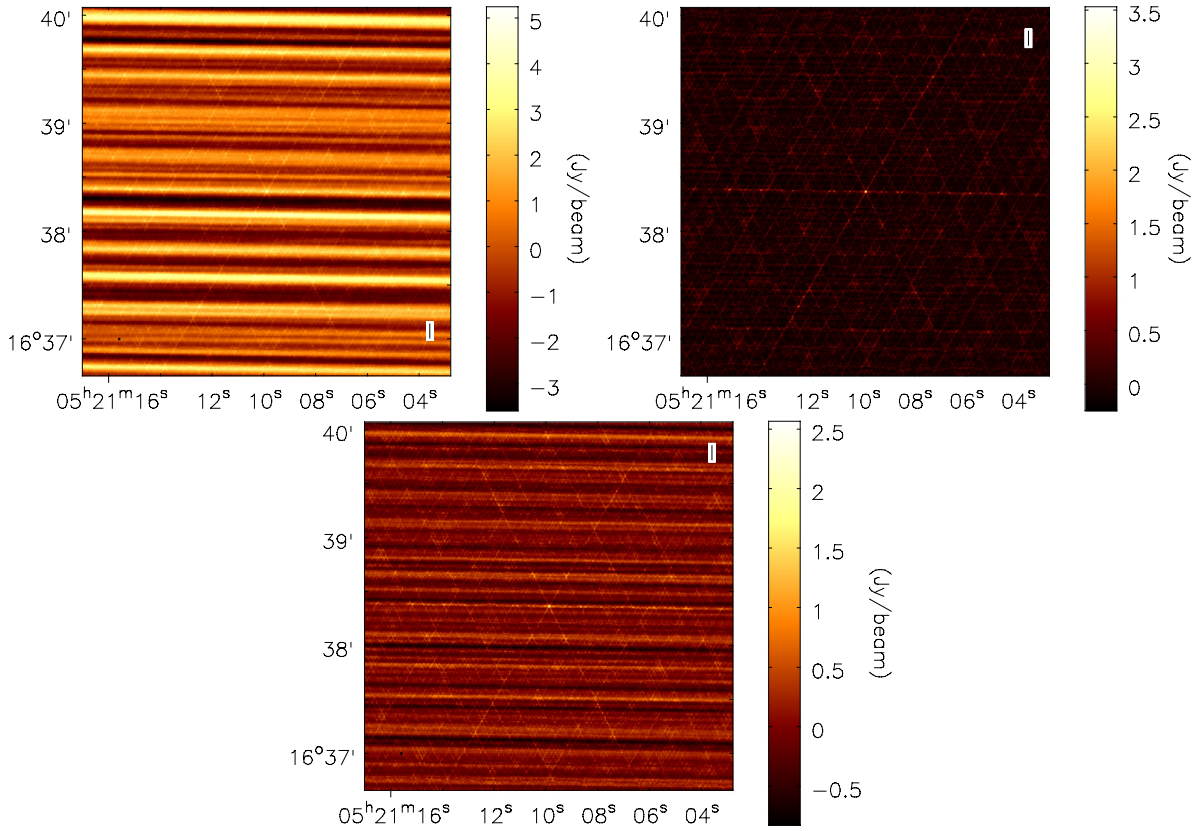

Figure 4.24: Image of 3C138 from the A-configuration data using a corrupted channel (left) and an RFI-free channel (right), and the corrupted channel after mitication (bottom). 3C138 is barely visible in the corrupted channel, but the RFI dominates. The source of interest is visible in the mitigated image, but there is still significant residual effects from the RFI.

4.6.1 Single Sub-Array

In an attempt to show that there is enough information in a reduced array to engage in sub-space projection, we selected the central core of antennas to do sub-space projection on. The central 9 antennas has a maximum baseline in A-configuration of about 5500 m, compared to the maximum baseline for C-configuration of 3300 m.

Sub-space projection was applied to the central core of the array, with moderate results. As shown in figure 4.26, using the central 9 antennas results in a different dirty beam pattern than the full array, but maintains a clear peak in the uncorrupted channel. The corrupted channel resembles a blurred version of the corrupted channel with the full array. While the dirty beam pattern is

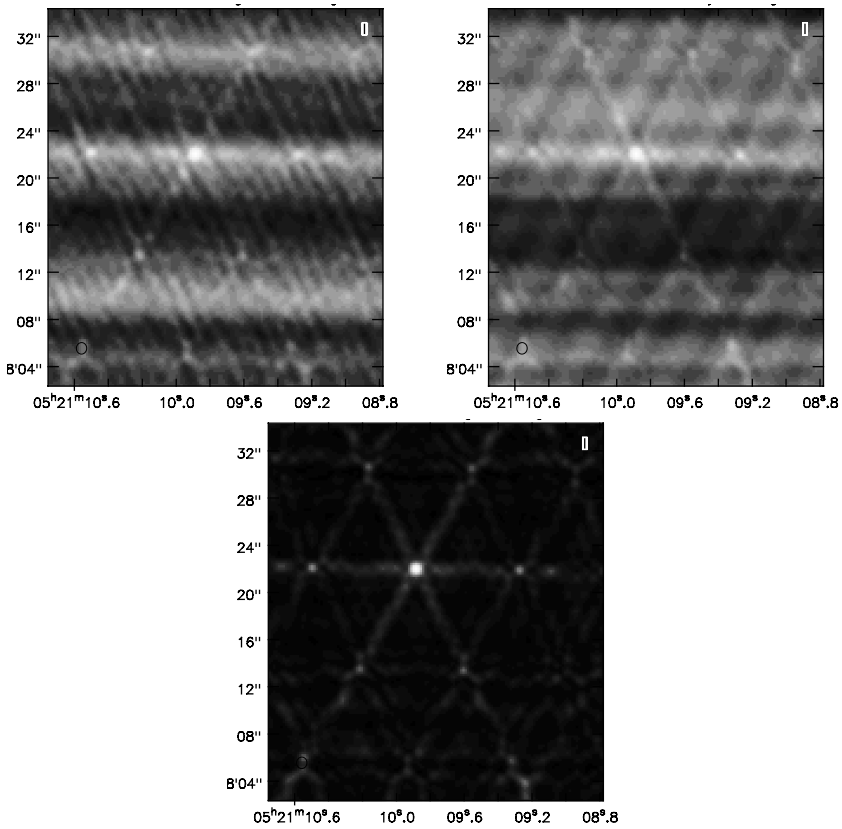

Figure 4.25: Images of 3C138 using normal subspace projection (left), sub-array processing modified subspace projection (right), and an uncorrupted channel. While the source is identifiable in all 3 images, neither subspace projection image has removed enough RFI influence to show the dirty beam pattern.

more visible in the mitigated image, the RFI has not been reduced as effectively as in full array c-configuration data.

4.6.2 Recommendations

To apply sub-array processing to the VLA, we needed to make a slight change to the algorithm: instead of simply drawing bands across the array and collecting the antennas in each band in a subarray, we drew bands such that the first edge of the band was the first antenna not assigned to a subarray. This had the immediate effect of increasing the size of subarrays.

One of the key differences between the VLA and the proposed ngVLA is the rate of radial symmetry. In the VLA, radial symmetry occurs about every 120°. In proposed ngVLA designs,

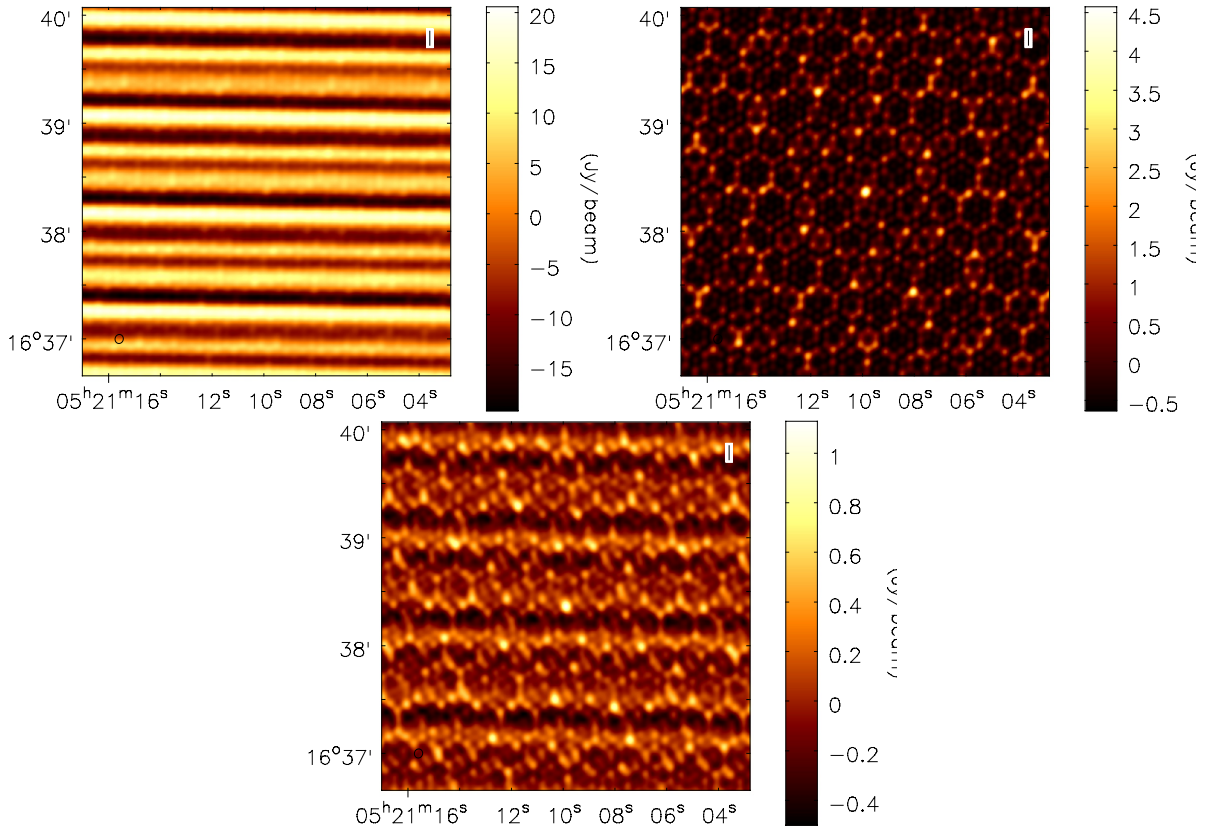

Figure 4.26: Image of 3C138 from the A-configuration data using only the central 9 antennas. Images formed from corrupted channel 22(left), RFI-free channel 132(right), and corrupted channel 22 with sub-space projection applied (bottom). The mitigated image has similar composition to the RFI-free channel, but has some residual RFI induced error.

radial symmetry occurys an order of magnitude more often. This will result in a more filled-in exterior than the VLA, allowing subarray processing the opportunity to have subarrays large enough to apply subspace projection to regardless of direction difference between the signal-of-interest and the RFI source. Albeit, the ngVLA will also have baselines orders of magnitude longer than the current VLA, but if bandwidth is chosen appropriately for baseline length and sky-separation, then direction of arrival will not impact subarray size.

CHAPTER 5. CONCLUSIONS

The scientific results shown herein shown the power and the limitations of using sub-space projection on a widely distributed array. On the one hand, when sub-space projection works, it removes interference with nary a trace. On the other hand, when sub-space projection fails it produces completely unusable data. If future radio telescopes have their correlators designed to handle the narrow bandwidths and high time resolution required to make sub-space projection work, they will be able to reject RFI and peer through the geo-stationary satellite belt.

This work represents some of the first application of subspace projection to actual data. Furthermore, it explores the relationship between short term integration window length and projected out eigen-vectors. Of particular note, is the good correspondence between short term integration window length and smearing caused by motion.

Ultimately, sub-space projection is a trade off between noise and interference. As we remove more interference, signal strength decreases while noise remains constant. On the other hand, artifacts from RFI drop off much faster than the signal of interest. We could re-calibrate levels after projection and be able to

5.1 Future Work

It was beyond the scope of this work to experiment with varying interferer levels and how various projection algorithms interacted therewith. The data set we examined had a range of interferer power levels on the edge of the interferer band. We could have gone through each channel, identified the interferer level, and analyzed the performance of the various algorithms.

A straightforward expansion of this work is applying sub-space projection to other VLA data sets. The primary problems with doing so are generally wider bandwidths used by astronomers and the lack of self power terms in general VLA data. VLA data normally only reports half of

the self power terms at every time sample. Full self power terms are collected somewhat less frequently. If an algorithm could make good estimates of the high resolution self power terms from the less frequent samples, VLA data could potentially be cleaned of RFI.

Finally, if more VLA time were available, getting a data set that included both a calibration source and another star-field would allow further testing. Particularly, while the results herein are useful, showing that recalibration is possible would seal the case that sub-space projection is worthwhile and useful.

REFERENCES

[1] A. R. Thompson, J. M. Moran, and G. W. Swenson, *Interferometry and synthesis in radio astronomy*. Wiley-VCH, 2004. 26

[2] J. M. Ford and K. D. Buch, "Rfi mitigation techniques in radio astronomy," in *Geoscience and Remote Sensing Symposium (IGARSS), 2014 IEEE International*. IEEE, 2014, pp. 231–234. 2

[3] P. Kalberla, N. McClure-Griffiths, D. Pisano, M. Calabretta, H. A. Ford, F. J. Lockman, L. Staveley-Smith, J. Kerp, B. Winkel, T. Murphy *et al.*, "Gass: the parkes galactic all-sky survey-ii. stray-radiation correction and second data release," *Astronomy & Astrophysics*, vol. 521, p. A17, 2010. 2

[4] P. M. Kalberla, "Rfi mitigation for the parkes galactic all-sky survey (gass)," *arXiv preprint arXiv:1102.4949*, 2011. 2

[5] C. Barnbaum and R. F. Bradley, "A new approach to interference excision in radio astronomy: Real-time adaptive cancellation," *The Astronomical Journal*, vol. 116, no. 5, p. 2598, 1998. [Online]. Available: http://stacks.iop.org/1538-3881/116/i=5/a=2598 2

[6] R. Athreya, "A new approach to mitigation of radio frequency interference in interferometric data," *The Astrophysical Journal*, vol. 696, no. 1, p. 885, 2009. [Online]. Available: http://stacks.iop.org/0004-637X/696/i=1/a=885 2

[7] J. Raza, A. J. Boonstra, and A. J. van der Veen, "Spatial filtering of rf interference in radio astronomy," *IEEE Signal Processing Letters*, vol. 9, no. 2, pp. 64–67, Feb 2002. 3

[8] "Resolution." [Online]. Available: https://science.nrao.edu/facilities/vla/docs/manuals/oss/performance/resolution 25

[9] A. S. Cohen, W. M. Lane, W. D. Cotton, N. E. Kassim, T. J. Lazio, R. A. Perley, J. Condon, and W. C. Erickson, "Vlss-redux postage stamp server." [Online]. Available: https://www.cv.nrao.edu/vlss/VLSSpostage.shtml 25

[10] M. C. Burnett, B. D. Jeffs, R. A. Black, and K. F. Warnick, "Subarray processing for projection-based rfi mitigation in radio astronomical interferometers," *The Astronomical Journal*, vol. 155, no. 4, p. 146, 2018. 46