A Novel Radio-Frequency Interference Flagging Technique for Radio-Interferometric Data from the Murchison Widefield Array

By

Theodora C. Kunicki

B.S., Massachusetts Institute of Technology, 2016

Thesis

Submitted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy in the Department of Physics at Brown University

PROVIDENCE, RHODE ISLAND

February 2025

This dissertation by Theodora Kunicki is accepted in its present form by the Department of Physics as satisfying the dissertation requirement for the degree of Doctor of Philosophy.

Date ____________ ______________________________
(JONATHAN C. POBER), Advisor

Recommended to the Graduate Council

Date ____________ ______________________________
(IAN P. DELL'ANTONIO), Reader

Date ____________ ______________________________
(JONATHAN C. POBER), Reader

Date ____________ ______________________________
(GREGORY S. TUCKER), Reader

Approved by the Graduate Council

Date ____________ ______________________________
(THOMAS A. LEWIS), Dean of the Graduate School

Curriculum Vitae

Education

2018–2024
Ph.D., (PHYSICS)
Brown University (Providence, RI, USA)

2008–2016
B.S., (PHYSICS)
Massachusetts Institute of Technology (Cambridge, MA, USA)

Publications

Papers

Kunicki, T., Pober, J.C. (2024). χ^2 from Redundant Calibration as a Tool in the Detection of Faint Radio-frequency Interference. Publications of the Astronomical Society of Australia.

Teaching

Fall 2018, Spring 2019, Spring 2022
Teaching Assistant, *Physics Department*, (Brown University, Providence, RI)
Course: Basic Physics B (Electricity and Magnetism)

Summer 2019
Instructor, *Summer@Brown*, (Brown University, Providence, RI)
Course: An Introduction to Electricity and Magnetism

Awards and Fellowships

2019: Award of Excellence as a Graduate Teaching Assistant
Brown University Department of Physics

2019: NASA Rhode Island Space Grant
NASA Rhode Island Space Grant Consortium

2022: Award of Excellence as a Graduate Teaching Assistant
Brown University Department of Physics

Acknowledgments

Thank you foremost to my mother, Catherine Kunicki, for her consistent support in helping me overcome my executive dysfunction struggles. Thank you to my advisor, Jonathan Pober, for his unwavering support and understanding. Thank you to Bryna Hazelton, Miguel Morales, Michael Wilensky, and Pyxie Star for their support, encouragement, and handy SSINS tips. Thank you also to our Cal-Bridge summer students Alexander Hawksley and Hal France for their initial investigations with χ^2 RFI detection.

I would also like to acknowledge the developers of pyuvdata and `hera_cal`. These software projects were indispensable to this work.

This scientific work uses data obtained from Inyarrimanha Ilgari Bundara / the Murchison Radio-astronomy Observatory. We acknowledge the Wajarri Yamaji People as the Traditional Owners and native title holders of the Observatory site. Establishment of CSIRO's Murchison Radio-astronomy Observatory is an initiative of the Australian Government, with support from the Government of Western Australia and the Science and Industry Endowment Fund. Support for the operation of the MWA is provided by the Australian Government (NCRIS), under a contract to Curtin University administered by Astronomy Australia Limited. This work was supported by resources provided by the Pawsey Supercomputing Research Centre with funding from the Australian Government and the Government of Western Australia.

This work was conducted at Brown University. Brown University is located in Providence, Rhode Island, on lands that are within the ancestral homelands of the Narragansett Indian Tribe. The Narragansett Indian Tribe, whose ancestors stewarded these lands with great care, continues as a sovereign nation today. We commit to working together to honor our past and build our future with truth.

Contents

List of Figures

List of Tables

INTRODUCTION

In the beginning the Universe was created. This has made a lot of people very angry and been widely regarded as a bad move.

DOUGLAS ADAMS

THE RESTAURANT AT THE END OF THE UNIVERSE

The first thing we know for sure that happened was the Big Bang. Anything prior to that is — in the absence of experimental provability — mere metaphysics. The Big Bang was the font of every constituent of our Universe: spacetime, radiation, baryonic matter, presumably dark matter, and so on. At first, the Universe was indescribably hot and dense, and all of these elements were in thermodynamic equilibrium for at least some time. As the Universe expanded, it cooled according to the laws of thermodynamics, and different particle species "froze out" one by one, depending on how strongly they interacted with every other species. Neutrinos, for instance, are expected to have frozen out roughly 1 second after the Big Bang; photons took another 370,000 years.

The frozen-out primordial photons are called the Cosmic Microwave Background, and they are visible as microwave radiation in our sky: a nearly perfect black-body spectrum with a temperature of 2.73 K. Studying this radiation has been fundamental to our understanding of the Early Universe; its initial discovery was the first experimental evidence for an evolving

Universe and a Big Bang, and modern study of anisotropies in its power spectrum gives us insight into the nature of the Early Universe and its development into the Universe we observe around us.

Studying the CMB continues to teach cosmologists much about our Universe, but there are some aspects of the Early Universe which are not probed by its light. For instance, when the CMB was emitted, the Universe was very different from how we see it today: luminous objects hadn't formed yet, and matter was presumably in the process of gravitationally clumping into the galactic structures that characterize the local Universe. This epoch in the Universe's history is known as the Dark Ages, since stars were not yet emitting light. However, *Dark* is a bit of a misnomer; there was, of course, the free-streaming light of the CMB, in addition to another source of light: the 21 cm line emission of neutral hydrogen.

0.1 21 cm Early Universe Cosmology

Neutral atomic hydrogen consists of a single proton with a single electron bound to it. Atomic hydrogen in its lowest energy state exhibits hyperfine splitting, which is the splitting of the atom's ground state energy into two very close but unequal energy levels. These energy levels come about as a consequence of spin-spin interactions between the proton and the electron within the atom. When their spins are added together, two distinguishable fermions like the proton and electron in neutral hydrogen form a triplet state where their spins are parallel, and a singlet state where their spins are antiparallel. As implied by their names, the degeneracy of the triplet state is $n = 3$, and the degeneracy of the singlet state is $n = 1$. The triplet state has just slightly higher energy than the singlet state, by about 6 μeV, equivalent to a microwave photon with a frequency of around 1420 MHz, or a wavelength of about 21 cm, which gives this radiation its name. Due to quantum mechanical selection rules, the spontaneous transition between these two states is considered "forbidden," but despite the nomenclature, this transition *does* occur naturally and spontaneously, only rarely.

Because all neutral hydrogen is prone to occasionally releasing 21 cm radiation, it can be a useful experimental probe into the presence and density of neutral hydrogen gas. In the 20^{th} Century, the 21 cm line was successfully used to map the neutral hydrogen clouds in our galactic neighborhood, for instance (van de Hulst et al., 1954). However, it is also a potential probe of the cosmological Dark Ages, as well as the Epoch of Reionization, which marks the Dark Ages' end.

During the Dark Ages, the main constituent of of intergalactic medium (IGM) was neutral hydrogen gas. This hydrogen first collapsed into its neutral bound state when the average kinetic energy of the particles which constituted the thermodynamic soup of the Early Universe had decreased enough for the bound, atomic state of hydrogen to be energetically preferable to the ionized state of unbound electrons and protons. This was a significant event in the history of the Universe, because it precipitated the release of the CMB photons, which interact much more strongly with free, charged particles than with neutral atoms.

The IGM remained almost completely neutral until ionizing radiation began to be emitted by stars and quasars, stripping electrons from hydrogen atoms, and leaving them as free charged particles alongside their previously bound protons. This is the "reionization" referenced in the name "Epoch of Reionization." The details of reionization are sensitive to other important cosmological factors such as the process of early galaxy formation, so its observation is of high potential value to experimental cosmology. Two review articles which describe the Epoch of Reionization in depth are Choudhury (2022) and Wise (2019).

0.2 Challenges in Detecting the Epoch of Reionization 21 cm signal

In cosmology, instead of referencing the age of the Universe in years, we often use the cosmological redshift of an object to describe its age. Cosmological redshift, z, is defined such that $z + 1$ is equal to the multiplicative factor comparing the Universe's scale at the time of emission to that at the time of observation (i.e., now). That is, if we observe a galaxy with $z = 2$, then we understand that the Universe was only $\frac{1}{3}$ its current scale when that light was emitted. The CMB has a redshift of $z \approx 1100$, and its release is considered the beginning of the Cosmological Dark Ages. The Dark Ages persist until the Epoch of Reionization begins, currently believed to be at around $z = 20$. At $z = 6$, the EoR is believed to have ended, with essentially all of the intergalactic medium having been ionized.

Like all distant cosmic radiation, the 21 cm photons emitted during the Dark Ages and EoR have redshifted since their emission due to the Universe's expansion. A Dark Ages 21 cm photon emitted at $z = 100$ will have an observed frequency of only ~ 14 MHz, while an EoR photon emitted at $z = 10$ will have an observed frequency ~ 130 MHz, both redshifted from their original frequency of about 1420 MHz. Thus we can theoretically use the frequency of observed 21 cm EoR photons as a time-axis — as we consider the structure revealed by probing the ancient 21 cm line, we can also observe its evolution over time by considering z.

Due to atmospheric effects, there is a lower limit to the radio frequencies observable from the Earth's surface. This limitation makes much of the frequency band corresponding to Dark Ages 21 cm radiation inaccessible to Earth-based observers. The Epoch of Reionization, however, corresponds to frequencies which *are* visible from Earth's surface. The Murchison Widefield Array, or MWA[1] (Tingay et al., 2013; Wayth et al., 2018), the instrument which

[1] `https://www.mwatelescope.org/`

is the primary data-source for this research, operates in a frequency band between 70 and 300 MHz, corresponding to 21 cm redshifts of roughly $20 < z < 4$. Other arrays built with observing the EoR in mind, like HERA[2] (DeBoer et al., 2017) and LOFAR[3] (van Haarlem et al., 2013), operate in similar frequency bands.

The primary long-term objective of experimental EoR research is to generate and analyze a power spectrum of the EoR's 21 cm radiation. A 21 cm power spectrum would shed light on the spatial distribution of neutral hydrogen gas, being related by a Fourier transform to the image-space correlation function. By mapping the EoR's 21 cm radiation, cosmologists hope to learn about the spatial distribution of neutral hydrogen in the IGM as the EoR progressed. For reviews of 21 cm cosmology, see, for example, Morales and Wyithe (2010), Furlanetto et al. (2006), Liu and Shaw (2020) and Pritchard and Loeb (2012).

Creating an EoR 21 cm power spectrum is a nontrivial task, however. 21 cm radiation, due to the transition's aforementioned "forbiddenness," is sparse even in dense clouds of atomic Hydrogen. Furthermore, between us and the Epoch of Reionization, there is a great deal of radiation in the same frequency band, but with very different origins, having nothing to do with the EoR. This obscuring radiation can be broken down further into astrophysical foregrounds and radio-frequency interference.

So-called foreground radiation is due largely to the interaction of free electrons in the local interstellar medium with galactic magnetic fields. Knowing something about the expected power spectra of both foregrounds and the EoR signal allows us to differentiate them. Because of their production mechanism — primarily synchrotron and bremsstrahlung radiation — foregrounds are spectrally smooth, which allows for them to be differentiated from the EoR, at least in theory.

[2] `https://reionization.org/`

[3] `https://www.astron.nl/telescopes/lofar/`

Foreground radiation is a natural phenomenon; radio-frequency interference, on the other hand, is anthropogenic. The frequency band used to observe the redshifted EoR signal from Earth is also used by broadcasters to convey information from place to place. The frequency band in which the MWA operates is considered a part of the VHF band, and according to Australian and international law, it is allocated for diverse uses. The MWA was built in a designated radio-quiet zone, but anthropogenic radiation still creeps in, and remains visible in MWA data. It is common, for instance, to see digital TV signal being picked up by the instrument. This data contamination is very deleterious to the effort to detect the 21 cm EoR signal; RFI doesn't have the spectral qualities that make foreground radiation easily distinguishable from the EoR 21 cm signal, and furthermore, it can be incredibly bright compared to any radiation coming from the sky. When RFI-contaminated data are used to create a power spectrum, the power from RFI contaminates the spectrum within the EoR window — the "window" being the portion of the power spectrum where the EoR signal is believed to be theoretically observable. Without careful identification and excision of even relatively weak RFI, the goal of creating a power spectrum that is sensitive to the EoR signal is impossible.

Removing RFI-contaminated data from a power-spectrum analysis is essential if the measurement is to be sensitive to the 21 cm EoR signal (Wilensky et al., 2020). While LOFAR is situated in the Netherlands, where RFI is omnipresent, HERA and the MWA are located in more radio-quiet sites in South Africa and Western Australia, respectively. Even in these radio-quiet sites, there are many sources of RFI, including digital television (DTV) (which may even be reflected by passing aircraft (Wilensky et al., 2019)), ORBCOMM satellite transmissions (Sokolowski et al., 2015), and FM radio, which may reflect from satellites (Zhang et al., 2018). The MWA, in particular, is located far from any transmitters, and still observes DTV regularly, as well as narrow- and broad-band RFI from time to time

(Offringa et al., 2015).

There are two software packages regularly used in RFI detection at the MWA, AOFlagger (Offringa et al., 2015), and SSINS (Wilensky et al., 2019). While these software tools are both powerful, it has been demonstrated that they fail to excise all RFI, leaving behind so-called "ultra-faint" RFI in the data (Wilensky et al., 2023). In Wilensky et al. (2023), the authors have used statistical methods to demonstrate that this "ultra-faint" RFI is present in existing datasets that have been used for analysis, and that even this very weak residual RFI is extremely deleterious to the effort to accurately measure the 21 cm EoR power spectrum.

The research presented in this document is largely concerned with the identification and excision of RFI-contamination from MWA data, centering on a novel method for RFI detection based on statistics produced by redundant calibration. The structure of this document is as follows: Chapter 1 is a brief overview of the basics of radio interferometry. Chapter 2 discusses the Murchison Widefield Array, and describes our dataset. The methods behind redundant calibration will be presented in Chapter 3. Chapter 4 details our novel technique for RFI detection using the outputs of redundant calibration. Chapter 5 is a brief conclusion. In the appendices, A, we present supplemental facts and figures which are important, but do not fit into the main document.

CHAPTER 1

An Overview of Radio Interferometry

I tried to discover, in the rumor of forests and waves, words that other men could not hear, and I pricked up my ears to listen to the revelation of their harmony.

Gustave Flaubert

November

It is possible to do impactful radio astronomy with a single antenna. The CMB was discovered with a single-antenna telescope, for instance, by Penzias and Wilson (1979). However, in order to have enough surface area to collect sufficient radiation to detect faint sources, single-antenna telescopes can become untenably large, presenting an engineering challenge. Furthermore, a single-dish instrument is not very good at resolving large-scale structure on the sky, its resolution being limited by the physical size of the antenna.

Although their design and operation is less intuitive than a single-dish instrument's, radio interferometers present a tidy solution to many of the problems which arise from single-dish instruments' limitations. An interferometric array's sensitivity is proportional to the sum of the collecting areas of every antenna in the array, and their resolution is determined by the distances and directions ("baseline vectors") of separation between antennas in the array. For a person accustomed to thinking of astronomy in terms of optical instruments, interferometers can be somewhat counter-intuitive. Briefly, instead of looking directly at the voltages measured by pointing a single radio antenna directly at the sky, we consider the interference patterns made by multiplying the voltages from two antennas, which are separated by some distance on the ground. By understanding this interference pattern, we can extract information about the radio sky, even being able to recreate an image-space representation with enough data.

1.1 A Simple Interferometer

Consider the sketch of a simple two-element interferometer in Figure 1.1. This instrument consists of two antennas separated by a baseline vector $\vec{b}$, and is observing a distant point source in the upper right of the image. For the purposes of simplifying the math, we consider the sky to be a plane parallel to the ground and infinitely far away, that is, we take the limit as the angle between the two beams of light originating at the point source approaches zero and those beams of light become parallel. Due to the limitations of drawing infinity, this assumption is loosened a little in Figure 1.1, where the two long legs of the triangle formed by the two antennas and the source are clearly converging — a little imagination will be required to visualize the system exactly as we will treat it mathematically.

We can calculate the difference in light-travel path length between the point source and both antennas by considering $\vec{b} \cdot \hat{s}$, where $\hat{s}$ represents the unit vector pointing directly at the point source from either antenna. This is illustrated by the waveform drawn over the short leg of the right triangle in the sketch — for illustrative purposes in Figure 1.1, these paths are about 2.5 wavelengths different in length. We arrive at the number of wavelengths

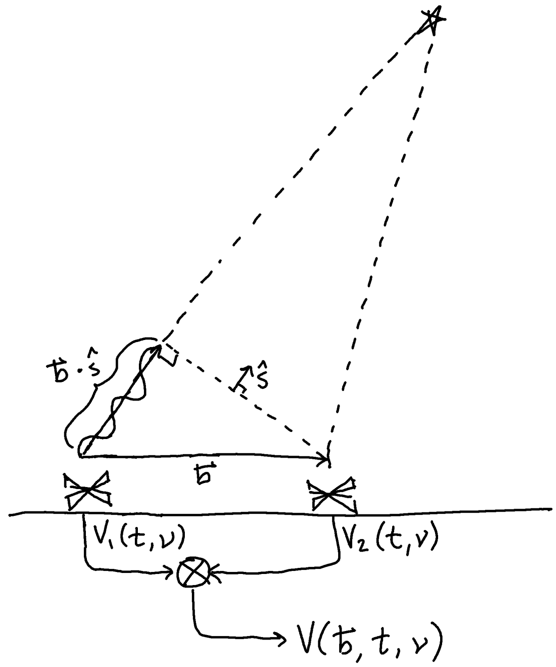


Figure 1.1: A simple two-element interferometer. See text for details.

difference by dividing $\vec{b} \cdot \hat{s}/\lambda$, where λ is the wavelength of radio light we are observing. This is the heart of interferometry: even though the distance to the source is enormous, the two distinct path lengths differ by a scale comparable to the wavelength of the observed light. From this fact, we can use a well-designed interferometer to deduce the structure of the sky from these interference patterns alone.

Both antennas produce a voltage measurement, represented in Figure 1.1 by $V_1(t, \nu)$ and $V_2(t, \nu)$, where we have made the dependence on both time and frequency explicit. To produce a visibility, represented in the sketch by $V(\vec{b}, t, \nu)$, we multiply the voltages $V_1 \times \tilde{V_2}$ at each time step ($\tilde{V_2}$ being the complex conjugate of V_2). $V(\vec{b}, t, \nu)$ is known as the "visibility" for that baseline, time, and frequency, and is generally a complex number.

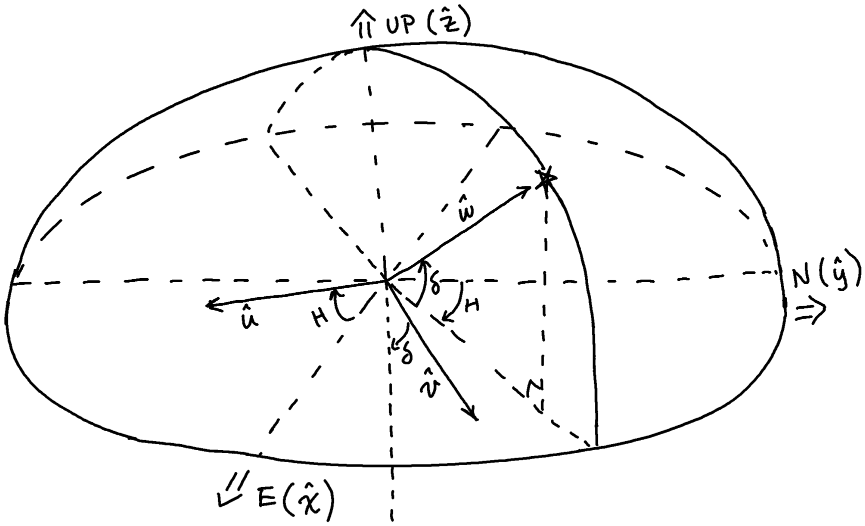


Figure 1.2: The *uvw* basis in context

1.2 The (u, v, w) basis

Our baseline $\vec{b}$ is a position vector, using units of length to describe the physical separation of the two antennas. The component of $\vec{b}$ which is perpendicular to $\hat{s}$ lies in what is called the *uv*-plane, which we can define as the locus of points perpendicular to $\hat{s}$. We can define a coordinate system in terms of the *uv*-plane and $\hat{w}$, which is defined to be identical to $\hat{s}$; our vector $\vec{b}$, when measured in units of wavelengths and expressed in this coordinate system, is called $\vec{u}$ and will be very important to our understanding of radio interferometry.

We characterize a visibility in terms of $\vec{u}$, which will differ with time of day and observing frequency, as the components are derived from both the static physical baseline $\vec{b}$ and the position of the source in the sky as a declination and hour-angle, (δ, H). Refer to Figure 1.2 to follow along with the transformations that follow. In order to define the orthonormal basis $(\hat{u}, \hat{v}, \hat{w})$, we begin with the basis aligned with our $(\hat{x}, \hat{y}, \hat{z})$ basis, where $\hat{x}$ points eastward (as is customary with the MWA) and rotate it through both angles H and δ sequentially, as follows[1]:

[1] note that the MWA is located in the Southern hemisphere, where the sky appears to rotate clockwise

First, we rotate about the z axis by the angle corresponding to H. H is negative if the source is still rising in the sky, and positive if it has already passed its highest point in the sky. If H is negative, the basis rotates clockwise when viewed from above. Likewise, if H is positive, the basis rotates counterclockwise when viewed from above. In the illustrated example, H is negative. Next, we rotate about the newly-defined u-axis by $\delta - \frac{\pi}{2}$ radians (the negative complement of δ), until $\hat{w}$, which had previously been aligned with $\hat{z}$, is pointed at an angle δ above the horizon. By construction, δ has a value smaller than $\frac{\pi}{2}$, so another way to say this is that we are rotating the basis about the u axis by an angle of magnitude $\frac{\pi}{2} - \delta$ in a clockwise direction if viewing the basis from directly above $\hat{u}$, thereby "dropping" the angle between $\hat{w}$ and the horizon from $\frac{\pi}{2}$ to δ. After rotating the $(\hat{u}, \hat{v}, \hat{w})$ basis by both of these angles, $\hat{w}$ should now be aligned with $\hat{s}$, pointing directly at the source. The uv-plane, perpendicular to $\hat{s}$ and its twin $\hat{w}$, is spanned by the $\hat{u}$ and $\hat{v}$ vectors.

By considering the rotations from in the previous paragraph, and Figure 1.2, we can explicitly derive the components of $(\hat{u}, \hat{v}, \hat{w})$ in terms of $(\hat{x}, \hat{y}, \hat{z})$.

$$\begin{aligned}
\hat{u} &= \hat{x}\cos\mathrm{H} + \hat{y}\sin\mathrm{H} \\
\hat{v} &= \sin\delta\left(-\hat{x}\sin\mathrm{H} + \hat{y}\cos\mathrm{H}\right) - \hat{z}\cos\delta \\
\hat{w} &= \cos\delta\left(-\hat{x}\sin\mathrm{H} + \hat{y}\cos\mathrm{H}\right) + \hat{z}\sin\delta
\end{aligned} \tag{1.1}$$

In order to find the uvw-components of $\frac{\vec{b}}{\lambda}$, $\frac{\vec{b}}{\lambda} = \frac{b_u}{\lambda}\hat{u} + \frac{b_v}{\lambda}\hat{v} + \frac{b_w}{\lambda}\hat{w}$, we break it down into its xyz-components, $\frac{\vec{b}}{\lambda} = \frac{b_x}{\lambda}\hat{x} + \frac{b_y}{\lambda}\hat{y}$ (here we are assuming that our baseline has no z-component, as is the case in the MWA[2]), and take the dot product with each vector in our

about the South celestial pole, so a source will be at its highest in the sky when it is its furthest North. The hour angle, H, measures the source's angle from this highest, northernmost point. Most text sources describe this transformation for an observer in the *Northern* hemisphere, which is subtly different, mostly due to the definition of the hour angle.

[2] In reality, due to the constraints of real-world construction, there are of course some deviations from coplanarity in the MWA, but in this work we will disregard them entirely, as they are comparatively very small.

new orthonormal basis.

$$
\begin{aligned}
u &= \frac{b_u}{\lambda} = \left(\frac{\vec{b}}{\lambda} \cdot \hat{u}\right) = \frac{b_x}{\lambda} \cos \mathrm{H} + \frac{b_y}{\lambda} \sin \mathrm{H} \\
v &= \frac{b_v}{\lambda} = \left(\frac{\vec{b}}{\lambda} \cdot \hat{v}\right) = \sin \delta \left(-\frac{b_x}{\lambda} \sin \mathrm{H} + \frac{b_y}{\lambda} \cos \mathrm{H}\right) \\
w &= \frac{b_w}{\lambda} = \left(\frac{\vec{b}}{\lambda} \cdot \hat{w}\right) = \cos \delta \left(-\frac{b_x}{\lambda} \sin \mathrm{H} + \frac{b_y}{\lambda} \cos \mathrm{H}\right)
\end{aligned} \tag{1.2}
$$

A given baseline at a given frequency is characterized by (u, v, w), which in our notation is identical to $(\frac{b_u}{\lambda}, \frac{b_v}{\lambda}, \frac{b_w}{\lambda})$.

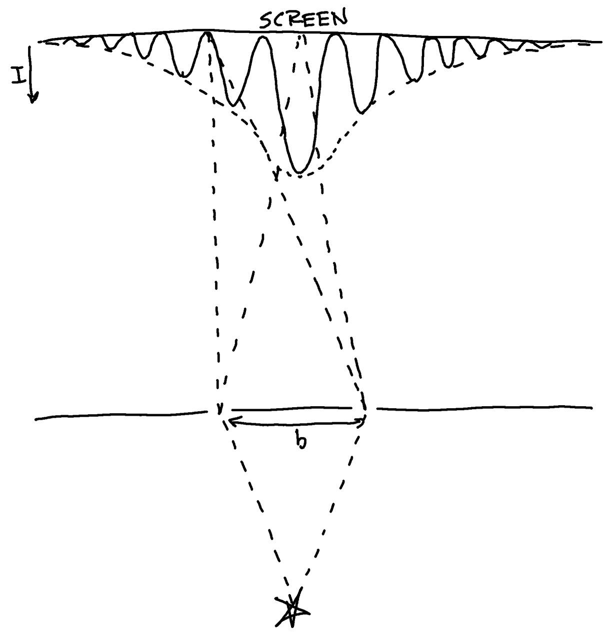


Figure 1.3: The double-slit experiment

1.3 The Primary Beam

The final subtlety of interferometry that we need to understand in order to establish the visibility equation is that of the primary beam. Put briefly, a pair of antennas does not measure the intensity of a single point on the sky as might have been implied by previous arguments, but rather samples the entire sky according to a sensitivity pattern called the "primary beam."

Consider Figure 1.3, illustrating the familiar double-slit experiment. The well-known result of viewing a coherent light source through two closely-spaced slits is the appearance of interference fringes. In Figure 1.3, we have traced the light-travel paths for one peak and one trough of the interference pattern. The path-length difference between the two paths for the peak is a whole number of wavelengths (the light arrives at the screen *in phase*), and the path-lengths that land on the trough differ by an odd number of half-wavelengths (the light arrives at the screen *out of phase*).

The spatial frequency of interference fringes is determined by the spacing, b, between the slits, whereas the width of the envelope describing the spatial attenuation of the sinusoid as its distance from the central bright fringe increases is determined by the size of the slits themselves. Reducing the distance between the slits, b, will reduce the spatial frequency of the interference pattern, whereas increasing their separation will make the fringes more spatially frequent. Increasing the individual slit sizes has the effect of making the function's envelope more sharply peaked in the center. The smaller the slits, the wider the interference pattern appears to spread.

A two-element interferometer can be understood as a double-slit experiment run in reverse. In place of the screen, we have the sky; the slits are the individual antennas; and what was the light-source is replaced by the correlator. With this new view, our interference fringes describe the sensitivity of our instrument to the sky. A peak represents the instrument's maximum sensitivity, whereas a trough represents a "blind spot" for this baseline.

Analogously to our description of what changing the size and separation of the slits does to the interference pattern in the double-slit experiment, our primary beam also changes as we change the sizes of antennas and the baseline lengths between them. The actual details become quite complex in real-world instruments (for one thing, our instrument is not two-dimensional), but the fundamental behavior is identical to that of the double-slit experiment. Increasing the size of the individual antennas will narrow the overall beam, and separating the antennas by a longer baseline will increase the frequency of its fringe pattern.

1.4 The Visibility Equation

Let's assume we are observing around a point of interest in the sky, which we will call the "phase center." As in Figure 1.1, we construct $\hat{s}$ such that it points directly at the phase center. In our (u, v, w) coordinate system, $\hat{s} = (0, 0, 1)$.

Now let us take one of the antennas of the baseline-pair and call it the "reference antenna," defining the phase delay of that antenna to be zero. When the (ideal, identical) antennas receive monochromatic, coherent radiation originating at the phase center, the voltages measured at the antennas can be expressed as the real parts of complex exponentials:

$$
\begin{aligned}
V_1 &= Ae^{i\omega t} \\
V_2 &= Ae^{i(\omega t + \varphi_0)}
\end{aligned}
\tag{1.3}
$$

The first voltage is then multiplied by the complex conjugate of the second to find the measured visibility:

$$
\begin{aligned}
V = V_1 \times \tilde{V_2} &= A^2 e^{i\omega t - i\omega t - i\varphi_0} \\
&= A^2 e^{-i\varphi_0}
\end{aligned}
\tag{1.4}
$$

Where A^2 is some real constant with units of [Volts]^2.

We can further break down φ_0, based on our understanding of the path length differences in Figure 1.1. As in Section 1.1, the total number of wavelengths difference between the path lengths will be $\vec{b} \cdot \hat{s}/\lambda$. To convert from a number of absolute wavelengths to an angle in radians, we simply multiply by 2π.

$$
\varphi_0 = 2\pi \frac{\vec{b} \cdot \hat{s}}{\lambda} \tag{1.5}
$$

Usually, we configure the instrument such that φ_0 at the phase center equals zero (i.e. waves arriving from the phase center are perfectly in phase when they reach the correlator, which multiplies the voltages together). For this discussion, though, we will remain in the generality that φ_0 may not equal zero.

Now, consider the case where the antennas are receiving a signal originating from a point *offset* from the phase center. We can define a new vector, $\hat{s}'$, which has length 1, and points in the direction of our source. Let the components of $\hat{s}' = (l, m, n)$ in the (u, v, w) coordinate system, with the condition that the vector has unit length, $1 = l^2 + m^2 + n^2$. For convenience, let us also define $\vec{\sigma} = \hat{s}' - \hat{s} = (l, m, n - 1)$, which we can visualize as pointing from the tip of $\hat{s}$ to the tip of $\hat{s}'$.

Much like above, we expect the measured visibility to have to do with the absolute phase difference between the radio light arriving at both antennas. As before,

$$\begin{aligned} V = V_1 \times \tilde{V_2} &= A^2 e^{-i\varphi} \\ &= A^2 \exp\left[-i2\pi\frac{\vec{b}\cdot\hat{s}'}{\lambda}\right] \end{aligned} \tag{1.6}$$

An absolute phase by itself doesn't mean very much, so we compare it to the phase at the phase center.

$$\begin{aligned} V = V_1 \times \tilde{V_2} &= A^2 e^{-i\Delta\varphi} \\ &= A^2 \exp\left[-i\left(\frac{2\pi\vec{b}\cdot\hat{s}'}{\lambda} - \frac{2\pi\vec{b}\cdot\hat{s}}{\lambda}\right)\right] \\ &= A^2 \exp\left[-i2\pi\frac{\vec{b}\cdot(\hat{s}' - \hat{s})}{\lambda}\right] \\ &= A^2 \exp\left[-i2\pi\frac{\vec{b}\cdot\vec{\sigma}}{\lambda}\right] \end{aligned} \tag{1.7}$$

Earlier, we described A^2 as a constant with units of [Volts]^2. Looking more closely, we know it is going to be related to both the intensity of the sky at that point, $I_\nu(l,m)$, and the instrument's spatial response to that intensity, the primary beam, $A(l,m)$. We can rewrite the visibility, this time considering the entire sky, rather than the contribution of a single point:

$$V(u,v,w) = \int_\Omega A(l,m) I_\nu(l,m) \exp\left[-i2\pi \frac{\vec{b}\cdot\vec{\sigma}}{\lambda}\right] d\Omega \tag{1.8}$$

Where $d\Omega$ is a differential solid angle of the sky, measured in steradians. Converting $d\Omega$ to $dldm$, we must multiply by the Jacobian:

$$d\Omega = \frac{dldm}{n} = \frac{dldm}{\sqrt{1-l^2-m^2}} \tag{1.9}$$

using the fact that $1 = l^2 + m^2 + n^2$.

Rewriting $\vec{b}\cdot\vec{\sigma}$ as $(u,v,w)\cdot(l,m,n-1)$, we have:

$$\vec{b}\cdot\vec{\sigma} = (ul + vm + w(n-1)) = (ul + vm + w(\sqrt{1-l^2-m^2} - 1)) \tag{1.10}$$

This brings us to our full Visibility Equation,

$$\boxed{\begin{aligned} V(u,v,w) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} & A(l,m) I_\nu(l,m) \times \\ & \exp\left[-i2\pi(ul + vm + w(\sqrt{1-l^2-m^2} - 1))\right] \frac{dldm}{\sqrt{1-l^2-m^2}} \end{aligned}} \tag{1.11}$$

Even though l and m are bounded by a maximum value, we can let the limits of the integral run from negative to positive infinity, and the primary beam function $A(l,m)$ can provide finite cutoffs for l and m by equalling zero outside of these cutoffs.

This full version of the Visibility Equation is what is used in all of the software built to measure the 21 cm EoR signal, with the w component of the dot-product included. However, if $l^2 + m^2 << 1$, or if $A(l, m)$ is strongly directional, being peaked at the phase center and small otherwise, we can make the so-called "flat sky" approximation, by disregarding the w component of the dot product entirely. This is equivalent to assuming that there is no information in the w-direction interference pattern, which is the same as assuming that the sky is a flat plane at infinity, rather than a hemisphere. In this approximation, Equation 1.11 simplifies to:

$$V(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A(l, m) I_\nu(l, m) \exp\left[-i2\pi(ul + vm)\right] dldm \tag{1.12}$$

Which is exactly the two-dimensional Fourier coefficient corresponding to the mode with indices (u, v). Thus, we can think of a baseline (u, v) as directly measuring the corresponding two-dimensional Fourier mode of the sky's intensity (convolved with the primary beam).

This insight grants us some intuition into practical design considerations for a radio interferometer. Because each individual baseline only measures one discrete Fourier mode at a given frequency and time, we want to have as many baselines of as many lengths and directions as possible, to increase our "uv-coverage" and make our picture of the sky more complete. Modern radio interferometers can have on the order of hundreds of antennas. If we denote the total number of antennas by N, then the total number of baselines will be $N(N-1)/2$.

In general, the goal of increasing uv-coverage encourages the layout of an interferometric array to be as randomized as possible. However, there are reasons why an array, such as the MWA, might have certain baseline vectors repeat between different pairs of antennas; these so-called *redundant* baselines increase the instrument's sensitivity to that Fourier mode by increasing its total collecting area, and also allow for a form of calibration which uses the fact that the visibilities of identical baseline vectors should be identical. Redundant calibration is discussed in more detail in chapter 3.

CHAPTER 2

The Murchison Widefield Array

Traveler, there are no paths. Paths are made by walking.

Indigenous Australian Proverb

2.1 Introduction

The Murchison Widefield Array is a low-frequency radio interferometer located in Western Australia. The array was commissioned and built partially as a precursor to the planned Square Kilometer Array (SKA). Despite being a proof-of-concept, the array has produced essential data for several important scientific projects in the ten years it has been operational, including work on the Epoch of Reionization (see Kolopanis et al. (2023)), alongside efforts including mapping the radio sky (see Wayth et al. (2015)), investigating pulsars (see Bhat et al. (2023)), and studying the sun's heliosphere (see Sharma et al. (2018)).

Figure 2.1: A single MWA tile. Photo credit ICRAR/Curtin.

2.2 Design

The MWA was upgraded to Phase II in the 2016-2017 season, as described in Wayth et al. (2018). Each MWA tile comprises 16 crossed dipoles arranged in a 4×4 square, on top of a ground-plate; a single MWA tile is visualized in Figure 2.1. The data used in this analysis were taken with the array in its "Phase II compact" configuration, which includes 128 of the array's antennas (alternatively called "tiles" in the context of the MWA) in a more closely-spaced spatial configuration than the alternative "Phase II extended" configuration.

In the compact configuration, the tiles are arranged as shown in Figure 2.2. As is apparent from that plot, some of the antennas are situated in two hexagons, which repeat certain baselines. In addition to increasing the sensitivity for those repeated (or "redundant") baselines, which should all receive identical signals, these two "hexes" allow for a type of calibration called "redundant calibration," which will be explored in Chapter 3.

Each dual-polarization "bowtie" antenna in a 4×4 tile has an onboard low-noise amplifier.

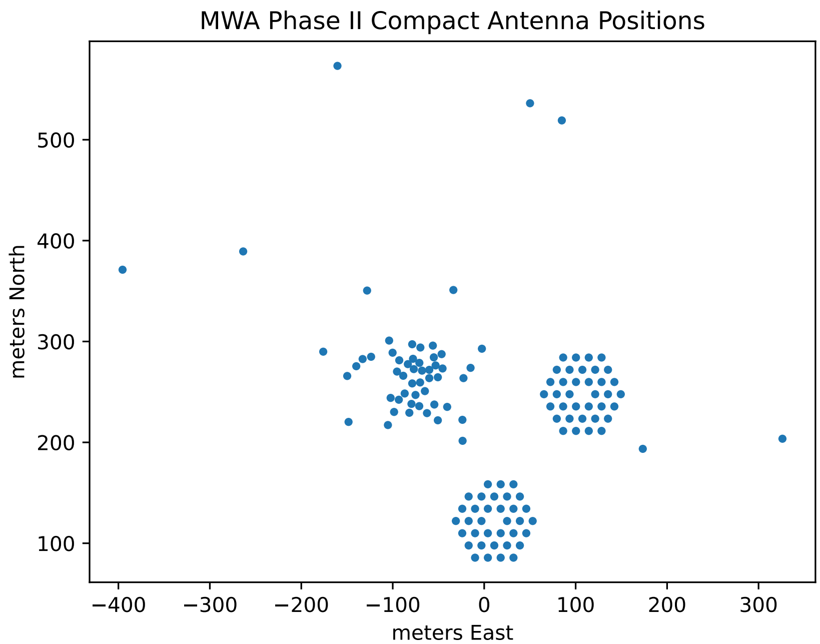


Figure 2.2: Tile Locations in the MWA Phase II compact configuration.

The amplified signal from each antenna is sent to the tile's beamformer, where the 16 signals of each polarization are added together. Each coaxial cable linking a dipole to the beamformer is of the same length, so that signals which arrive simultaneously at all antennas in the tile arrive in phase at the beamformer. The beamformer has the capability to change the relative delays between individual dipoles within the tile, by sending signal through physical delay loops on a printed board. These delays are used to phase the antenna to different angles by introducing an artificial time-delay to the signal depending on the positions of the individual antenna elements, with zero delay phasing to the zenith.

Following the pointing numbering scheme introduced in Beardsley et al. (2016), the tile pointings used in this work are numbered according to the schematic diagram in Figure 2.3. The data used in this work were taken with pointings between −3 and +3, inclusive.

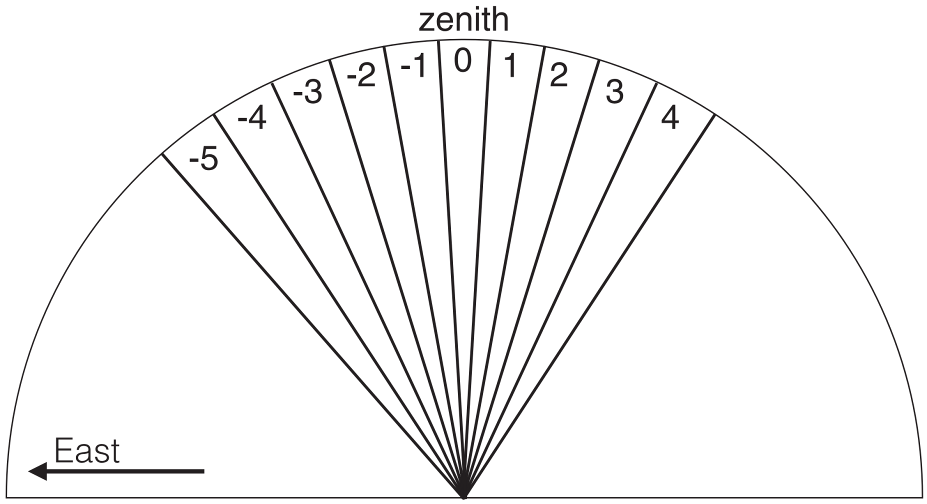


Figure 2.3: A schematic depiction of the MWA pointing-scheme used in this work, borrowed from Beardsley et al. (2016). The 0th pointing corresponds to the zenith, positive pointing numbers angle to the West of zenith, and negative pointing numbers angle to the East.

2.3 Data

The data used in this work were downloaded from the MWA All-Sky Virtual Observatory (ASVO)[1] service in MWA correlator FITS format. The data lie in the band between 167.055 MHz and 197.735 MHz, and were collected by the MWA between the dates of October 15th and December 15th of 2016. This period coincides with the beginning of MWA Phase II operations, and these data were taken with the array in its 128-antenna "compact" configuration, which includes 71 antennas in two redundant, compact hexagons of 36 and 35 antennas (both hexagons were intended to comprise 36 antennas, but one antenna location was inaccessible on site (Li et al., 2018)).

The dataset used in this work consists entirely of so-called "drift-and-shift" observations. A "drift-and-shift" observation is one in which the instrument's beam is pointed toward a particular source, which, as time passes, will move through the primary beam due to the Earth's rotation. Once the object has left the center of the beam, the instrument's pointing

[1] `https://asvo.mwatelescope.org/`

is incremented, and the source is allowed to drift through the primary beam once again. The areas of the sky tracked by the drift-and-shift data in our dataset are called EoR0 and EoR1, after the Epoch of Reionization. EoR0 is located at R.A. = 0.0°, Dec. = -27.0°, and EoR1 is located at R.A. = 60.0°, Dec = -30.0°. These two fields have been chosen for their fitness for attempts to observe the Epoch of Reionization. EoR0 is a mostly-empty field in terms of nearby radio sources. EoR1, on the other hand, includes Fornax A, a bright radio galaxy.

The EoR0 observations in this set are the same data analyzed in Li et al. (2019), and (apart from three nights of data analyzed in Zhang et al. (2020)) the EoR1 data are heretofore unanalyzed. The dataset comprises 51.8 hours of observations, spanning 1665 approximately two-minute observation files.

The MWA correlator used during Phase II of operations introduced a coarse-band structure to the data, which is discussed in detail in Section 4.1.4. Briefly, due to the design of the correlator, there are certain frequency bins (corresponding to the edges and centers of 24 coarse bands) which always contain corrupted data. Throughout this work, we will mask out corrupted data in these bins, and will always note when an algorithm has been modified to work around these masked sections of data.

2.4 Data Preprocessing

Before they are run through redundant calibration software, the data are downsampled in time using the python package pyuvdata[2] version 2.4.0 (Hazelton et al., 2017). The files are time-averaged from a time cadence of one visibility every 2 seconds to one visibility every 18 seconds; this is to reduce noise in the observations, which at a 2-second cadence prove too noisy to allow the redundant calibration algorithm to arrive at reasonable gain solutions. The

[2]Available at `https://github.com/RadioAstronomySoftwareGroup/pyuvdata`

rationale for this time-averaging is discussed further and example images are provided in A.1.

Pyuvdata is likewise used to select down to only the 71 antennas in the redundant section of the array. This is because the software package used in this work for running redundant calibration, `hera_cal`, was developed for use with the HERA telescope, which is fully redundantly laid out (Dillon and Parsons, 2016), as opposed to the MWA which has some redundant antennas alongside other, pseudorandomly positioned antennas. The software is therefore likely to calculate spuriously "redundant" baselines when all of the pseudorandomly positioned antennas are included.

Once the data have been pre-processed, they are run through redundant calibration software called `hera_cal`, which applies the algorithms detailed in Chapter 3. Although a fairly off-the-shelf version of `hera_cal` is run on the data, we are not performing the calibration to produce a good calibration for the array; rather, we are looking for a statistical measure of how similarly nominally "redundant" baselines actually behave in the instrument. This measure is called χ^2, and it turns out to be sensitive to radio-frequency interference. We were able to use χ^2 to detect RFI which was not picked up by other techniques, namely SSINS and AOFlagger, both of which are software in common use for RFI detection and excision in MWA data.

CHAPTER 3

Redundant Calibration

Not that I have anything much against redundancy.

But I said that already.

Larry Wall

creator of the Perl Programming Language

3.1 Introduction

As alluded to earlier, redundant calibration takes advantage of the fact that baselines which have the same vector separating their component antennas should in principle measure the same visibility from the sky. We call these repeated baselines "redundant." We assume that a measured visibility, v_{ij}, for a given time t, frequency ν, and polarization p, between antennas i and j, can be expressed as the product of two antenna-dependent complex gains, g_i and g_j and a "true" sky visibility y_{ij}, with an added noise term n_{ij}. The gains, g_i and g_j, encode complex physical conditions within the antenna, simplifying them into an added amplitude and phase which is imparted to the signal. The "true" sky visibility is the theoretical visibility which would be measured by a perfect instrument; i.e., the complex number we are trying to measure as accurately as possible with our manifestly imperfect instrument.

$$v_{ij}(t,\nu,p) \approx g_i(t,\nu,p)g_j^*(t,\nu,p)y_{ij}(t,\nu,p) + n_{ij} \tag{3.1}$$

We will drop the explicit time, frequency, and polarization dependencies in much of the following discussion, but it should be emphasized that redundant calibration provides an independent solution for the gains and visibilities at every unique combination of time, frequency, and polarization. This also means that we will extract a value for χ^2 at each time, frequency, and polarization.

The goal of calibration is to solve for the gains g_i. We go about this by minimizing χ^2, defined as

$$\chi^2 = \sum_{i<j} \frac{\left|v_{ij} - g_i g_j^* y_{ij}\right|^2}{\sigma_{ij}^2} \tag{3.2}$$

Where y_{ij} has subtly changed in meaning to be a "consensus" visibility calculated for the baseline group (one of several free parameters to be solved for in our system of linear equations), and σ_{ij}^2 is an estimation of the noise variance of that baseline.

Minimizing these χ^2 values is our objective as our calibration algorithms adjust parameters in a large system of linear equations, adjusting gains until they produce the best possible agreement between redundant baselines. The main difference between calibration algorithms is the method by which the system is linearized. In the redundant subsection of the MWA, there are 71 antennas, and if we only consider redundant baseline groups containing two or more baselines, that leaves 181 unique baseline types; 71 complex gains and 181 complex visibilities make for 252 complex free parameters to solve for, but the number of equations is much larger at $(71 \times 70)/2 = 2485$ (the total number of baselines in the redundant subsection of the array), making this system overdetermined.

χ^2 and its sensitivity to radio-frequency interference will be the focus of the remainder of this work. In order to obtain χ^2 for analysis, we must run redundant calibration, as described in sections 3.2, 3.3, and 3.4. However, it is vital to note that the results of this redundant calibration process — the gains and visibilities generated by the algorithm — will not be used in our analysis. Only the χ^2 values are of interest in finding and flagging RFI.

There are three different algorithms used to solve this system of equations in `hera_cal`. The first is aptly named `firstcal`, the second is known as `logcal`, and the third is called `omnical`.

3.2 firstcal

In order for our later iterative algorithms to converge, we need a reasonable starting guess for our gains; we obtain this guess from `firstcal`. The goal of `firstcal` — first introduced in Li et al. (2018) — is not accuracy, but reasonableness. Providing a reasonable first guess at gains is an important first step to prevent phase-wrapping and convergence errors with `logcal` and `omnical`.

`firstcal` begins with the assumption that a gain g_j at a given time can be broken down in terms of a few components:

$$g_j(\nu) = A_j(\nu)e^{i\phi_j(\nu)+2\pi i\nu\tau_j+i\theta_j} \tag{3.3}$$

With A, ϕ, τ, and θ as real parameters. A is an overall amplitude, ϕ represents second-order and higher functional terms describing how the phase changes with frequency, τ represents a linear time-dependence relating frequency to phase angle, and θ is a factor representing constant offsets in phase. The goal of `firstcal` will be to solve for the set of τ_i and the set of θ_i, which represent delays introduced by differences in cable length and an overall phase shift that might be due to improperly installed cables (which introduce a 180° phase shift), respectively. The issue of improperly installed cables is specific to the HERA hardware, and so will be disregarded in the following summary of the algorithm.

If we consider a redundant baseline group that contains the antenna pairs ij and kl, we can calculate the normalized product of their visibilities,

$$\begin{aligned}\frac{v_{ij}v_{kl}^*}{|v_{ij}||v_{kl}|} &\approx \frac{g_i g_j^* y_{ij} g_k^* g_l y_{ij}^*}{|g_i||g_j||g_k||g_l||y_{ij}|^2} \\ &\approx e^{(2\pi i\nu(\tau_i - \tau_j - \tau_k + \tau_l))}\end{aligned} \tag{3.4}$$

Here, we've made an approximation that all frequency dependence in the gains can be expressed as a delay term τ (i.e., we have ignored the contribution of $\phi(\nu)$). For the purposes of obtaining a rough estimation of our final parameters, this is a reasonable assumption, equivalent to ignoring second-order and higher terms.

The left-hand side of equation (3.4) is a measurable quantity, which can be expressed as the phase angle between two visibilities v_{ij} and v_{kl}. If v_{ij} is expressed as

$$v_{ij} = A_{ij} e^{i\varphi_{ij}} \tag{3.5}$$

then the left hand side of equation 3.4 can be equivalently written as

$$\frac{v_{ij}v_{kl}^*}{|v_{ij}||v_{kl}|} = e^{i(\varphi_{ij} - \varphi_{kl})} \tag{3.6}$$

Substituting this into equation (3.4), we get

$$e^{i(\varphi_{ij}-\varphi_{kl})} \approx e^{(2\pi i\nu(\tau_i-\tau_j-\tau_k+\tau_l))} \tag{3.7}$$

Taking the logarithm, this simplifies to

$$\varphi_{ij} - \varphi_{kl} \approx 2\pi\nu(\tau_i - \tau_j - \tau_k + \tau_l) \tag{3.8}$$

There are several equations of this type for each redundant baseline group, and so this system of equations may be solved for the τ terms, which are then applied before running the next steps of the process.

3.3 logcal

Beginning with equation (3.1), we can take the natural logarithm of both sides to linearize:

$$\begin{aligned} \mathbb{R}\mathrm{e}[\ln v_{ij}] &= \eta_i + \eta_j + \mathbb{R}\mathrm{e}[\ln y_{ij}] \\ \mathbb{I}\mathrm{m}[\ln v_{ij}] &= \varphi_i - \varphi_j + \mathbb{I}\mathrm{m}[\ln y_{ij}] \end{aligned} \tag{3.9}$$

where we have factored g_j as follows: $g_j = e^{\eta_j + i\varphi_j}$ and ignored the noise term n_{ij}. Considering the vector of all $\ln v_{ij}$ in the instrument, we can formulate this as a system of linear equations which may be solved for the gains and the "true" visibilities y_{ij}.

We can construct this vector of logarithms of visibilities, $\ln v_{ij}$ and call it $\mathbf{d}$, as well as a vector containing all of the $\ln y_{ij}$, η_i, and φ_i, called $\mathbf{x}$. Using these, we can encode equation

(3.9) as

$$\begin{aligned} \mathbb{R}\mathrm{e}[\mathbf{d}] &= \mathbf{A}\ \mathbb{R}\mathrm{e}[\mathbf{x}] \\ \mathbb{I}\mathrm{m}[\mathbf{d}] &= \mathbf{B}\ \mathbb{I}\mathrm{m}[\mathbf{x}] \end{aligned} \tag{3.10}$$

Where $\mathbf{A}$ and $\mathbf{B}$ are matrices with elements of only 0's, 1's, and in the case of $\mathbf{B}$, −1's.

This system of linear equations can be directly solved via conventional methods for the unknown vector $\mathbf{x}$, yielding estimations of both gains and visibilities for the array.

This approach unfortunately proves to produce biased results, as shown in Liu et al. (2010), so it can't be used for a final calibration algorithm. However, `firstcal` only produces estimates for gains, and not visibilites, but `omnical` needs gains and visibilities to begin solving. Therefore, `logcal`, which solves for visibilities as well as gains, is used as a transition between `firstcal` and `omnical`.

3.4 omnical

Another approach to linearizing equation (3.1) is to perform a first-order approximation of the terms. This approach was proposed by Liu et al. (2010), where it was called `lincal`, and improved in Zheng et al. (2014), in an algorithm called `omnical` (which was unfortunately not properly described in the literature until Dillon et al. (2020)).

Starting from equation (3.1) and disregarding the noise term n_{ij}, we can set $y_{ij} = y_{ij}^0 + \Delta y_{ij}$ and $g_i = g_i^0 + \Delta g_i$. For the sake of iteration, y_{ij}^0 and g_i^0 are taken to be the algorithm's most recent iteration's best solutions (beginning in the first step with `lincal`'s solutions), while the Δ terms are solved for as a correction in the current iteration, to be included in y_{ij}^0 and g_i^0 in the next iteration of the algorithm.

Rewriting equation (3.1) in terms of these linearized approximations, disregarding terms which are second-order in the Δs, we find

$$v_{ij} - g_i^0 g_j^{0*} y_{ij}^0 = g_j^{0*} y_{ij}^0 \Delta g_i + g_i^0 y_{ij}^0 \Delta g_j^* + g_i^0 g_j^{0*} \Delta y_{ij} \tag{3.11}$$

Both `lincal` and `omnical` solve this system of linear equations for the Δs, and then iterate. `omnical` differs from the original implementation of `lincal` in how it updates gains and visibilites; `omnical` updates each individual gain and visibility as though all other gains and visibilities were held static. As a result, `omnical` executes in $\mathcal{O}(N^2)$ time, where N is the number of antennas in the array, while the original implementation of `lincal` executes in $\mathcal{O}(N^3)$ time[1] (Dillon et al., 2020).

In each iteration of `omnical`, new Δ's are found by minimizing χ^2, as described in equation 3.2. In a sense, the minimization of χ^2 — representing a maximization of redundancy in redundant baseline groups — is the entire goal of `omnical`. As a result, χ^2 values for both the overall array and for individual baselines (relative to their baseline groups) are produced by the software when it runs. These χ^2 values are intended mainly as a proxy for the overall quality of the final calibration, higher values representing the presence of data which were inconsistently valued between nominally redundant baselines, even in the final calibration, signifying worse convergence and perhaps poor calibration solutions.

In the course of our investigations, it became obvious that these χ^2 values produced by `omnical` demonstrate sensitivity to some RFI, the details and utility of which will be explored

[1]Further complicating the history of these algorithms, the software package described in Zheng et al. (2014) is called `OMNICAL`, and contains three algorithms: `logcal`, `lincal`, and `omnical` (where we use lowercase letters to distinguish the algorithm `omnical` from the software package `OMNICAL`). `lincal` and `omnical` perform the same task with different solvers. Previous work applying redundant calibration to the MWA (Li et al. (2018), Li et al. (2019), and Zhang et al. (2020)) use the `OMNICAL` package to run `logcal` and `lincal` but not `omnical`. To be explicit, in this work we use the `hera_cal` software package (not `OMNICAL`) and the `omnical` algorithm (not `lincal`).

in Chapter 4.

CHAPTER 4

χ^2 from Redundant Calibration in RFI detection

And now here is my secret, a very simple secret:
It is only with the heart that one can see rightly;
what is essential is invisible to the eye.

Antoine de Saint-Exupéry
Le Petit Prince

4.1 RFI identification and flagging with χ^2

Once our data have been processed by `hera_cal`, we are left with the outputs of redundant calibration, including χ^2, which we will demonstrate presently to have a sensitivity to radio-frequency interference. Furthermore, we demonstrate that a fairly simple two-step flagging algorithm can effectively flag this interference.

4.1.1 χ^2 and Degrees of Freedom

As `hera_cal` implements the algorithms described in Chapter 3 to estimate gains g_i and visibilities y_{ij}, it does so with the goal of minimizing the metric χ^2, defined in equation (3.2). This is the overall χ^2 for the entire array (at each time, frequency, and polarization), found by summing over each unique baseline ij, comparing a baseline's measured visibility with the visibility calculated by multiplying the antenna gains g_i and g_j^* by the calculated "consensus" visibility y_{ij}. The difference of these quantities is squared, and then divided by an estimation of the noise variance σ_{ij}^2. If the measured visibility and calculated visibility only differ by a factor of noise, this quotient should be of order 1. When summed over the entire system, we would expect χ^2 of a well-calibrated array to be of order nDoF, or the number of degrees of freedom in the system, calculated by Dillon et al. (2020) to be

$$\text{nDoF} = N_{bl} - N_{ubl} - N_{\text{ants}} + 2.5 \tag{4.1}$$

Where N_{bl} is the total number of baselines, N_{ubl} is the number of unique baseline types, or alternatively the number of y_{ij} that need to be solved for, and N_{ants} is the total number of antennas, understood as the number of individual gains g_i to be solved for. The additive factor of 2.5 accounts for the degeneracies of redundant calibration in the array, which reduce the total number of parameters which can be solved for in our system of linear equations, increasing the total number of degrees of freedom.

In Dillon et al. (2020), this constant factor is calculated to be 2 in a classical redundant array. In such an array, these 2 complex numbers or 4 real parameters correspond to overall amplitude, overall phase, tip/tilt of the array in the N/S direction, and in the E/W direction. However, due to having two spatially separated redundant hexagons, the MWA introduces an additional real degenerate parameter — the phase separation between the two redundant hexagons — which increases the number of real degeneracies in the system from 4 to 5, and the number of complex degenerate parameters in turn from 2 to 2.5 (Dillon, 2024).

When considering values of χ^2/nDoF, we expect a number close to 1.0 for each time and frequency if redundant calibration has been successfully run and the resulting calibration is accurate.

χ^2 is intrinsically a measure of how redundant the array actually is. Said another way, χ^2 measures how well the calibrated baseline groups agree with each other on a measured visibility; in a perfectly calibrated, perfectly constructed ideal array, each redundant baseline in a group should measure identical visibilities. Higher χ^2 values for a given time or frequency indicate that nominally "redundant" baseline groups do not agree on their measured visibilities.

We have observed that χ^2/nDoF is elevated by radio-frequency interference (c.f. the lower plot in Figure A.1). There are two main ways this could occur: either the presence of RFI increases χ^2 by increasing non-redundancy (the numerator of χ^2), or the presence of RFI leads to an underestimation of noise variance σ_{ij}^2 (the denominator of χ^2).

In order to determine which of these mechanisms was responsible for heightened χ^2 in the presence of RFI, we separated the numerator and denominator of equation (3.2). Within `hera_cal`, the antenna autocorrelations, which measure the total power incident on each antenna element, are used as a proxy for σ^2. We do find that some powerful RFI events slightly elevate the total power seen in the autocorrelations, and hence elevate `hera_cal`'s estimation of σ^2. However, this has the effect of suppressing χ^2, opposite to the overall effect in χ^2 we report here. It follows that RFI elevates χ^2 by magnifying or exacerbating non-redundancy in the array, i.e., it increases the numerator of equation (3.2).

4.1.2 Modified z-score as a tool to flag outlying χ^2 data

The simplest approach to flagging RFI in χ^2 data would be to set an absolute threshold for χ^2/nDoF, and to flag data which exceed that threshold. However, both the EoR0 and EoR1 fields have an overall Local Sidereal Time (LST) dependence in the χ^2 values. This makes it difficult to define an absolute χ^2/nDoF cutoff for flagging data, since what may be an elevated χ^2 value for one observation can be the χ^2 value of perfectly uncontaminated data for another. Figure 4.1 illustrates this LST-dependence with the North-North polarizations of EoR1 observations. Y-axis locations represent the mean of χ^2/nDoF for each observation, and the error bars signify their standard deviations. Points are color-coded by integer modified Julian date, such that data taken during the same 24-hour period are of the same color. Different marker shapes represent different pointings of the MWA beam.

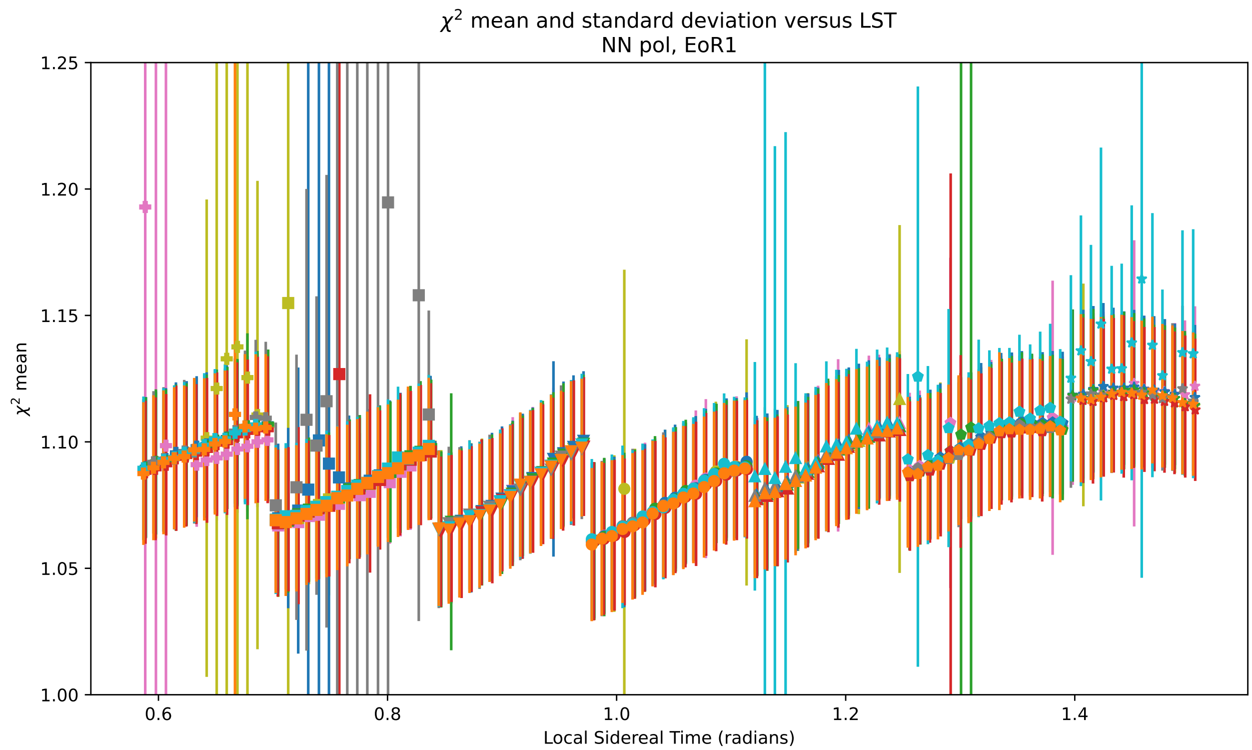
In the EoR1 data especially, there is a clear upward trend in each individual pointing's χ^2/nDoF means. This sub-pointing trend is likely due to bright celestial objects moving through the sidelobes of the individual tile beams, which are much less redundant between the tiles of the MWA than the beam's main lobe (Line et al., 2018; Chokshi et al., 2021); seeChoudhuri et al. (2021) for a similar effect seen in simulations of HERA.

Taking the modified z-scores of the χ^2/nDoF data greatly reduces their LST-dependence, as shown in Figure 4.2, which represents the same data as Figure 4.1. Modified z-score for the i^{th} element of a dataset x is calculated according to the following formula:

$$z_i = 0.6745 \times \frac{(x_i - \tilde{x})}{(\text{MAD})} \tag{4.2}$$

where $\tilde{x}$ represents the dataset's median, and MAD is the median absolute deviation of the

Figure 4.1: The mean value of χ^2/nDoF versus local sidereal time; each point represents one two-minute observation, with the standard deviation of χ^2/nDoF within the observation represented as an error bar. Each color represents data from a different night, while each shape represents data from a different pointing. The χ^2 means have a pronounced LST dependence. Within individual instrument pointings, there is also an upward trend. Because of the dependence on LST, a single χ^2/nDoF cutoff value applied across all observations in sub-optimal for flaging RFI.

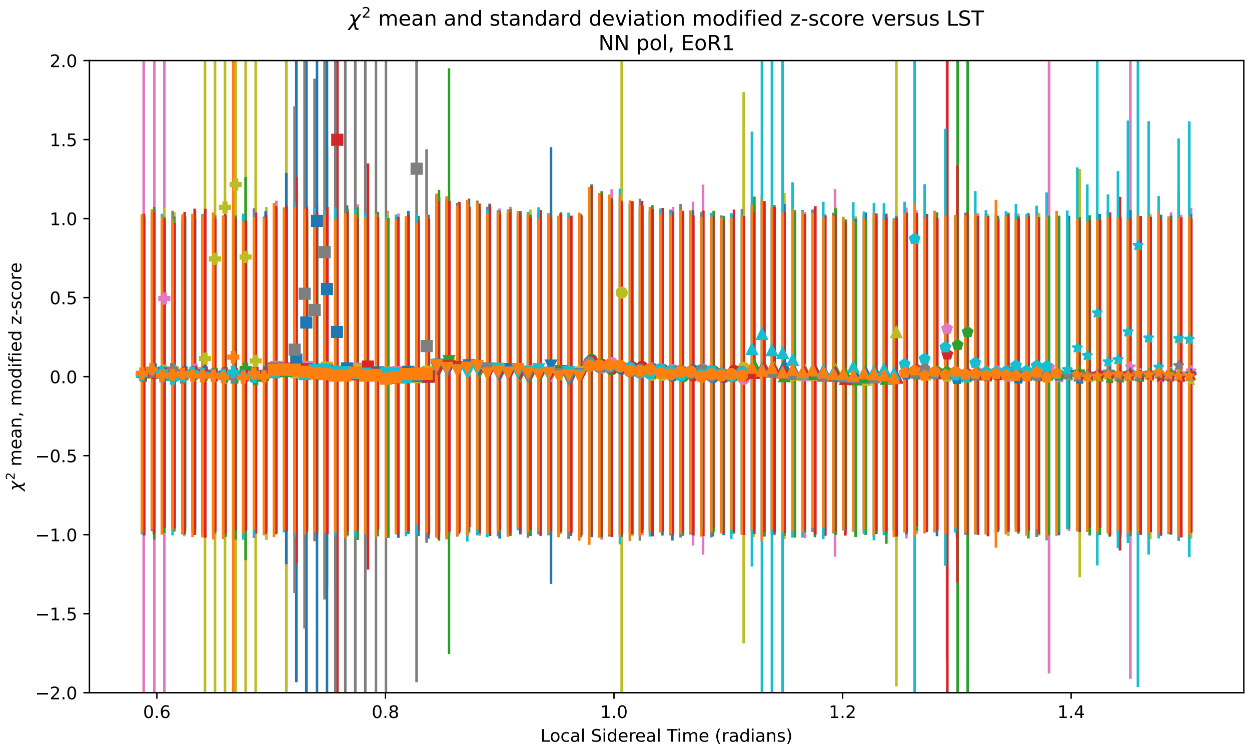


data from the median.

$$\mathrm{MAD} = \widetilde{|x_i - \tilde{x}|} \tag{4.3}$$

Here, the tilde (~) operator has been used to signify "take the median of the dataset." The factor of 0.6745 is included to scale modified z-scores to have the same step size as standard z-score, i.e., if the standard z-score and modified z-score are both taken of the same perfectly Gaussian dataset, each integer step in modified z-score is the same size as a step in standard z-score, and corresponds to 1 standard deviation of the dataset.

Figure 4.2: This plot represents the same data as Figure 4.1, except now modified z-scores have been taken of the χ^2/nDoF data. The mean of the modified z-scores of χ^2/nDoF is plotted against local sidereal time, with the standard deviation of the modified z-scores of the data represented as error bars. The LST dependence of χ^2 present in Figure 4.1 is effectively eliminated by using modified z-score. As with Figure 4.1, color denotes the day data were taken, and different point shapes represent different antenna pointings of the instrument.



As opposed to the standard z-score, modified z-score uses medians instead of means. As a result, the measure is less affected by outlier values, and more appropriate for datasets that have a mostly Gaussian distribution, aside from some high z-score outliers (in this case caused by RFI).

Modified z-score assumes that the data have a distribution that is close to Gaussian; this is a valid assumption for these data because, even though a χ^2 distribution is not Gaussian, due to the central limit theorem, as nDoF increases, so does the distribution's resemblance to a Gaussian. For the redundant sub-array of the MWA, as calculated with equation (4.1), nDoF $= 2228.5$, which is more than enough to warrant making this simplification.

Because it almost completely eliminates LST-dependence of χ^2, we base our RFI-flagging algorithm on the modified z-scores of χ^2/nDoF generated by redundant calibration.

As shown in Figure 4.2, most files have a mean z-score which is near 0.0 and a standard deviation of z-scores which is close to 1.0. There are also files with noticeably higher mean z-scores, and wider standard deviations, which are almost always caused by strong RFI events. However, rather than further analyze the distribution of χ^2 values within a single observation, we turn to the time, frequency, and polarization information of the χ^2's, which help us localize and characterize the source of the interference.

4.1.3 Examples of χ^2 and its Sensitivity to RFI

To demonstrate the sensitivity of χ^2 to RFI, we have included here a few χ^2 time-frequency waterfall plots and histograms of the corresponding data in Figure 4.3.

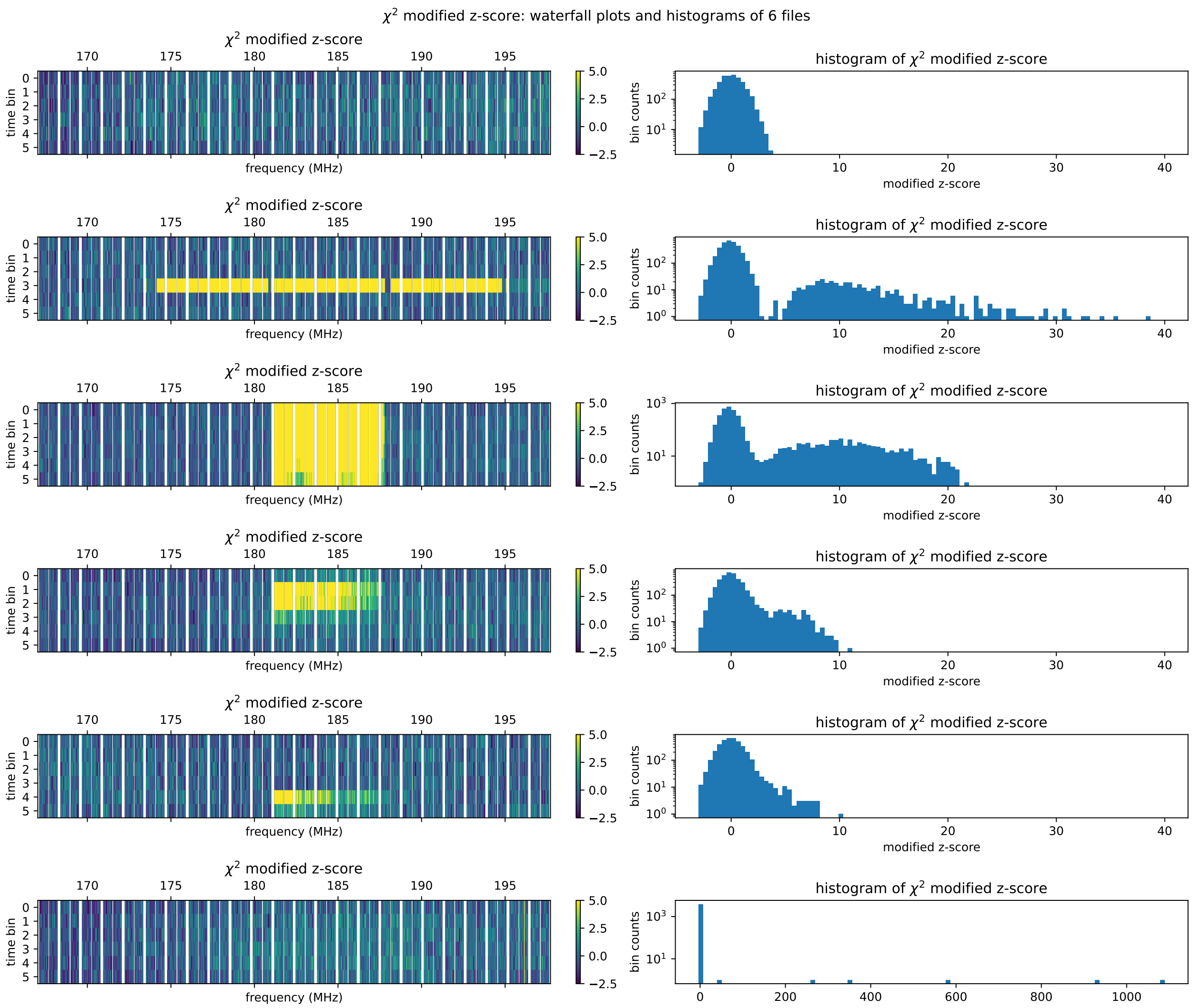
While the RFI-free file produces a distribution of χ^2 that is close to Gaussian, the presence of RFI leads to outliers. Included are four examples of DTV RFI of differing strengths. Common to all four of them is a bimodal distribution of χ^2, which generally remains below $z = 40$, but, even in the example of the weakest RFI, have bins above $z = 4$.

On the other hand, the RFI in the final waterfall plot of Figure 4.3 may be difficult to see, because it only spans one frequency-bin at 196.175 MHz. It is worth noting, however, that the χ^2 z-score values are exceptionally high for these six bins, as demonstrated in the histogram.

4.1.4 RFI Flagging

RFI appearing in the χ^2 data is flagged with a two-step process consisting of iterative modified z-score flagging followed by watershed flagging. The entirety of our flagging code,

Figure 4.3: Waterfalls and histograms of the modified z-scores of χ^2/nDoF of 6 two-minute observations. The waterfall plots, left, show the modified z-score of χ^2/nDoF as a function of time and frequency. Elevated values are indicative of RFI. To the right of each waterfall plot is the corresponding histogram, showing the distribution of the modified z-scores of the data. The first observation is RFI-free and shows a nearly Gaussian normal distribution, whereas the subsequent examples all contain RFI and have high-z-score outliers of varying degree. The second through fifth images all represent DTV events, while the sixth is a narrow-band event at 196.175 MHz.



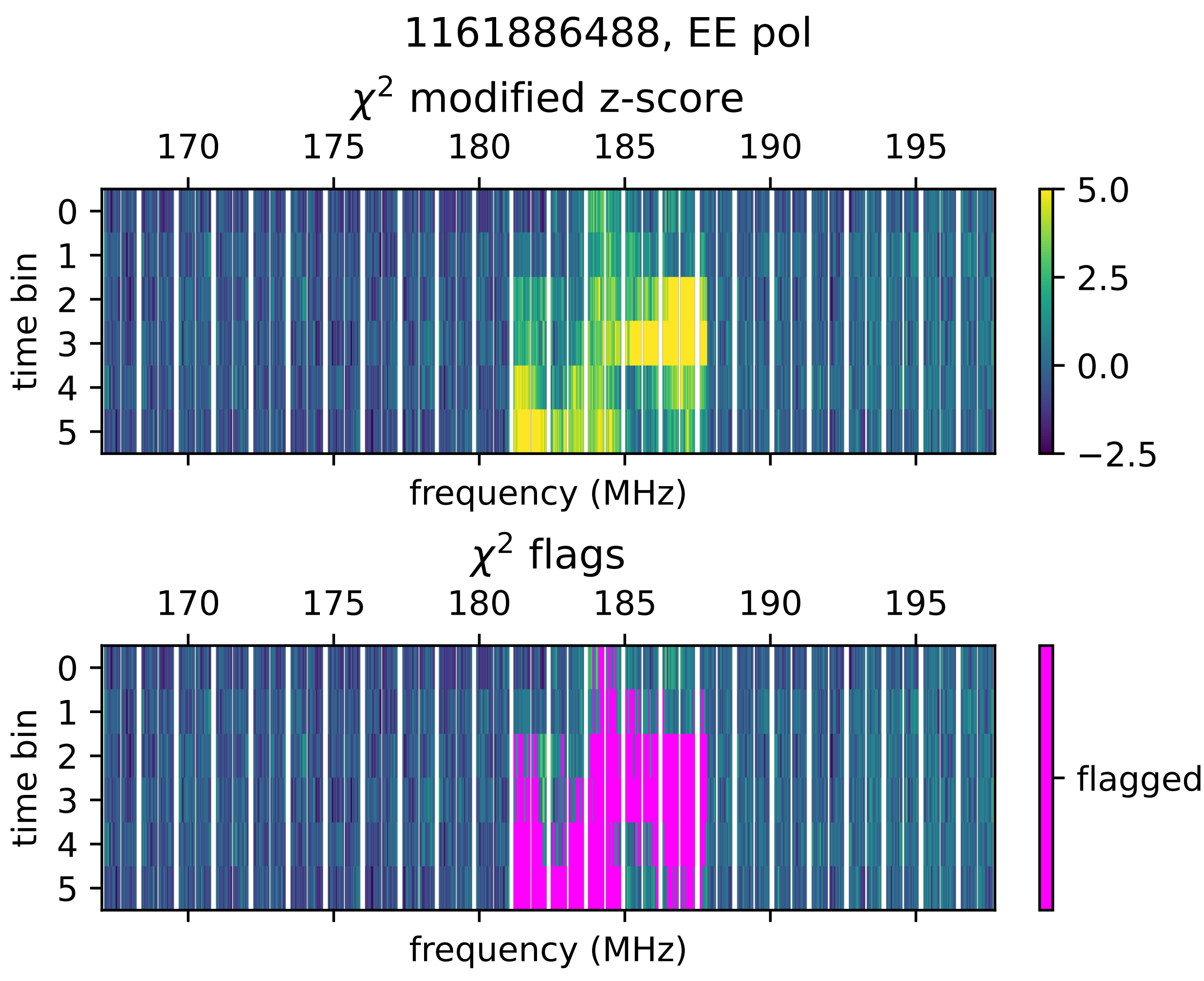
including algorithms for calculating modified z-score, iteratively flagging with modified z-score, and watershed flagging, is included in A.3.

First, the modified z-scores of the χ^2 values are calculated for the dataset. Each point with a modified z-score greater than 4.0 is first masked out of the dataset. After high-z-score datapoints have been flagged, z-scores are recalculated without the outlier points, and the process is repeated until it produces no update in flags from its most recent iteration. This part of the algorithm is implemented in our code, available in A.3, in the function `iteritively_flag()`.

Starting from this preliminary set of flags, we follow the watershed RFI flagging algorithm described in Appendix A of Kerrigan et al. (2019). Briefly, this flagging algorithm works on the assumption that RFI-contaminated bins tend to be contiguous, and flags every bin with a modified z-score greater than 2.0 which lies orthogonally adjacent to an already-flagged bin. This process is iteratively repeated until the flags remain unchanged from the last iteration, thereby "flooding" the lower-z-score data which are directly connected to higher outliers. This part of the algorithm is implemented in our code, available in A.3, in the function `watershed()`.

This algorithm has to be slightly modified in order to handle the coarse-band flagging which is necessary for Phase II MWA data. In general, when working with MWA Phase II data, within each 1.28 MHz, 32-channel coarse band, two channels on either edge of the band as well as the central channel are flagged and excluded from analysis due to data corruption arising from aliasing from the poly-phase filter bank (for the edges) and DC offsets (for the center) (Offringa et al., 2015). These flags are apparent in the masked values in the waterfall plots of Figure 4.3. As a result, the function for finding "adjacent" bins has to be modified to ignore the coarse band edges. When a bin being processed by the watershed algorithm is adjacent to a coarse band flag, the algorithm skips the flagged coarse feature and considers the next bin beyond it as adjacent to the flagged bins on the other side for "flooding." This part of the algorithm is implemented in our code, available in A.3, in the function `adjacent_cells()`.

Figure 4.4: An example χ^2 waterfall plot of an incompletely-flagged channel 7 DTV event. The top panel shows the calculated z-scores, while the bottom panel shows the derived flags in magenta. This event was flagged using the algorithm described in Section 4.1.4, but the channel 7 DTV band is only partially flagged for any given time.



This two-step flagging algorithm is far from perfect. Even after watershed flagging, there are datapoints which are clearly contiguously part of an RFI event which remain unflagged due to their low z-scores. One such event is shown in Figure 4.4. In this work, we focus on demonstrating the potential for redundant calibration χ^2 to detect RFI events. Future work is likely to improve on the exact algorithm for calculating flags from the χ^2 values themselves. Especially interesting is the prospect of using a machine-learning approach to flagging RFI events, based on the heuristic that they are easy for the human eye to identify, while still being difficult to algorithmically define.

4.2 Other RFI-Flagging Algorithms: AOFlagger and SSINS

In order to compare χ^2 flagging with existing RFI flagging algorithms, we must obtain these algorithms' flags for the same dataset.

For the purposes of this work, we have compared the performance of χ^2 flagging with AOFlagger as part of the MWA `cotter` pipeline (Offringa et al., 2015) and SSINS (Wilensky et al., 2019), run locally.

4.2.1 AOFlagger

AOFlagger operates in three broad steps. First, a high-pass filter is applied to the visibility amplitudes. The filtered visibility amplitudes are then run through a straight-line detecting algorithm called SumThreshold, introduced in Offringa et al. (2010), based on the observation that RFI often has sharp edges in the frequency-time domain. Finally, the scale-invariant rank (SIR) operator, introduced in Offringa et al. (2012), is applied to the SumThreshold-flagged visibilities, which acts to propagate those flags into adjacent contaminated bins.

AOFlagger flags are generated by `cotter` (the processing pipeline in use by the MWA during Phase II operations), and included when downloading MWA data from the data download service ASVO. The AOFlagger algorithm generates flags by baseline, so in order to visualize the flagging patterns for an entire individual observation, we average over the baseline axis and view the average occupation fraction for each frequency/time bin for a given polarization.

4.2.2 SSINS

SSINS is short for Sky-Subtracted Incoherent Noise Spectra, and is an RFI-flagging algorithm first introduced in Wilensky et al. (2019)[1]. SSINS begins with "sky-subtraction," whereby every visibility is subtracted from the next visibility adjacent in time, removing features that vary much more slowly than the two-second cadence of most MWA data, including the sky. Once sky-subtraction has been applied to every baseline's visibilities, the data should consist of only noise and RFI. The sensitivity to RFI is increased by forming an "incoherent noise spectrum," or INS, of the sky-subtracted visibilities. This is done by averaging the visibility amplitudes (discarding phase) over all baselines in the array. One INS per polarization is produced.

SSINS uses a frequency-matched flagging algorithm to find features in the sky-subtracted incoherent noise spectrum. This algorithm is aware of DTV-bands and will flag out the entire channel band if enough DTV RFI is detected in that channel. The flagging algorithm is also sensitive to narrow-band RFI and broad-band streaks.

SSINS version 1.4.7 was run locally on the downloaded data. In order to reduce the number of spurious flags, the incoherent noise spectrum (INS) object was calculated with a 2^{nd} order polynomial fit for each frequency channel during mean subtraction.

[1] `https://github.com/mwilensky768/SSINS`

4.3 Results

4.3.1 Average flagging occupation fractions

When the χ^2 flags are averaged across all of the files in the dataset, as visualized in Figure 4.5, it becomes clear that RFI falling in the band allocated to Australia's digital television channel 7 (181-188 MHz) comprises the bulk of the RFI which is flagged. Looking closely, we see that DTV channels 6 (174-181 MHz) and 9 (195-202 MHz) are also visible as slight elevations in flagging fraction.

In addition to the strong peak at channel 7 frequencies, there are some noticeable narrow-band peaks. Most of these narrow-band peaks correspond to the edges of the MWA correlator's coarse bands, which are widely known to contaminate data, as described in Section 4.1.4. Coarse band flags are represented by the shaded zones in Figure 4.5. Most of the observed narrow-band peaks in flagging fraction are adjacent to these flagged channels, which perhaps indicates that the data corruption extends beyond the commonly-flagged channels, although the flagging fraction is only $1 - 2\%$, so it follows that their χ^2 values are not consistently elevated. The cause of the elevated χ^2 numbers that occasionally appear around the coarse band edges remains a topic for further investigation.

Of the narrow-band flagging peaks, two prominent examples are *not* adjacent to coarse band edges. These represent real narrow-band RFI, each contained within a single 40 kHz frequency-bin at central frequencies of 196.175 MHz and 194.455 MHz for the higher and lower peaks, respectively.

The last obvious feature of Figure 4.5 is a small increase in flagging at low frequencies, which is due to the spurious flagging of low-frequency bins whose χ^2 values are elevated — not by RFI, but by the presence of the Galactic plane in the far sidelobes of the primary beam. Affected observations, like the one visualized in Figure 4.6, are uniformly from gridpoint

Figure 4.5: The fraction of data flagged using χ^2 across all observations as a function of frequency. The most commonly flagged RFI (in approximately 5 to 6% of the data) is channel 7 DTV, as indicated by the large feature between 181 and 188 MHz. Also visible are slight amounts of flagging in channels 6 (174-181 MHz) and 9 (195-202 MHz). Narrow band RFI is apparent at 196.175 MHz and 194.455 MHz. Most other narrow-band peaks correspond to MWA coarse band edges, indicated by the shaded regions and are unlikely to be actual RFI. The slight increase in flagging at the lowest frequencies is due to spurious flagging of elevated χ^2 stemming from the presence of the Galactic plane in the sidelobes of the primary beam (see Figure 4.6).

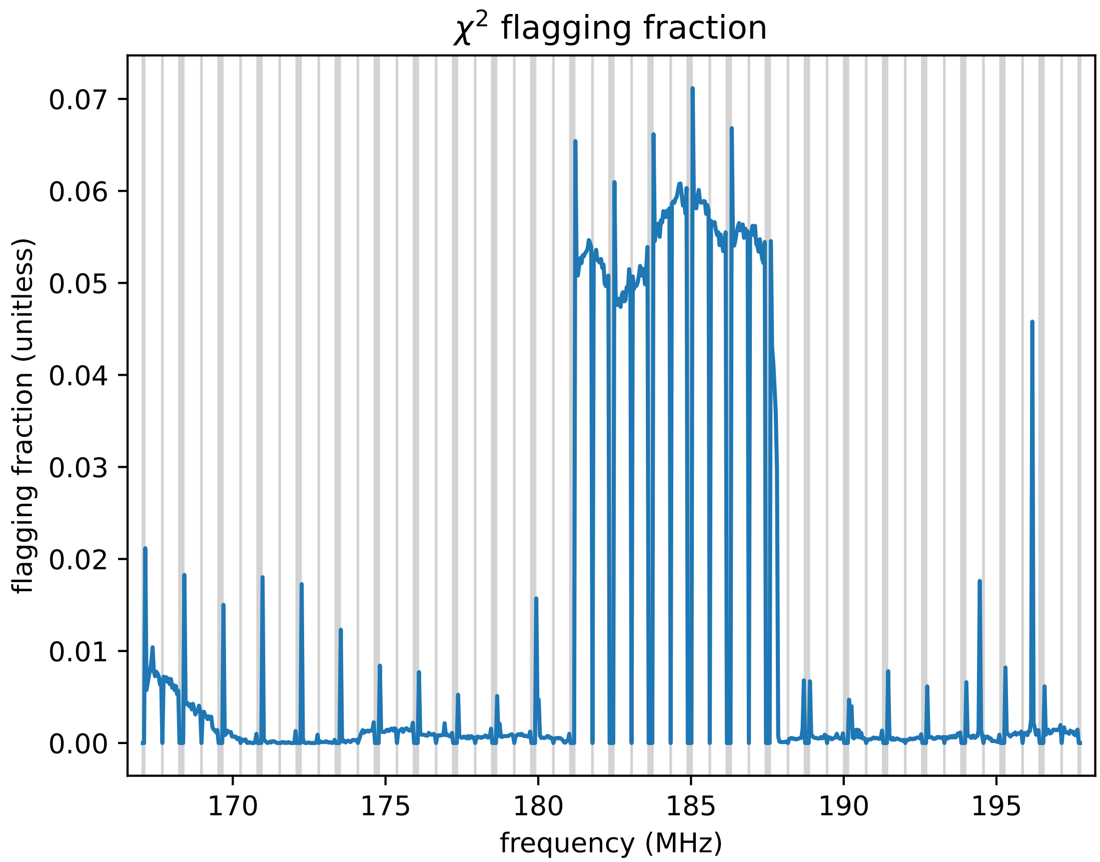
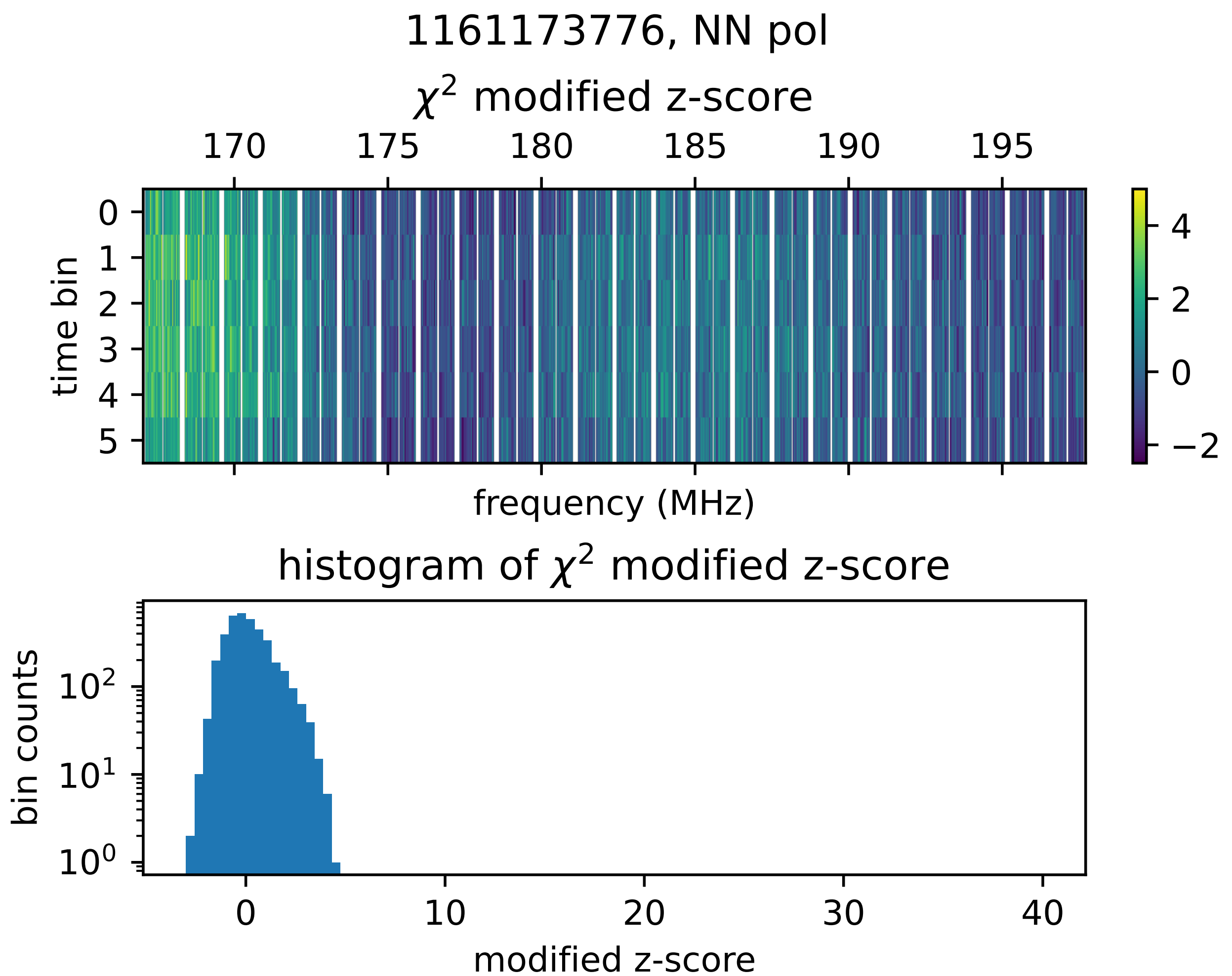


Figure 4.6: An example of an observation with elevated low-frequency NN χ^2 values due to the presence of the Galactic plane in the sidelobes of the MWA primary beam. In the corresponding histogram, we can see that the distribution deviates from a Gaussian normal. Observations like this (all corresponding to the -3 pointing with the instrument phased to EoR0) cause the uptick in flagging at very low frequencies in Figure 4.5.



number −3 of the grid pointing scheme used in Beardsley et al. (2016), with az = 90°, el = 69.1655° — one of the more extreme pointings of the instrument used to observe the EoR0 field which appear in this dataset. The affected data were taken when the Galactic anti-center was still above the horizon, where it could be picked up by sensitivity in the sidelobes of the beam. Although using modified z-score for flagging decreases the number of spurious flags in this pointing, it does not reduce them to zero. This results in the uptick in χ^2 flagging at the

	flagged by none	flagged by χ^2 only	flagged by SSINS only	flagged by AOFlagger only	flagged by χ^2 and SSINS only	flagged by χ^2 and AOFlagger only	flagged by SSINS and AOFlagger only	flagged by all
number of observations	1209	169	143	0	30	13	14	87
% of total	72.61%	10.15%	8.59%	0.00%	1.80%	0.78%	0.84%	5.23%
% of flagged	n/a	37.06%	31.36%	0.00%	6.58%	2.85%	3.07%	19.08%

Table 4.1: Number of files flagged in the DTV channel 7 band for χ^2, SSINS, and AOFlagger.

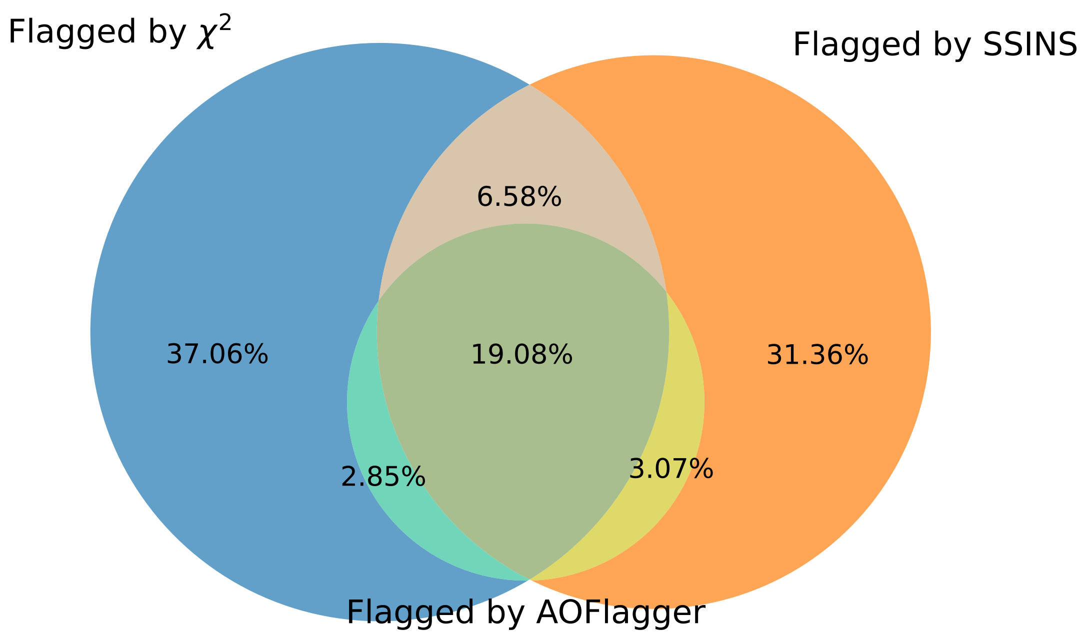
lowest frequencies observable by the MWA. Future work looking to apply RFI flags generated from χ^2 values to MWA data may need to handle the analysis of this pointing separately.

4.3.2 Comparing χ^2 flagging with SSINS and AOFlagger

The most common type of RFI detected in this dataset was channel 7 DTV, which is broadcast between 181 MHz and 188 MHz. This section focuses on flagging RFI when only this band is considered. Other types of RFI are discussed in Section 4.4.

In Figure 4.8, a filled dot indicates that channel 7 DTV RFI was detected in a particular observation. In order to qualify as "detected" by χ^2 or AOFlagger, a file must clear a threshold value of 5% of channel 7 bins being flagged by the algorithm in question, as well as verifying that the DTV channel 7 band is more densely flagged than the file as a whole (which distinguishes channel-7-only RFI from broad-band events). To qualify as "detected" by SSINS, a file must be tagged as containing channel 7 RFI by SSINS's match filter. Although Figure 4.5 indicates that between 5% and 6% of the data are flagged for channel 7 RFI by χ^2, as indicated in Table 4.1, roughly 27.4% of the two-minute observation *files* are flagged for

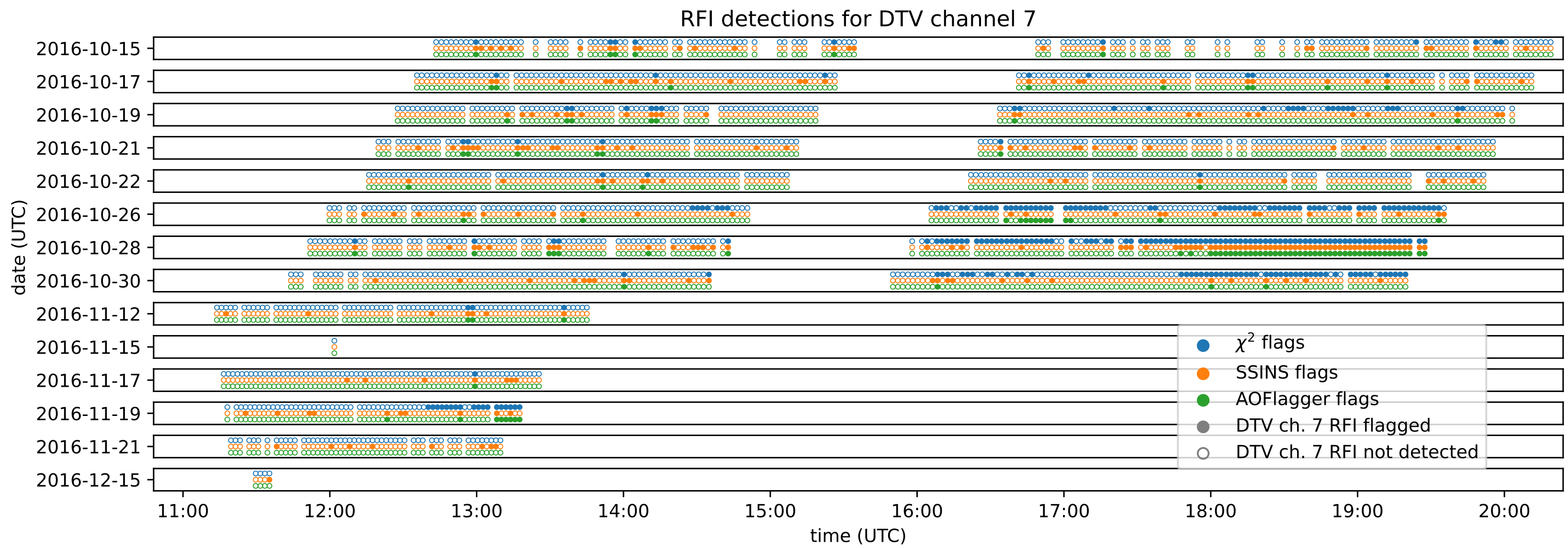
Figure 4.7: A Venn diagram comparing the percentage of DTV events flagged by the χ^2, SSINS, and AOFlagger algorithms. Most of the events flagged by AOFlagger are detected by either χ^2 or SSINS, but there are a significant number which are flagged by χ^2 and missed by SSINS, and vice-versa. The circles in this diagram are to scale, although overlaps are not.



this type of RFI. That is to say, 27.4% of files have 5% or more of the DTV channel 7 band flagged (if the detection was with χ^2 or AOFlagger) or was tagged by SSINS's match filter as containing DTV channel 7, but if all of the data is considered in bulk, between 5% and 6% of the total data are contaminated by channel 7 DTV (according to χ^2-based flagging).

When the flagging fraction of the channel 7 band is compared across algorithms and plotted versus time, as in Figure 4.8, we observe that there are certain DTV events which are flagged by χ^2 but not the other algorithms, and that the same could be said for SSINS.

Figure 4.8: A comparison of χ^2, SSINS, and AOFlagger performance on flagging DTV channel 7, the most commonly-detected source of RFI in our dataset. The x-axis represents the local time of the observation in UTC, while each vertical panel represents observations from a different date. Within a panel, each observation is represented by three circles: the top row (blue circles) is our χ^2 method, the middle row (orange circles) is SSINS, and the bottom row (green circles) is AOFlagger. Filled circles indicate that RFI was detected by an algorithm in the channel 7 band for that file. "Detected" meaning, in the case of χ^2 and AOFlagger, that $> 5\%$ of possibly contaminated bins were flagged, and in the case of SSINS, that the match filter identified DTV channel 7 in this file. χ^2 seems well-suited to detecting long-duration DTV events which are picked up only sporadically by SSINS and AOFlagger. Meanwhile, SSINS detects isolated events which are not detected by χ^2 and, often, AOFlagger.



The number of files flagged in the channel 7 band for each algorithm is summarized in Table 4.1 and in the Venn diagram in Figure 4.7.

While every event caught by AOFlagger was also found by either SSINS, χ^2, or both, there is less overlap between the events flagged by SSINS and χ^2. Of all events flagged by χ^2, about 56.5% are unflagged by other algorithms, and of all the events flagged by SSINS, roughly 52.2% are unflagged by other algorithms. This indicates that the SSINS and χ^2 algorithms are both suited to spotting exclusive categories of events, whereas there is a significant overlap between AOFlagger and either SSINS or χ^2.

4.4 Discussion

4.4.1 RFI classification

Several distinct types of RFI can be seen in χ^2 data. The most common type of RFI is digital television (DTV) signals, usually in Australian channel 7 (181-188 MHz), although channel 9 (195-202 MHz) is sometimes also present. Channels 6 (174-181 MHz) and 8 (188-195 MHz) are the rarest to see, usually only visible in very strong DTV RFI events which contain 7 and 9 as well. Channel 7 is flagged in roughly 5-6% of the data, channels 6, 8, and 9 are all flagged in less than 1% of the data.

We also observe broad-band, short-duration RFI which is visible through most of the frequencies observed by the MWA. In general, SSINS seems much more sensitive to these events, which seem difficult to flag using χ^2.

Finally, we can see long-duration, narrow-band RFI which fits entirely into a single 40 kHz frequency-bin at central frequencies of 196.175 MHz, and 194.455 MHz, as discussed in section 4.3.1. An example event is visualized in Figure A.8 in A.4.

When flagging occupancy for the 196.175 MHz narrow-band RFI is plotted against time, as in Figure 4.9, it seems to represent a repeating signal with a period of around 105 minutes. The source of this radiation is unclear, although the 105 minute period does fall in a range of periods typical for low Earth orbit satellites. Further investigation is needed to track down this signal's provenance.

4.4.2 Comparing events flagged by different algorithms

Referring to Figure 4.8, we observe a few patterns:

- Observations flagged by only χ^2 tend to occur clustered in time, and are more common in the data taken later at night (which correspond to EoR1 observations).
  - an example of an observation that was flagged by χ^2 alone, and was a part of one of these late-night clusters is Figure A.2 (this and other examples in this section may be found in A.4). This observation is from the second half of the night on 2016-10-26, around 18:15 UTC.
- SSINS flags also cluster in time somewhat, but less consistently than χ^2 flags, and the clusters seem shorter in duration on average (consistent with the time clustering properties found in (Wilensky et al., 2023)).
  - Figures A.3 and A.4 are part of a short cluster of SSINS flags, at around 14:30 on 2016-10-30.
  - While Figure A.4 is a bright signal in SSINS, it is brief enough to be reduced in strength by time-averaging. Figure A.3, on the other hand, is much weaker when visualized in SSINS, even though it has a longer duration. It is unclear why the event in Figure A.3 was visible to SSINS but not to χ^2.
- There are also many observations flagged by SSINS alone which are isolated from other

flags in time.

  - An example of an event which is isolated from other detections may be found at Figure A.5. Like Figure A.4, it is bright and brief. It occurred on 2016-11-12 shortly after 11:45.

- AOFlagger flags for this dataset overlap entirely with SSINS flags, χ^2 flags, or both.

  - It seems as though AOFlagger is not uniquely sensitive to any of the RFI in this dataset.

  - An event which was flagged by AOFlagger and χ^2 but not seen by SSINS can be seen at Figure A.6. This event occurred just after 13:15 on 2016-11-19.

In general, it seems that χ^2 is more consistent in finding longer-timescale RFI than SSINS, and conversely, SSINS seems better at detecting short-timescale RFI. An example of the former can be seen in figure A.2, and an example of the latter in Figure A.4. Given that the first step to processing the χ^2 values of MWA data is to time-average the files, and this would have the tendency to "wash out" shorter-timescale events, this pattern makes sense. Conversely, since SSINS relies on time-differencing to remove slowly varying signals, it is perhaps not surprising that it misses events with long durations.

SSINS also sees certain wide-band, short duration RFI, termed "streaks", which are not as visible to χ^2. An example is Figure A.9.

4.4.3 χ^2 and RFI

In Section 4.1, we discussed how χ^2 is calculated and how it could be influenced by RFI. We came to the conclusion that the heightened χ^2 seen with RFI are not due to an underestimation of the noise variance σ^2, but instead due to genuinely less redundant visibilities between nominally redundant baselines.

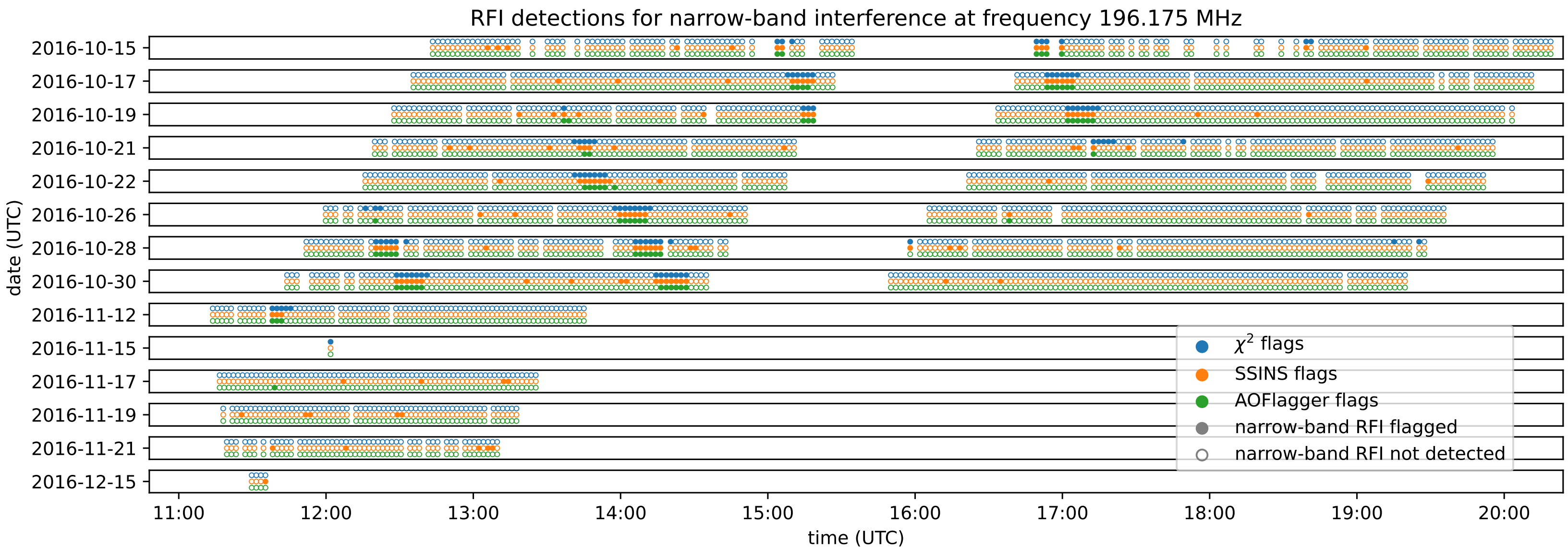
RFI detected by χ^2 has in the past been hypothesized to be locally produced, which would lead to certain antennas detecting the signal more strongly than others, thus increasing the χ^2 metric. In fact, in our analysis, it became clear that the RFI picked up by χ^2 does not seem to be associated with local transmitters, as it doesn't peak in particular areas of the array for individual observations — when χ^2 is elevated for a certain observation, it is increased for every antenna in the array.

Anecdotally, our results suggest the following mechanism through which χ^2 is heightened by RFI: RFI is generally brighter than sky signal, so any differences between antenna beams will be magnified, especially in the sidelobes, which vary more between antennas than the main lobe of the beam. Furthermore, since the MWA is located far from any DTV transmitters, and exposure to direct (non-reflected) DTV signals could be due to tropospheric ducting over the horizon, we would expect this bright RFI to be located near the horizon, in the sidelobes of the MWA where non-redundancy is most pronounced (Line et al., 2018; Chokshi et al., 2021).

As a result, we might expect this χ^2 flagger to have a lower success-rate when RFI is located near the zenith, i.e., within the main lobe, which is much more redundant between redundant baselines. RFI appearing within the main lobe is more likely to be reflected from airplanes or other objects, a detection-case which SSINS has been shown to handle well (Wilensky et al., 2019).

In order to verify this hypothesis and its implications, direct imaging of RFI caught by a χ^2-based flagger may be necessary. This is an avenue for future work on this topic.

Figure 4.9: The same as Figure 4.8, except for the 196.175 MHz narrow-band RFI signal (as opposed to channel 7 DTV). This signal appears to be periodic with a period of around 105 minutes.



CHAPTER 5

Conclusion

Four things do not come back: the spoken word, the sped arrow,
the past life, and the neglected opportunity.

TED CHIANG
THE MERCHANT AND THE ALCHEMIST'S GATE

RFI detection and flagging remain a formidable challenge in the realm of 21 cm Epoch of Reionization cosmology. Without an algorithm (or suite of algorithms) which can detect and excise RFI at a sensitivity heretofore unachieved, instruments like the MWA and HERA may be unable to deliver an accurate power spectrum of the Epoch of Reionization's 21 cm radiation.

In this work, we have presented χ^2 from redundant calibration as a new approach to flagging RFI in data from the MWA. We have demonstrated the effectiveness of this method and compared its efficacy at flagging different kinds of RFI than AOFlagger and SSINS.

χ^2 flagging as implemented here seems effective at flagging the DTV RFI events which are all too common at the MWA site. In the realm of long-duration, weak DTV events, χ^2

finds RFI that's missed by the current state-of-the-art algorithm SSINS. However, χ^2 should not be taken as a replacement for or improvement on SSINS; in fact, the algorithms seem quite complementary, each being able to detect RFI that remains invisible to the other.

There are many potential avenues for future work using this tool. The flagging algorithm based on the modified z-scores of χ^2/nDoF could be improved with features like those found in the SSINS flagging suite, such as applying a threshold to fully flag a DTV channel when sufficient RFI has been detected within it. Furthermore, a machine-learning approach to flag generation could potentially be sensitive to even weaker and more tenuous events, which are visible to the human eye but difficult to pick out algorithmically.

In the course of this work, a repeating narrow-band RFI source was identified at 196.175 MHz with a period of around 105 minutes. Identifying the source of this RFI is another opportunity for further work.

Producing a cosmological 21 cm power spectrum using data screened with χ^2 RFI flagging alongside SSINS is another near-term goal. Furthermore, it would be instructive to investigate whether this hypothetical power spectrum still carries the statistical footprint of ultra-faint RFI as described in Wilensky et al. (2023).

That the χ^2 values produced by redundant calibration are sensitive to weak RFI is a happy accident, which will hopefully become useful in improving the quality of the data that goes into serious analyses in the future. This work represents the first step in bringing this promising new method into the pipelines of new 21 cm data-processing efforts.

Appendix A

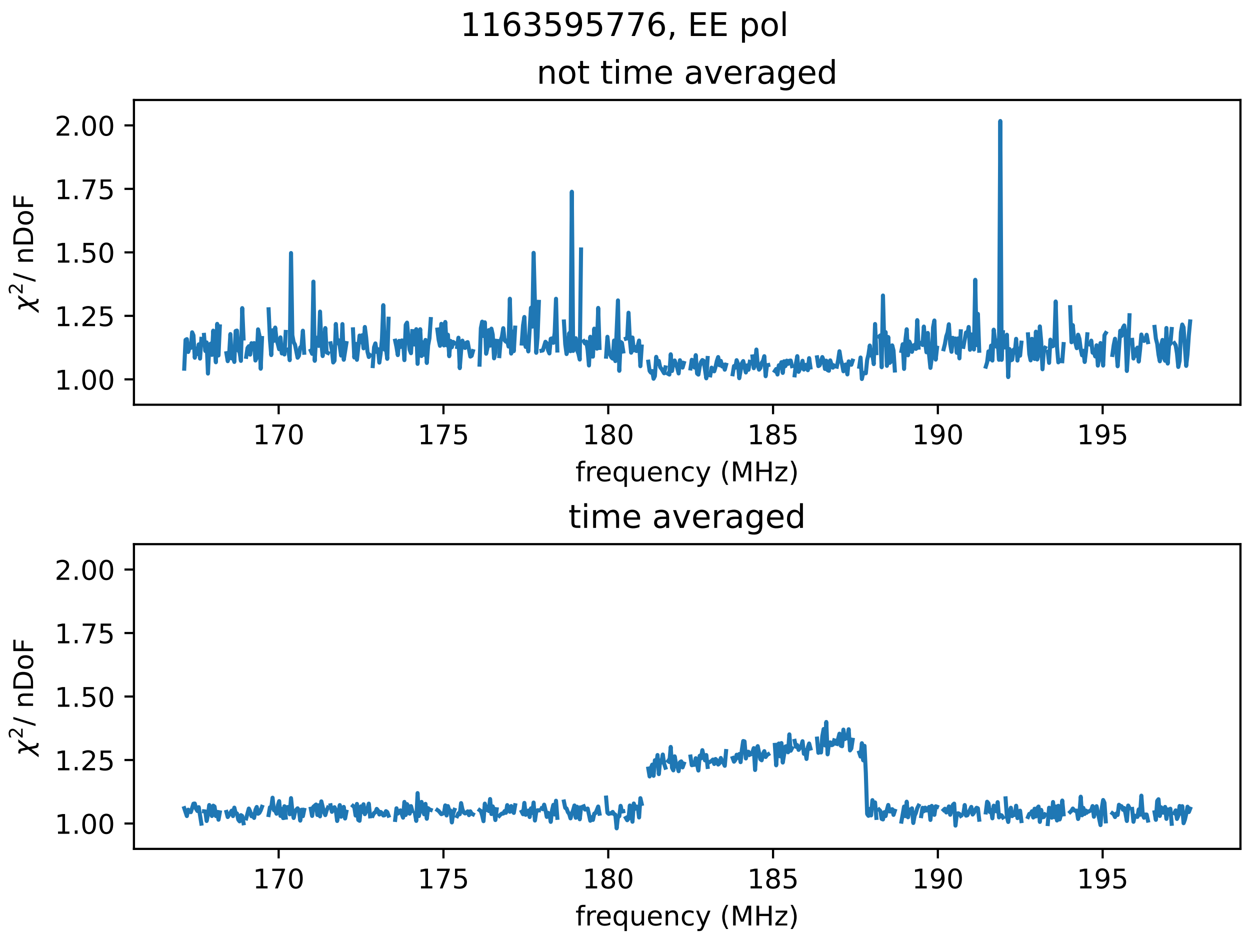
A.1 The effects of time-averaging on χ^2/nDoF

In an initial analysis, the data were processed with `hera_cal` without the additional time averaging step described in Section 2.4. The result was that the χ^2 values were noisy and elevated in regions where there is no RFI.

The upper plot in Figure A.1 is a single time-slice of the χ^2/nDoF generated by `hera_cal` from a file that was not first downsampled in time. This file is known to have DTV channel 7 RFI present, visible between 181 MHz and 188 MHz. Notably, the presence of RFI increases the signal-to-noise level in those frequencies, and so there is much less noise propagated through to χ^2 in that band.

The lower plot of Figure A.1, on the other hand, is also a single time-slice of the χ^2/nDoF generated by `hera_cal` from the same file at approximately the same time, when the data had been time-averaged before calibrating. This plot showcases the lower noise levels afforded by time-averaging, and the stark contrast between frequencies with and without RFI.

Figure A.1: A single time-slice of χ^2/nDoF, generated by hera_cal without downsampling in time first (top) and after downsampling in time (bottom). The result without downsampling is noisy and difficult to process to identify RFI. After downsampling, noise in the χ^2 metric is significantly reduced, and the RFI is visibly elevated over the noise floor.



A.1.1 Weighting χ^2 by NSamples to correct noise discrepancy

When pyuvdata downsamples in time, it does not include flagged data in the calculated average. Furthermore, when working with MWA data, it is good practice to flag a few times at the beginning and end of the file, which often contain bad data. Noise is reduced by time-averaging, so we observe that χ^2 generated from MWA data which have been flagged at the beginning and end of the file in this manner have an elevated noise-floor.

We can correct for this discrepancy by multiplying each χ^2 value by the corresponding element in pyuvdata's NSample_array, which tracks the fraction of unflagged data which were averaged into a given time/frequency bin.

When averaged, noise is decreased by a factor of $\sqrt{N}$, where N is the total number of data-points averaged together. Because σ^2 scales with the square of the noise level, and forms the denominator of equation (3.2), we expect χ^2 to scale inversely with N. It is sufficient, then, to normalize the noise level by multiplying χ^2 by the corresponding element in the UVData object's NSample_array.

A.2 redcal_run() parameters for use on MWA data

A few custom parameters are necessary to make `hera_cal` work with MWA data. The parameters we used were

```
max_dims=3, oc_conv_crit=1e-10, oc_maxiter=2000, check_every=10,
    check_after=500, ant_z_thresh=100.0, gain=0.3
```

A few of these warrant further explanation – `max_dims=3` is absolutely necessary, because the spatial separation between the two hexagons introduces an extra dimension to the degenerate subspace inherent to redundant calibration solutions, corresponding to an arbitrary phase separation between the hexagons, in addition to the three other degenerate parameters, which correspond to the tip/tilt of the array and an overall phase. Setting `ant_z_thresh=100` eliminates the software's throwing out of high-z-score antennas when calculating solutions, because we are *not* planning to use the gain solutions for calibration. The remaining parameters are the result of some by-hand fine-tuning, but could probably be improved with a more systematic parameter-space search.

A.3 chisq_flagging.py

```
import numpy as np
import numpy.ma as ma

def modified_z_scores(data):
    """
    Computes the modified z-scores for an array of data.

    This function calculates the modified z-scores for each data point in the input array,
    which is a robust measure of outliers. The modified z-score uses the median absolute deviation
    (MAD) instead of the standard deviation, making it more resilient to outliers in the data.

    Parameters:
    - data (numpy.ma.MaskedArray): A masked array of data points for which to calculate the modified Z-scores.

    Returns:
    - numpy.ma.MaskedArray: An array of modified z-scores for each data point in the input array.
                            If MAD is 0, returns an array of zeros with the same shape as the input.
    """

# data = ma.masked_invalid(data) # Mask any invalid (NaN) values
    median = ma.median(data)
    deviations = ma.abs(data - median)
    mad = ma.median(deviations)
    if mad == 0:
        return ma.zeros_like(data)
    modified_z_scores = 0.6745 * (data - median) / mad
    return modified_z_scores

def modified_z_score(point, data):
    """
    Calculates the modified z-score of a single data point given an array of data.

    This function computes the modified z-score for a specific data point against an array of data,
    using the median absolute deviation (MAD) for robustness. It is useful for assessing how
    anomalous a single point is within the context of a given dataset.

    Parameters:
    - point (float): The data point for which the modified z-score is to be calculated.
    - data (numpy.ma.MaskedArray): The array of data points against which the score is calculated.

    Returns:
    - float: The modified z-score of the specified data point.
    """

# data = ma.masked_invalid(data) # Mask any invalid (NaN) values
    median = ma.median(data)
    deviations = ma.abs(data - median)
    mad = ma.median(deviations)
    score = 0.6745 * (point - median) / mad
    return score
```

```python
def value_from_modified_z_score(data, desired_z_score):
    """
    Determines the data value that corresponds to a specified modified z-score in a given dataset.

    Given an array of data and a desired modified z-score, this function calculates the data value that
    would have that z-score in the context of the provided dataset. It utilizes the median and
    median absolute deviation (MAD) of the dataset for the calculation, providing a robust measure
    even in the presence of outliers.

    Parameters:
    - data (numpy.ma.MaskedArray): The array of data points used for the calculation.
    - desired_z_score (float): The modified z-score for which the corresponding data value is desired.

    Returns:
    - float: The data value that corresponds to the specified modified z-score within the dataset.
             Returns the median of the dataset if MAD is 0.
    """
# data = ma.masked_invalid(data) # Mask any invalid (NaN) values
    median = ma.median(data)
    deviations = ma.abs(data - median)
    mad = ma.median(deviations)
    if mad == 0:
        return median
    value = median + (desired_z_score * mad) / 0.6745
    return value

def iteratively_flag(chisqs):
    """
    Iteratively flags outliers in chi-square datasets based on modified z-scores.

    This function applies an iterative flagging process to two chi-square datasets ('Jee' and 'Jnn').
    It calculates modified z-scores for each data point in the datasets and flags those with a z-score above 4.0.
    The process repeats until no new flags are added.

    Parameters:
    - chisqs (dict): A dictionary containing two numpy masked arrays under the keys 'Jee' and 'Jnn',
                     representing chi-square datasets of two polarizations.

    Returns:
    - dict: A dictionary with the same structure as the input, where data points have been flagged
            (masked) if deemed outliers based on the iterative flagging process.
    """
    chisq_ee = chisqs['Jee']
    chisq_nn = chisqs['Jnn']
    last_chisq_ee = ma.zeros(chisq_ee.shape)
    last_chisq_nn = ma.zeros(chisq_nn.shape)

    while True:
        if np.all(chisq_ee.mask == last_chisq_ee.mask) and np.all(chisq_nn.mask == last_chisq_nn.mask):
            break
        last_chisq_ee = chisq_ee.copy()
        last_chisq_nn = chisq_nn.copy()
```

```
        xx_new_flags = (modified_z_scores(chisq_ee)>4.0).data
        yy_new_flags = (modified_z_scores(chisq_nn)>4.0).data

        chisq_ee = ma.masked_array(chisq_ee.data, mask=(chisq_ee.mask + xx_new_flags))
        chisq_nn = ma.masked_array(chisq_nn.data, mask=(chisq_nn.mask + yy_new_flags))
    return {'Jee': chisq_ee, 'Jnn': chisq_nn}

def adjacent_cells(chisqs, i, j):
    """
    Identifies unflagged (non-masked) adjacent cells around a specified cell in a chi-square dataset,
    considering special handling near coarse-band edges or centers.

    This function takes into account the layout of coarse bands, avoiding the inclusion of cells
    that are directly adjacent to coarse-band edges or centers. It explores left, up, right, and down
    directions from the specified cell, extending the search if a coarse band is encountered,
    until unflagged cells are found or edges of the dataset are reached.

    Parameters:
    - chisqs (numpy.ndarray): A numpy masked array representing a chi-square dataset.
    - i (int): The row index of the specified cell.
    - j (int): The column index of the specified cell.

    Returns:
    - list of tuples: A list containing tuples for each unflagged adjacent cell, where each tuple contains
                      the cell's value, row index, and column index.

    Raises:
    - IndexError: If the specified indices are out of the bounds of the dataset or are negative.
    """
    if i >= chisqs.shape[0] or j >= chisqs.shape[1]:
        raise IndexError('one or both of your indices are greater than the extent of the array')
    if i < 0 or j < 0:
        raise IndexError('indices must be non-negative for this function')

    adjacent_cells = []

    coarse_bands = np.zeros(chisqs.shape, dtype=bool)
    for time in coarse_bands:
        for k, element in enumerate(time):
            remainder = k%32
            if remainder in [0, 1, 16, 30, 31]:
                time[k] = True

    # left
    m, n = i, j-1
    if n < 0: # we are on the leftmost edge
        pass
    elif coarse_bands[m, n]: # we are adjacent to a coarse-band edge or center
        for k in range(1,4):
            if n-k >= 0 and not coarse_bands[m, n-k]:
                n = n-k
```

```
                if not chisqs.mask[m, n]:
                    adjacent_cells.append((chisqs[m, n], m, n))
                break
    else:
        if not chisqs.mask[m, n]:
            adjacent_cells.append((chisqs[m, n], m, n))

    # up
    m, n = i-1, j
    if m < 0: # we are on the top
        pass
    else:
        if not chisqs.mask[m, n]:
            adjacent_cells.append((chisqs[m, n], m, n))

    # right
    m, n = i, j+1
    if n >= chisqs.shape[1]: # we are on the rightmost edge
        pass
    elif coarse_bands[m, n]: # we are adjacent to a coarse-band edge or center
        for k in range(1, 4):
            if n+k < chisqs.shape[1] and not coarse_bands[m, n+k]:
                n = n+k
                if not chisqs.mask[m,n]:
                    adjacent_cells.append((chisqs[m, n], m, n))
                break
    else:
        if not chisqs.mask[m, n]:
            adjacent_cells.append((chisqs[m, n], m, n))

    # down
    m, n = i+1, j
    if m >= chisqs.shape[0]: # we are on the bottom
        pass
    else:
        if not chisqs.mask[m, n]:
            adjacent_cells.append((chisqs[m, n], m, n))

    return adjacent_cells


def watershed(chisqs):
    """
    Flags outliers in chi-square datasets using a watershed-like algorithm, following an initial iterative
    flagging process.

    The function starts by applying `iteratively_flag` to preliminarily flag outliers in the 'Jee' and 'Jnn'
    datasets based on modified z-scores. It then employs a watershed-like algorithm, which iteratively
    expands the flagging from already flagged cells to their adjacent unflagged cells based on a modified
    z-score criterion. This expansion mimics the watershed technique in image processing, where regions grow
    from seeded points. The process continues until no new cells meet the criteria for flagging.

```

```
    This method is particularly effective in flagging contiguous outlier regions that might not be identified
    in the initial flagging phase due to their local context within the dataset.

    Parameters:
    - chisqs (dict): A dictionary containing two numpy masked arrays under the keys 'Jee' and 'Jnn',
                     representing two different chi-square datasets.

    Returns:
    - dict: A dictionary with the same structure as the input, where additional data points have been
            flagged based on the watershed-like algorithm, in addition to the initial iterative flagging.
            This includes updates to the masks of the 'Jee' and 'Jnn' arrays to reflect the expanded
            flagging regions.
    """
    chisqs = iteratively_flag(chisqs) # preliminary flagging by iterative z-score

    chisq_ee = chisqs['Jee']
    chisq_nn = chisqs['Jnn']

    coarse_bands = np.zeros(chisq_ee.shape)
    for time in coarse_bands:
        for k, element in enumerate(time):
            remainder = k%32
            if remainder in [0, 1, 16, 30, 31]:
                time[k] = True

    existing_flags_ee = chisq_ee.mask - coarse_bands
    existing_flags_nn = chisq_nn.mask - coarse_bands

    added_flags_xx = True
    while added_flags_xx:
        added_flags_xx = False
        for i, time in enumerate(chisq_ee):
            for j, element in enumerate(chisq_ee[i]):
                if existing_flags_ee[i,j] == True:
                    for cell in adjacent_cells(chisq_ee, i, j):
                        if modified_z_score(cell[0], chisq_ee) > 2:
                            # flag the point
                            if not existing_flags_ee[cell[1], cell[2]]:
                                existing_flags_ee[cell[1], cell[2]] = True
                                added_flags_xx = True

    added_flags_yy = True
    while added_flags_yy:
        added_flags_yy = False
        for i, time in enumerate(chisq_nn):
            for j, element in enumerate(chisq_nn[i]):
                if existing_flags_nn[i,j] == True:
                    for cell in adjacent_cells(chisq_nn, i, j):
                        if modified_z_score(cell[0], chisq_nn) > 2:
                            # flag the point
                            if not existing_flags_nn[cell[1], cell[2]]:
                                existing_flags_nn[cell[1], cell[2]] = True
```

```
                    added_flags_yy = True


    flags_ee = existing_flags_ee
    flags_nn = existing_flags_nn

    return {'Jee': ma.masked_array(chisq_ee.data, mask=flags_ee+coarse_bands),
            'Jnn': ma.masked_array(chisq_nn.data, mask=flags_nn+coarse_bands)}
```

A.4 A selection of interesting RFI events

In the course of this research, we encountered many interesting individual RFI events. In this appendix, we have included a selection of particularly illustrative or interesting events.

Figures A.2 through A.10 show time-frequency waterfall plots of χ^2/nDoF modified z-scores (top row), SSINS's Incoherent Noise Spectrum waterfall plot (INS, middle row), and AOFlagger flags averaged over baseline as described in section 4.2.1 (bottom row). Each column shows one of the two linear polarizations measured by the MWA.

Figure A.2: A typical example of an observation with RFI in DTV channel 7 which is detectable with χ^2, and not with other flagging algorithms. This observation is a part of a long string of DTV detections seen by χ^2 alone.

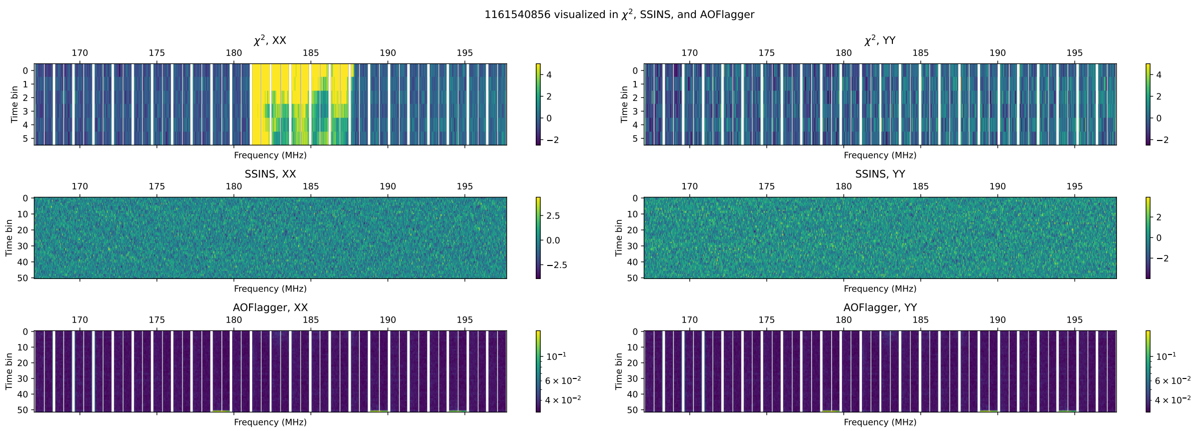


Figure A.3: This SSINS DTV detection occurs as part of a small cluster of files flagged by SSINS alone. The event is also visible in AOFlagger, and a hint of it is visible in χ^2, but it did not reach the threshold for being listed as being "detected" by any algorithm besides SSINS.

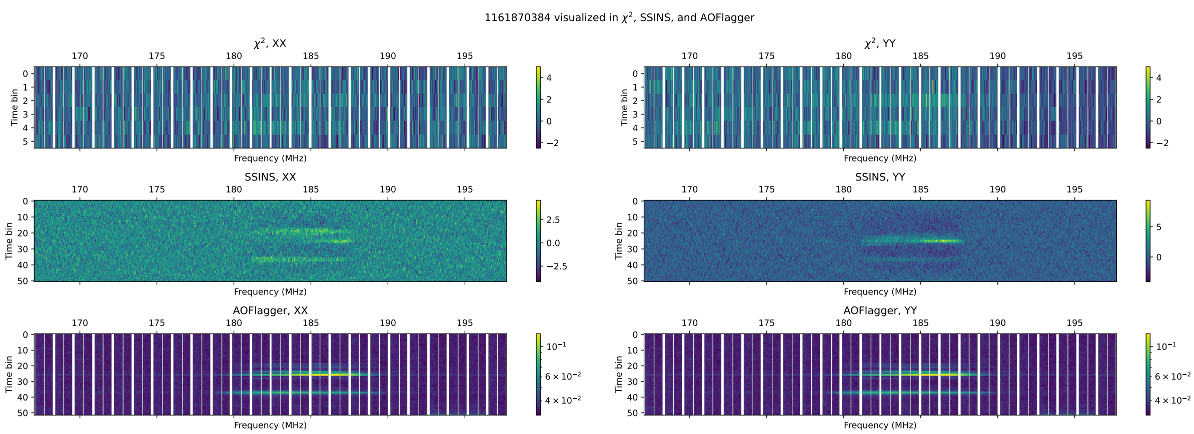


Figure A.4: This SSINS DTV detection occurs as part of the same small cluster of SSINS detections as Figure A.3. The event is also visible in AOFlagger, but it did not reach the threshold for being listed as being "detected" by AOFlagger or χ^2.

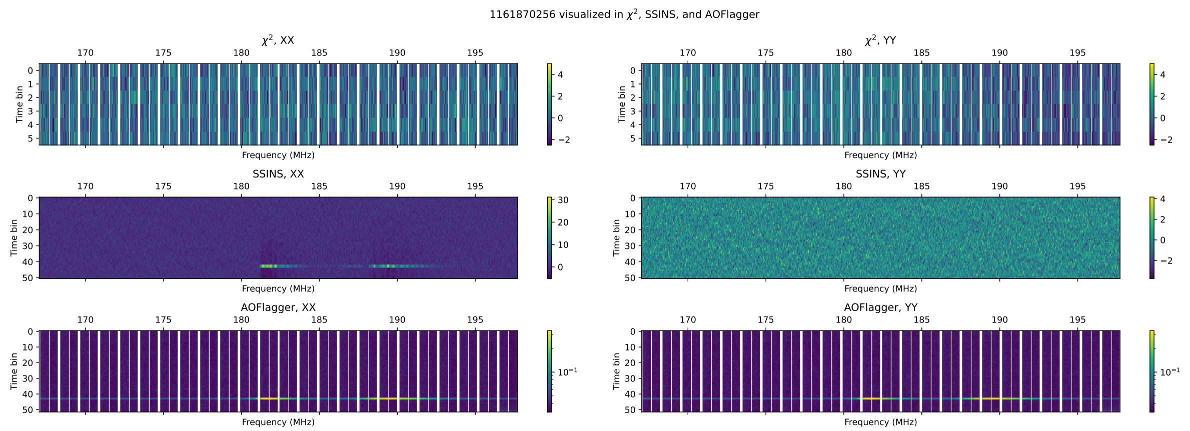
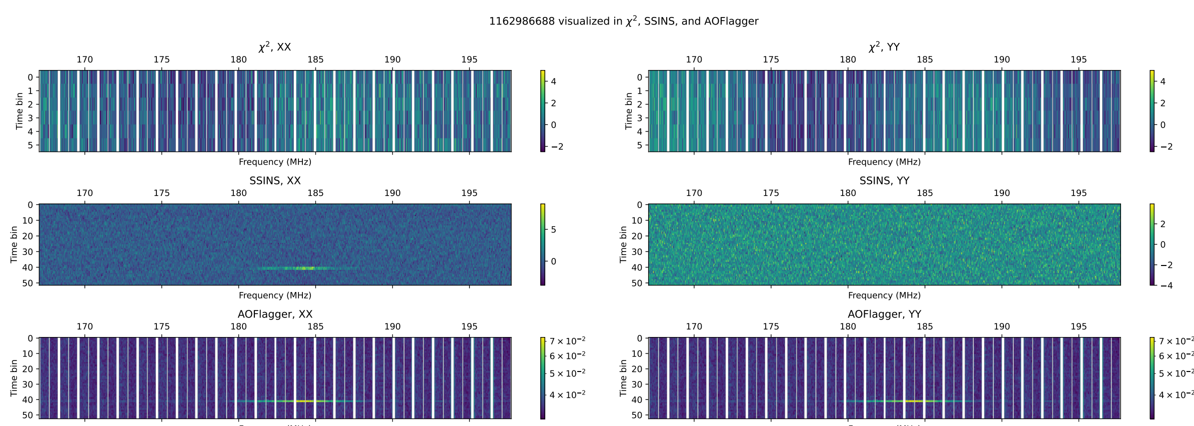


Figure A.5: This SSINS DTV detection occurs isolated in time from other RFI events.



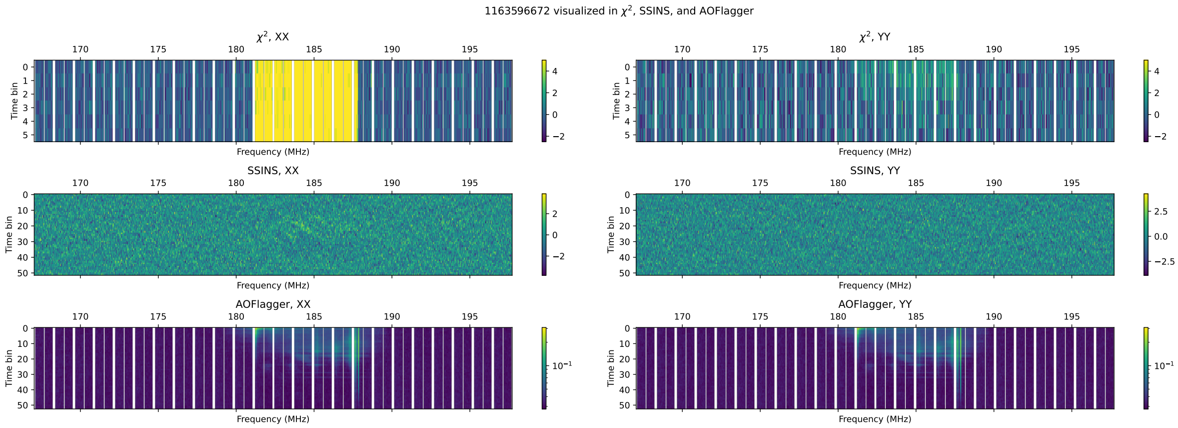


Figure A.6: This DTV channel 7 event was detected by χ^2 and AOFlagger, but missed by SSINS.

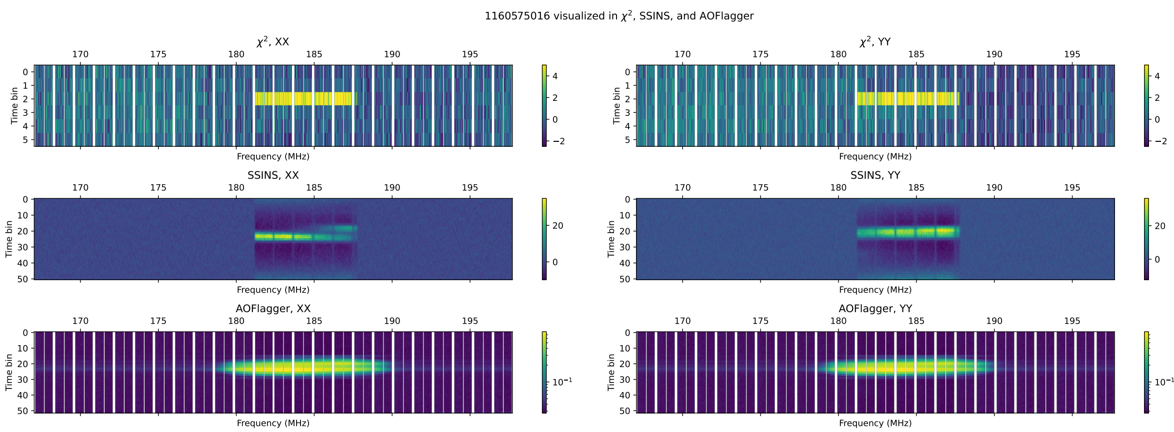


Figure A.7: An example of DTV channel 7, detected by all three algorithms.

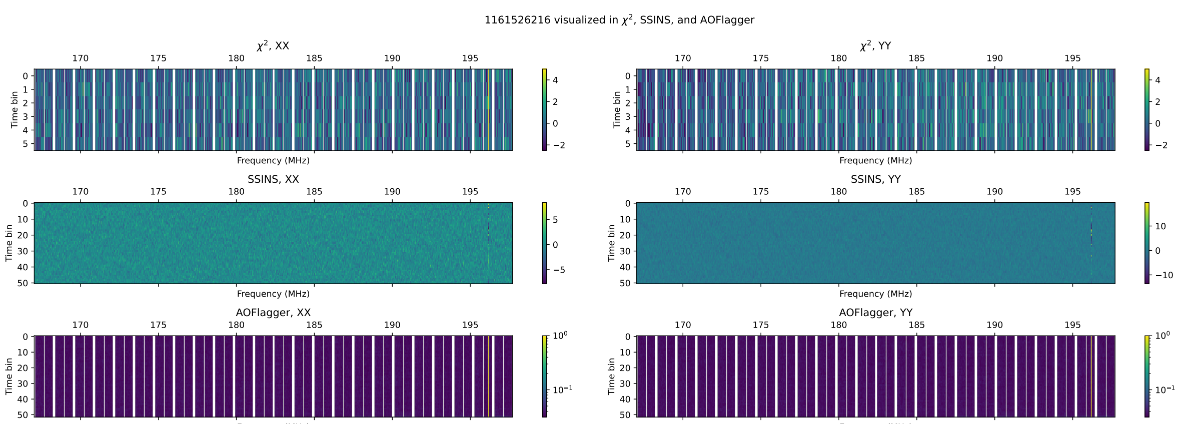


Figure A.8: An example of narrow-band RFI at 196.175 MHz, detected by all three algorithms.

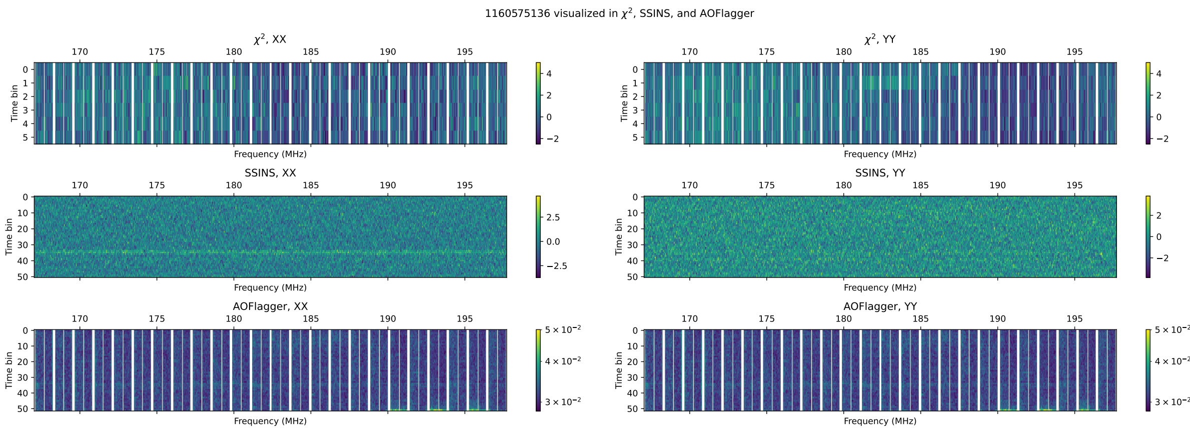


Figure A.9: An example of a "streak" found in SSINS, but invisible to χ^2.

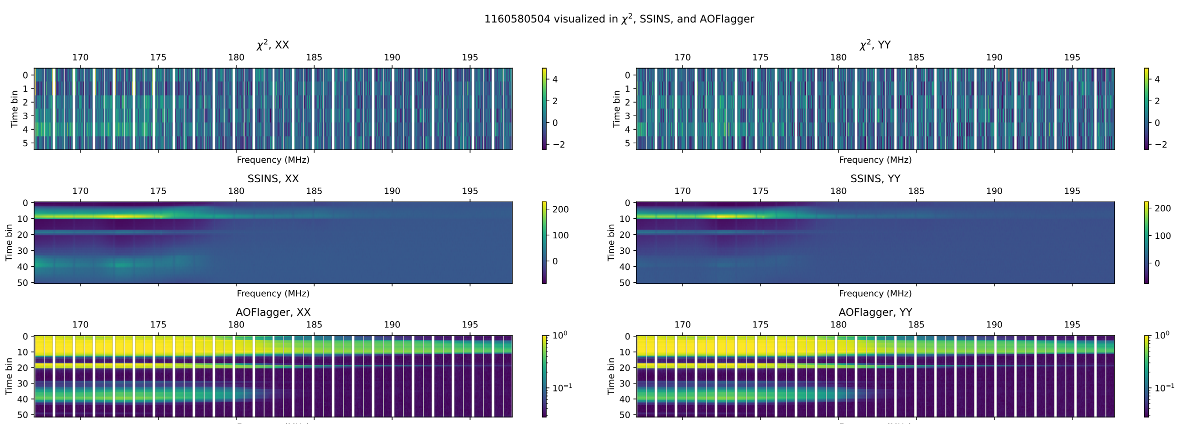


Figure A.10: An unusual event, which might have been due to weather, is also nearly invisible to χ^2.

Bibliography

Beardsley, A., Hazelton, B., Sullivan, I., Carroll, P., Barry, N., Rahimi, M., Pindor, B., Trott, C., Line, J., Jacobs, D. C., et al. (2016). First season mwa eor power spectrum results at redshift 7. *The Astrophysical Journal*, 833(1):102.

Bhat, N. D. R., Swainston, N. A., McSweeney, S. J., Xue, M., Meyers, B., Kudale, S., Dai, S., Tremblay, S. E., van Straten, W., Shannon, R. M., Smith, K. R., Sokolowski, M., Ord, S. M., Sleap, G., Williams, A., Hancock, P. J., Lange, R., Tocknell, J., Johnston-Hollitt, M., Kaplan, D. L., Tingay, S. J., and Walker, M. (2023). The southern-sky mwa rapid two-metre (smart) pulsar survey—ii. survey status, pulsar census, and first pulsar discoveries. *Publications of the Astronomical Society of Australia*, 40.

Chokshi, A., Line, J. L. B., Barry, N., Ung, D., Kenney, D., McPhail, A., Williams, A., and Webster, R. L. (2021). Dual polarization measurements of mwa beampatterns at 137 mhz. *Monthly Notices of the Royal Astronomical Society*, 502(2):1990–2004.

Choudhuri, S., Bull, P., and Garsden, H. (2021). Patterns of primary beam non-redundancy in close-packed 21 cm array observations. *Monthly Notices of the Royal Astronomical Society*, 506(2):2066–2088.

Choudhury, T. R. (2022). A short introduction to reionization physics. *General Relativity and Gravitation*, 54(9).

DeBoer, D. R., Parsons, A. R., Aguirre, J. E., Alexander, P., Ali, Z. S., Beardsley, A. P., Bernardi, G., Bowman, J. D., Bradley, R. F., Carilli, C. L., Cheng, C., Acedo, E. d. L., Dillon, J. S., Ewall-Wice, A., Fadana, G., Fagnoni, N., Fritz, R., Furlanetto, S. R., Glendenning, B., Greig, B., Grobbelaar, J., Hazelton, B. J., Hewitt, J. N., Hickish, J., Jacobs, D. C., Julius, A., Kariseb, M., Kohn, S. A., Lekalake, T., Liu, A., Loots, A., MacMahon, D., Malan, L., Malgas, C., Maree, M., Martinot, Z., Mathison, N., Matsetela, E., Mesinger, A., Morales, M. F., Neben, A. R., Patra, N., Pieterse, S., Pober, J. C., Razavi-Ghods, N., Ringuette, J., Robnett, J., Rosie, K., Sell, R., Smith, C., Syce, A., Tegmark, M., Thyagarajan, N., Williams, P. K. G., and Zheng, H. (2017). Hydrogen epoch of reionization array (hera). *Publications of the Astronomical Society of the Pacific*, 129(974):045001.

Dillon, J. S. (2024). personal communication.

Dillon, J. S., Lee, M., Ali, Z. S., Parsons, A. R., Orosz, N., Nunhokee, C. D., La Plante, P., Beardsley, A. P., Kern, N. S., Abdurashidova, Z., Aguirre, J. E., Alexander, P., Balfour, Y., Bernardi, G., Billings, T. S., Bowman, J. D., Bradley, R. F., Bull, P., Burba, J., Carey, S., Carilli, C. L., Cheng, C., DeBoer, D. R., Dexter, M., de Lera Acedo, E., Ely, J., Ewall-Wice, A., Fagnoni, N., Fritz, R., Furlanetto, S. R., Gale-Sides, K., Glendenning, B., Gorthi, D., Greig, B., Grobbelaar, J., Halday, Z., Hazelton, B. J., Hewitt, J. N., Hickish, J., Jacobs, D. C., Julius, A., Kerrigan, J., Kittiwisit, P., Kohn, S. A., Kolopanis, M., Lanman, A., Lekalake, T., Lewis, D., Liu, A., Ma, Y.-Z., MacMahon, D., Malan, L., Malgas, C., Maree, M., Martinot, Z. E., Matsetela, E., Mesinger, A., Molewa, M., Morales, M. F., Mosiane, T., Murray, S., Neben, A. R., Nikolic, B., Pascua, R., Patra, N., Pieterse, S., Pober, J. C., Razavi-Ghods, N., Ringuette, J., Robnett, J., Rosie, K., Santos, M. G., Sims, P., Smith, C., Syce, A., Tegmark, M., Thyagarajan, N., Williams, P. K. G., and Zheng, H. (2020). Redundant-baseline calibration of the hydrogen epoch of reionization array.

Monthly Notices of the Royal Astronomical Society, 499(4):5840–5861.

Dillon, J. S. and Parsons, A. R. (2016). Redundant Array Configurations for 21 cm Cosmology. ApJ, 826(2):181.

Furlanetto, S. R., Oh, S. P., and Briggs, F. H. (2006). Cosmology at low frequencies: The 21 cm transition and the high-redshift universe. *Physics reports*, 433(4-6):181–301.

Hazelton, B. J., Jacobs, D. C., Pober, J. C., and Beardsley, A. P. (2017). pyuvdata: an interface for astronomical interferometeric datasets in python. *The Journal of Open Source Software*, 2(10):140.

Kerrigan, J., Plante, P. L., Kohn, S., Pober, J. C., Aguirre, J., Abdurashidova, Z., Alexander, P., Ali, Z. S., Balfour, Y., Beardsley, A. P., et al. (2019). Optimizing sparse rfi prediction using deep learning. *Monthly Notices of the Royal Astronomical Society*, 488(2):2605–2615.

Kolopanis, M., Pober, J. C., Jacobs, D. C., and McGraw, S. (2023). New eor power spectrum limits from mwa phase ii using the delay spectrum method and novel systematic rejection. *Monthly Notices of the Royal Astronomical Society*, 521(4):5120–5138.

Li, W., Pober, J., Barry, N., Hazelton, B., Morales, M., Trott, C., Lanman, A., Wilensky, M., Sullivan, I., Beardsley, A., et al. (2019). First season mwa phase ii epoch of reionization power spectrum results at redshift 7. *The Astrophysical Journal*, 887(2):141.

Li, W., Pober, J., Hazelton, B., Barry, N., Morales, M., Sullivan, I., Parsons, A., Ali, Z., Dillon, J., Beardsley, A., et al. (2018). Comparing redundant and sky-model-based interferometric calibration: A first look with phase ii of the mwa. *The Astrophysical Journal*, 863(2):170.

Line, J. L. B., McKinley, B., Rasti, J., Bhardwaj, M., Wayth, R. B., Webster, R. L., Ung, D., Emrich, D., Horsley, L., Beardsley, A., Crosse, B., Franzen, T. M. O., Gaensler, B. M., Johnston-Hollitt, M., Kaplan, D. L., Kenney, D., Morales, M. F., Pallot, D., Steele, K., Tingay, S. J., Trott, C. M., Walker, M., Williams, A., and Wu, C. (2018). In situ measurement of mwa primary beam variation using orbcomm. *Publications of the Astronomical Society of Australia*, 35.

Liu, A. and Shaw, J. R. (2020). Data analysis for precision 21 cm cosmology. *Publications of the Astronomical Society of the Pacific*, 132(1012):062001.

Liu, A., Tegmark, M., Morrison, S., Lutomirski, A., and Zaldarriaga, M. (2010). Precision calibration of radio interferometers using redundant baselines. *Monthly Notices of the Royal Astronomical Society*, 408(2):1029–1050.

Morales, M. F. and Wyithe, J. S. B. (2010). Reionization and cosmology with 21-cm fluctuations. *Annual review of astronomy and astrophysics*, 48:127–171.

Offringa, A., De Bruyn, A., Biehl, M., Zaroubi, S., Bernardi, G., and Pandey, V. (2010). Post-correlation radio frequency interference classification methods. *Monthly Notices of the Royal Astronomical Society*, 405(1):155–167.

Offringa, A., Van De Gronde, J., and Roerdink, J. (2012). A morphological algorithm for improving radio-frequency interference detection. *Astronomy & astrophysics*, 539:A95.

Offringa, A., Wayth, R., Hurley-Walker, N., Kaplan, D., Barry, N., Beardsley, A., Bell, M., Bernardi, G., Bowman, J., Briggs, F., et al. (2015). The low-frequency environment of the murchison widefield array: radio-frequency interference analysis and mitigation. *Publications of the Astronomical Society of Australia*, 32:e008.

Penzias, A. A. and Wilson, R. W. (1979). A measurement of excess antenna temperature at 4080 mhz. In *A Source Book in Astronomy and Astrophysics, 1900–1975*, pages 873–876. Harvard University Press.

Pritchard, J. R. and Loeb, A. (2012). 21 cm cosmology in the 21st century. *Reports on*

Progress in Physics, 75(8):086901.

Sharma, R., Oberoi, D., and Arjunwadkar, M. (2018). Quantifying weak nonthermal solar radio emission at low radio frequencies. *The Astrophysical Journal*, 852(2):69.

Sokolowski, M., Wayth, R. B., and Lewis, M. (2015). The statistics of low frequency radio interference at the murchison radio-astronomy observatory. In *2015 IEEE Global Electromagnetic Compatibility Conference (GEMCCON)*. IEEE.

Tingay, S. J., Goeke, R., Bowman, J. D., Emrich, D., Ord, S. M., Mitchell, D. A., Morales, M. F., Booler, T., Crosse, B., Wayth, R. B., et al. (2013). The murchison widefield array: The square kilometre array precursor at low radio frequencies. *Publications of the Astronomical Society of Australia*, 30:e007.

van de Hulst, H. C., Muller, C. A., and Oort, J. H. (1954). The spiral structure of the outer part of the Galactic System derived from the hydrogen emission at 21 cm wavelength. Bull. Astron. Inst. Netherlands, 12:117.

van Haarlem, M. P., Wise, M. W., Gunst, A. W., Heald, G., McKean, J. P., Hessels, J. W. T., de Bruyn, A. G., Nijboer, R., Swinbank, J., Fallows, R., Brentjens, M., Nelles, A., Beck, R., Falcke, H., Fender, R., Hörandel, J., Koopmans, L. V. E., Mann, G., Miley, G., Röttgering, H., Stappers, B. W., Wijers, R. A. M. J., Zaroubi, S., van den Akker, M., Alexov, A., Anderson, J., Anderson, K., van Ardenne, A., Arts, M., Asgekar, A., Avruch, I. M., Batejat, F., Bähren, L., Bell, M. E., Bell, M. R., van Bemmel, I., Bennema, P., Bentum, M. J., Bernardi, G., Best, P., Bîrzan, L., Bonafede, A., Boonstra, A.-J., Braun, R., Bregman, J., Breitling, F., van de Brink, R. H., Broderick, J., Broekema, P. C., Brouw, W. N., Brüggen, M., Butcher, H. R., van Cappellen, W., Ciardi, B., Coenen, T., Conway, J., Coolen, A., Corstanje, A., Damstra, S., Davies, O., Deller, A. T., Dettmar, R.-J., van Diepen, G., Dijkstra, K., Donker, P., Doorduin, A., Dromer, J., Drost, M., van Duin, A., Eislöffel, J., van Enst, J., Ferrari, C., Frieswijk, W., Gankema, H., Garrett, M. A., de Gasperin, F., Gerbers, M., de Geus, E., Grießmeier, J.-M., Grit, T., Gruppen, P., Hamaker, J. P., Hassall, T., Hoeft, M., Holties, H. A., Horneffer, A., van der Horst, A., van Houwelingen, A., Huijgen, A., Iacobelli, M., Intema, H., Jackson, N., Jelic, V., de Jong, A., Juette, E., Kant, D., Karastergiou, A., Koers, A., Kollen, H., Kondratiev, V. I., Kooistra, E., Koopman, Y., Koster, A., Kuniyoshi, M., Kramer, M., Kuper, G., Lambropoulos, P., Law, C., van Leeuwen, J., Lemaitre, J., Loose, M., Maat, P., Macario, G., Markoff, S., Masters, J., McFadden, R. A., McKay-Bukowski, D., Meijering, H., Meulman, H., Mevius, M., Middelberg, E., Millenaar, R., Miller-Jones, J. C. A., Mohan, R. N., Mol, J. D., Morawietz, J., Morganti, R., Mulcahy, D. D., Mulder, E., Munk, H., Nieuwenhuis, L., van Nieuwpoort, R., Noordam, J. E., Norden, M., Noutsos, A., Offringa, A. R., Olofsson, H., Omar, A., Orrú, E., Overeem, R., Paas, H., Pandey-Pommier, M., Pandey, V. N., Pizzo, R., Polatidis, A., Rafferty, D., Rawlings, S., Reich, W., de Reijer, J.-P., Reitsma, J., Renting, G. A., Riemers, P., Rol, E., Romein, J. W., Roosjen, J., Ruiter, M., Scaife, A., van der Schaaf, K., Scheers, B., Schellart, P., Schoenmakers, A., Schoonderbeek, G., Serylak, M., Shulevski, A., Sluman, J., Smirnov, O., Sobey, C., Spreeuw, H., Steinmetz, M., Sterks, C. G. M., Stiepel, H.-J., Stuurwold, K., Tagger, M., Tang, Y., Tasse, C., Thomas, I., Thoudam, S., Toribio, M. C., van der Tol, B., Usov, O., van Veelen, M., van der Veen, A.-J., ter Veen, S., Verbiest, J. P. W., Vermeulen, R., Vermaas, N., Vocks, C., Vogt, C., de Vos, M., van der Wal, E., van Weeren, R., Weggemans, H., Weltevrede, P., White, S., Wijnholds, S. J., Wilhelmsson, T., Wucknitz, O., Yatawatta, S., Zarka, P., Zensus, A., and van Zwieten, J. (2013). Lofar: The low-frequency array. *Astronomy & Astrophysics*, 556:A2.

Wayth, R. B., Lenc, E., Bell, M. E., Callingham, J. R., Dwarakanath, K. S., Franzen, T. M. O., For, B. Q., Gaensler, B., Hancock, P., Hindson, L., Hurley-Walker, N., Jackson, C. A., Johnston-Hollitt, M., Kapińska, A. D., McKinley, B., Morgan, J., Offringa, A. R., Procopio, P., Staveley-Smith, L., Wu, C., Zheng, Q., Trott, C. M., Bernardi, G., Bowman, J. D., Briggs, F., Cappallo, R. J., Corey, B. E., Deshpande, A. A., Emrich, D., Goeke, R., Greenhill, L. J., Hazelton, B. J., Kaplan, D. L., Kasper, J. C., Kratzenberg, E., Lonsdale, C. J., Lynch, M. J., McWhirter, S. R., Mitchell, D. A., Morales, M. F., Morgan, E., Oberoi, D., Ord, S. M., Prabu, T., Rogers, A. E. E., Roshi, A., Shankar, N. U., Srivani, K. S., Subrahmanyan, R., Tingay, S. J., Waterson, M., Webster, R. L., Whitney, A. R., Williams, A., and Williams, C. L. (2015). GLEAM: The GaLactic and Extragalactic All-Sky MWA Survey. PASA, 32:e025.

Wayth, R. B., Tingay, S. J., Trott, C. M., Emrich, D., Johnston-Hollitt, M., McKinley, B., Gaensler, B. M., Beardsley, A. P., Booler, T., Crosse, B., Franzen, T. M. O., Horsley, L., Kaplan, D. L., Kenney, D., Morales, M. F., Pallot, D., Sleap, G., Steele, K., Walker, M., Williams, A., Wu, C., Cairns, I. H., Filipovic, M. D., Johnston, S., Murphy, T., Quinn, P., Staveley-Smith, L., Webster, R., and Wyithe, J. S. B. (2018). The phase ii murchison widefield array: Design overview. *Publications of the Astronomical Society of Australia*, 35.

Wilensky, M. J., Barry, N., Morales, M. F., Hazelton, B. J., and Byrne, R. (2020). Quantifying excess power from radio frequency interference in epoch of reionization measurements. *Monthly Notices of the Royal Astronomical Society*, 498(1):265–275.

Wilensky, M. J., Morales, M. F., Hazelton, B. J., Barry, N., Byrne, R., and Roy, S. (2019). Absolving the ssins of precision interferometric radio data: A new technique for mitigating faint radio frequency interference. *Publications of the Astronomical Society of the Pacific*, 131(1005):114507.

Wilensky, M. J., Morales, M. F., Hazelton, B. J., Star, P. L., Barry, N., Byrne, R., Jordan, C., Jacobs, D. C., Pober, J. C., and Trott, C. (2023). Evidence of ultrafaint radio frequency interference in deep 21 cm epoch of reionization power spectra with the murchison wide-field array. *The Astrophysical Journal*, 957(2):78.

Wilensky, M. J., Morales, M. F., Hazelton, B. J., Star, P. L., Barry, N., Byrne, R., Jordan, C. H., Jacobs, D. C., Pober, J. C., and Trott, C. M. (2023). Evidence of Ultrafaint Radio Frequency Interference in Deep 21 cm Epoch of Reionization Power Spectra with the Murchison Wide-field Array. ApJ, 957(2):78.

Wise, J. H. (2019). Cosmic reionisation. *Contemporary Physics*, 60(2):145–163.

Zhang, X., Hancock, P., Devillepoix, H. A. R., Wayth, R. B., Beardsley, A., Crosse, B., Emrich, D., Franzen, T. M. O., Gaensler, B. M., Horsley, L., Johnston-Hollitt, M., Kaplan, D. L., Kenney, D., Morales, M. F., Pallot, D., Steele, K., Tingay, S. J., Trott, C. M., Walker, M., Williams, A., Wu, C., Ji, J., and Ma, Y. (2018). Limits on radio emission from meteors using the MWA. *Monthly Notices of the Royal Astronomical Society*, 477(4):5167–5176.

Zhang, Z., Pober, J. C., Li, W., Hazelton, B. J., Morales, M. F., Trott, C. M., Jordan, C. H., Joseph, R. C., Beardsley, A., Barry, N., et al. (2020). The impact of tandem redundant/sky-based calibration in mwa phase ii data analysis. *Publications of the Astronomical Society of Australia*, 37:e045.

Zheng, H., Tegmark, M., Buza, V., Dillon, J. S., Gharibyan, H., Hickish, J., Kunz, E., Liu, A., Losh, J., Lutomirski, A., et al. (2014). Miteor: a scalable interferometer for precision 21 cm cosmology. *Monthly Notices of the Royal Astronomical Society*, 445(2):1084–1103.