The HERA RFI environment

Saul A. Kohn*

University of Pennsylvania

September 4, 2015

Abstract

In this memo I describe the RFI properties of the HERA site, as recorded by PAPER-128 during the 2014 EoR observation season. A key result is a close-to-complete characterization of detected RFI transmission frequencies present in $\geq 1\%$ of the season (Table 1).

1 Introduction

The PAPER-128 2014 observation season ran from 18th June 2014 through the 30th April 2015. During this run, some 200 nights of data were recorded. A "night", which I will refer to using the JD at the start of observations, consists of twelve hours of observation from 6pm to 6am South African Standard Time (SAST). Observations as processed by the PAPER/HERA correlator are recorded in MIRIAD uv files. These files contain visibilities for each antenna pair in the array. Each integration is 20 seconds long over 1024 frequency bins from 100 to 200 MHz. Each uv file contains 56 integrations per antenna pair, and 72 uv files are recorded per linear polarization (xx, xy, yx, yy) per night.

Early in the PAPER data compression process, visibilities are flagged for radio frequency interference (RFI). This is accomplished by the AIPY script *xrfi_simple.py*, which takes the derivative of the frequency axis of all baselines associated with antenna 1 (2nd row, 4th column of the redundant grid), and flags any frequencies with a derivative $\geqslant 6\sigma$ above the mean. We always flag the band-edges ($\sim$7 MHz on each side), since these frequencies are not useful to us, and always flag the 137 ± 0.6 MHz band associated with ORBCOMM satellite network transmissions (See Section 2.1.3). This process is repeated per integration within each uv file and stored in a PYTHON npz file. This means that any baseline associated with antenna 1 can contribute a flag to the resultant npz file, which in turn is applied to the data.

The result is 280 files of high-time and -frequency resolution files per night per linear polarization containing information about the RFI environment of the HERA site. In this memo I report on the properties of these flags in time- and frequency-space over the 2014 observation season. I have organised it as follows: in Section 2, I analyse the average properties of RFI over the season by stacking flags in local time and normalizing appropriately. In Section 3 I address nights with particularly strange RFI properties. I discuss the implications of my findings and conclude in Section 4.

2 Average Properties

In order to assess the average properties of the HERA RFI environment, I calculated a weighted average of flags over the season. Over 200 nights, one-time occurrences are washed-out beneath the 1% level, allowing me to assess persistent issues.

Nominally, each night should grant 3920 integrations-worth of flags over 1024 frequency bins, per linear polarization. In reality, most of the time this holds true, but occasionally not all files are compressible (hence failing to generate flags) or observations fail to start at the correct time (so there are no data to flag). Also, in the event of an X-engine failure within the correlator, contiguous chunks of the band (in eighths, i.e. 25 MHz across) are flagged-out, usually for the rest of the night.

For this reason, I calculated a weighted average of the flags across the season, but neglected nights with correlator failures or late starts. Weights were simply the number of nights that contained that integration-bin in SAST. The resultant "flag density waterfall" is shown in Figure 1. The color scale is indicative of flagging frequency across the season, and line plots above and to the right of the the waterfall showing the percentage of times and frequencies that were flagged, respectively.

A summary of the persistent (flagged $\geq 1\%$ of the time per channel) RFI frequencies can be found in Table 1. I have investigated each frequency and tried to find the most likely source for each. In most cases, this required looking at the properties in time as well as frequency. Others were more obvious from frequency alone, e.g. the 149.8 MHz transmission frequency from the ISS. Still others I could not track down a convincing explanation for, and these are listed with a '?'. A '?' next to a possible cause indicates that the listed cause is the most prevalent at that frequency,

---

*saulkohn@sas.upenn.edu

Table 1: RFI frequencies and brief characterization for the averaged flags

ν (MHz)	flagged (%)	Cause (Possible)	Notes/ SAST Characterization
103 ± 3	100	BAND EDGE	Built-in to flagger.
107.25 ± 0.25	2.6	FM radio	Constant background at 2% level
107.55 ± 0.05	1.9	FM radio	Constant background at 2% level
108.1 ± 0.4	9	FM radio?	Rises with time, peaking at midnight and 4am
109 ± 0.4	11.5	FM radio?	Rises with time, peaking around 4am
112.8 ± 0.1	1.4	Aircraft?	Constant background at 1% level
114.05 ± 0.85	3.7	?1	Decreases till midnight; peak at 4am
116.55 ± 0.35	2.2	?2	Peak at midnight
120.15 ± 0.35	3.2	Aircraft	Roughly follows CPT$\leftrightarrow$JNB flight times
124.95 ± 0.35	5.5	Aircraft	Roughly follows CPT$\leftrightarrow$JNB flight times
130.25 ± 0.55	4.3	?3	Falls (7pm) and rises (3am) steeply
131.75 ± 0.35	10.3	Aircraft?	Peaks at 6:30, 7:30, 8:30, 9:30, 10 and then a steep falloff
136.05 ± 0.45	33.1	Radar?	Decreases over night
137.35 ± 0.85	100	ORBCOMM	
141.45 ± 0.35	2.1	Mobile phones?	High until 9pm, then at background 1% level
145.85 ± 0.45	10.7	Amateur radio	Strong 9pm-1am – this is the official downlink for ISS-HAM
149.75 ± 0.55	90.7	ISS	"Beeps", but in stacked data peaks 2am
175.15 ± 0.35	20.5	VHF TV (video)	Channel 4. Peaks at 8:30pm, then falls to background 7%
181.15 ± 0.15	1.6	VHF TV (audio)	Channel 4. 2% level turns-off at 10pm
182.15 ± 0.35	75	?4	Decreases until 10pm (to 15%), when it begins a slow rise again
183.2 ± 0.5	89.7	VHF TV (video)	Channel 5. Rises throughout night.
186.25 ± 0.35	4.6	?5	Extreme turn-off at 9:45
189.15 ± 0.35	41.4	VHF TV (audio)	Channel 5. Rises throughout night.
189.9 ± 0.4	100	VHF TV	Channel 6. Built-in to flagger.
191.1 ± 0.3	100	VHF TV	Channel 7. Built-in to flagger.
196 ± 4	100	BAND EDGE	

but that the temporal properties of that cause do not necessarily make sense.

2.1 Characterization

Table 1 provides a brief characterization of each frequency in SAST. In this subsection I elaborate on these characterizations. Many of the characterizations arise from the South African Table of Frequency Allocations (SATFA; [7]).

Figure 2 shows the detail of the top panel of Figure 1. This figure highlights the broad swath of the band from roughly 150 to 180 MHz that is completely clear of RFI. This roughly corresponds to 21 cm redshifts $z = 6.9$ to 8.5. This is one of the reasons that the [6] and [1] limits on the 21 cm power spectrum concentrated on this redshift range – there are simply more unflagged data to average-down with. [3] show that the $z \sim 8$ universe can be considered roughly coeval over an ~ 8 MHz bandwidth. As such, the 30 MHz chunk could be used to create ~ 3 power spectra, as demonstrated in [4]. As we show in Section 2.1.4, the deactivation of VHF TV broadcasts would enable measurements up to the band edge.

2.1.1 FM radio

SATFA lists the frequency band 87.5–108 MHz as available for FM radio broadcasts, leading me to postulate that the low-level RFI we observe in the 107.25 ± 0.25 and 107.55 ± 0.05 MHz bands has FM radio as the leading cause. The 108.1 ± 0.4 and 109 ± 0.4 MHz bands are outside of the official range, and exhibit odd temporal properties for human activity – two peaks at midnight and 4am with a increasing number of flags throughout the average night (see Figure 3).

2.1.2 Aircraft communications

It is difficult to argue that the 112.1 ± 0.1 MHz signal is caused by aircraft communications since it maintains a constant background level. However, SATFA lists this frequency as reserved for aircraft communications and it has been used in the past as a calibration frequency for aircraft instruments [2].

The other aircraft frequencies are obvious, because they closely trace the 2-hour flight from Cape Town to Johannesburg[1]. An example (120.15 ± 0.35 MHz) is shown in Figure 4. SATFA reserves frequencies 108–117.975 MHz for aeronautical radionavigation and 117.975–137 MHz for aeronautical mobile. In Table 1 I list 131.75 ± 0.35 MHz as caused by aircraft since it falls in the aeronautical mobile band, but it does not follow the flight patterns as closely as the other bands.

---

[1]Credit to Danny Jacobs for first spotting this and noting it in an internal PAPER circular in December 2009.
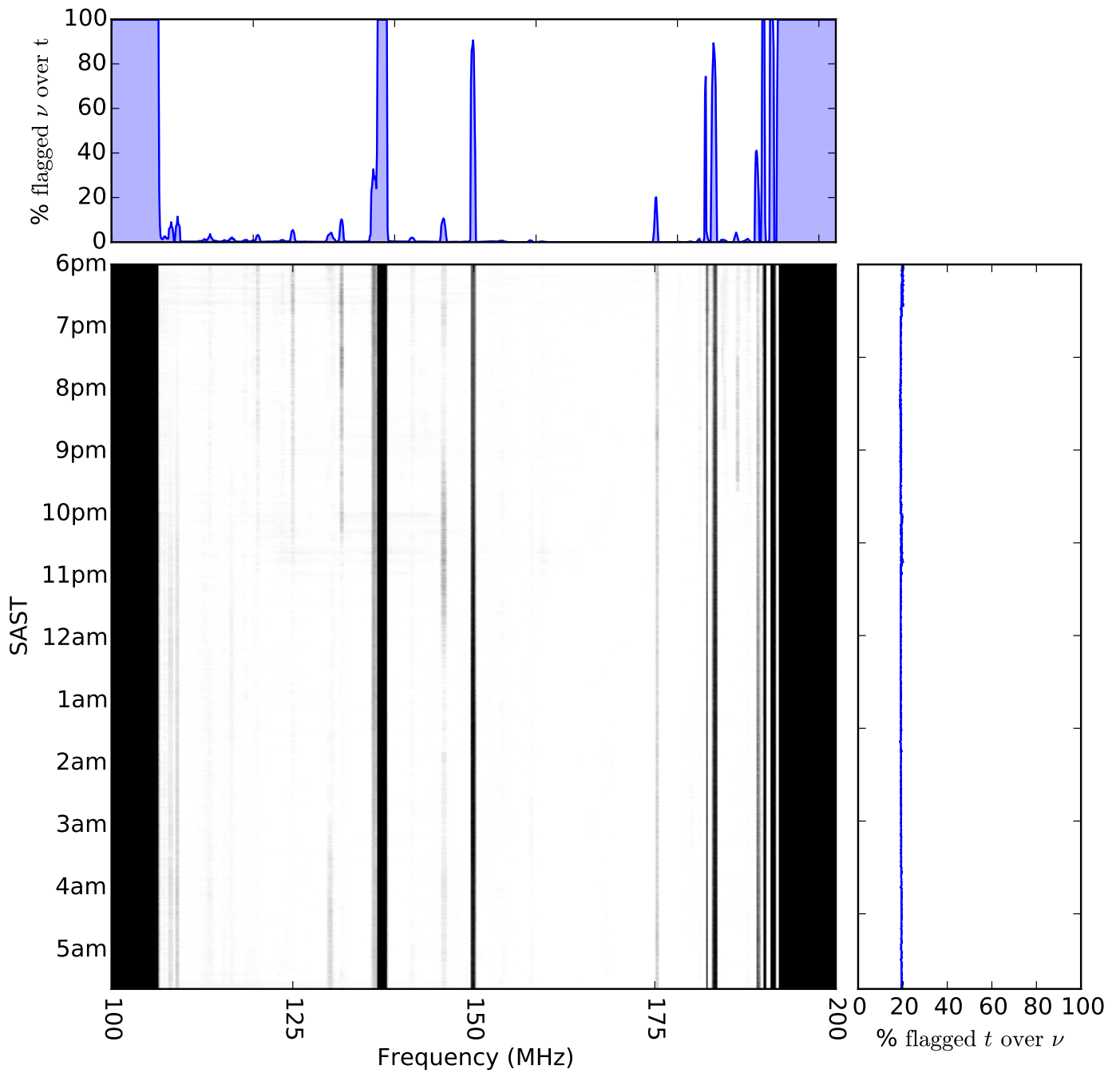
Figure 1: A waterfall plot of RFI flags averaged over 150 days of data. The gridding process is described in the text. Above the waterfall I show the percentage of the season each frequency is flagged, and to the right I show the percentage of frequencies that are flagged per integration.
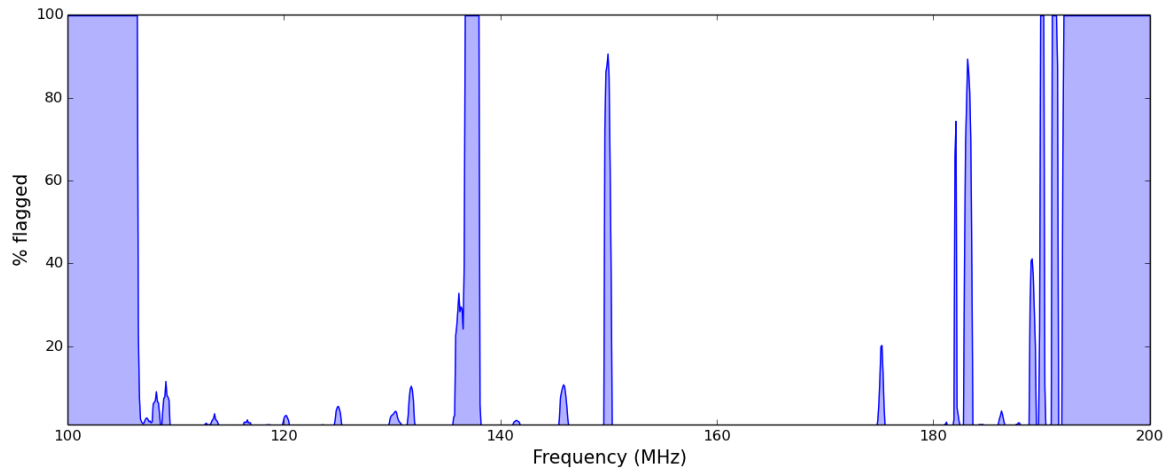
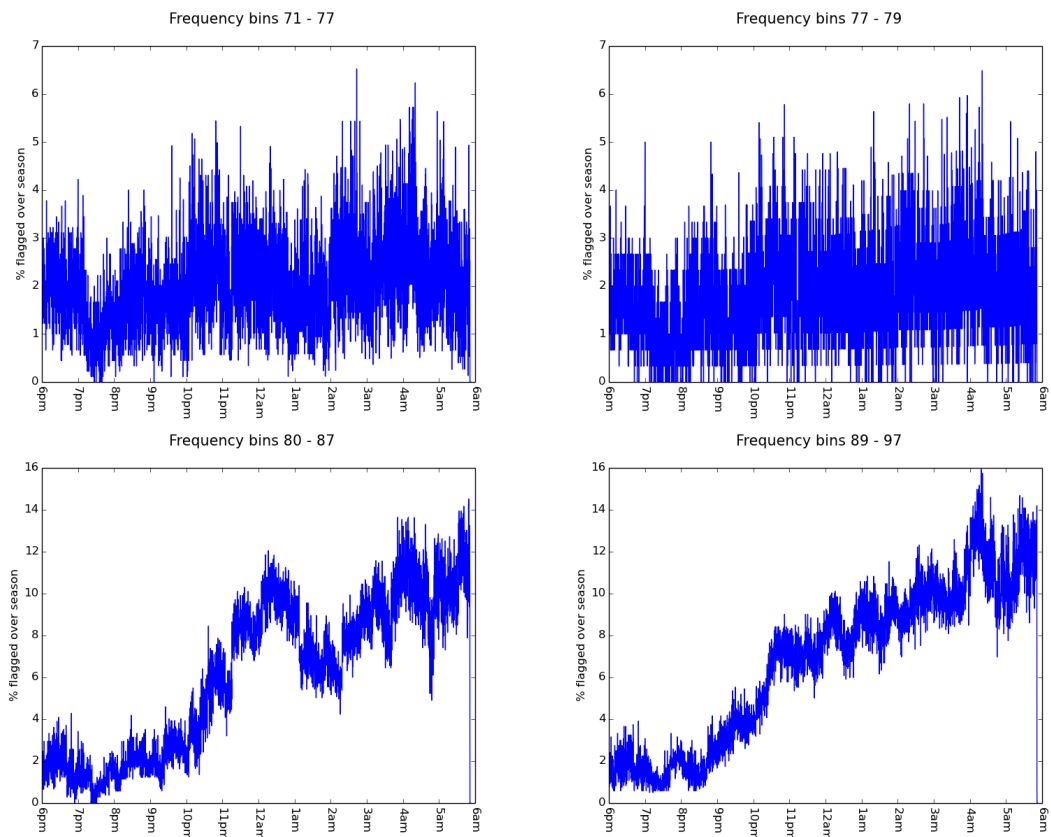Figure 2: The percentage of time that each frequency was flagged over the season.

Figure 3: Possible FM radio contamination in the *Top, left to right*: 107.25±0.25 and 107.55±0.05 MHz bands, and *Bottom, left to right*: 108.1±0.4 and 109±0.4 MHz bands.
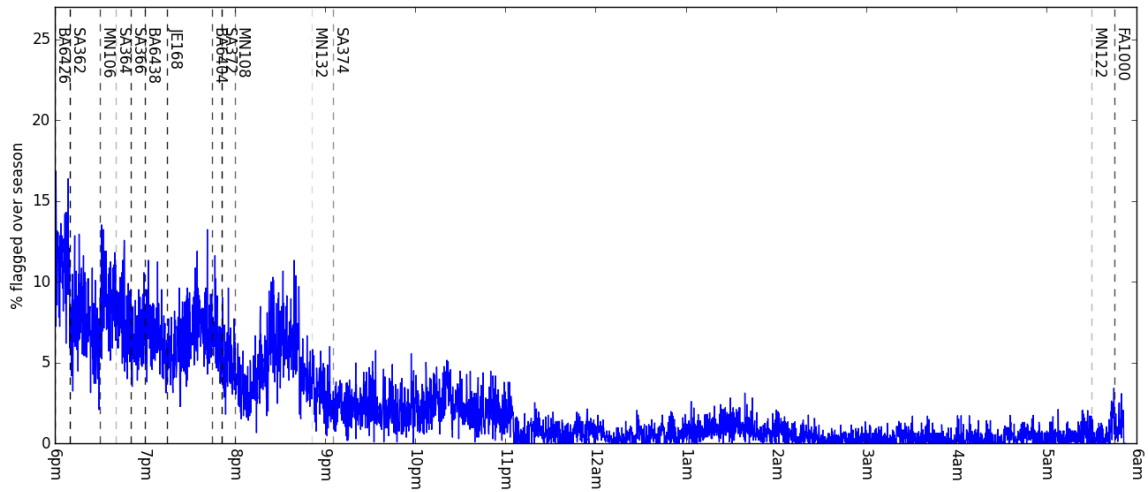
Figure 4: Flights from Cape Town to Johannesburg correspond to RFI in the 120.15±0.35 MHz channels. Vertical dashed lines indicate a flight leaving Cape Town (flights from Johannesburg are roughly concurrent) and the flight code is listed. The transparency of a line is inversely proportional to how many days a week that flight is scheduled for. The flight is 2 to 2.5 hours long – and about 2 hours after the last flight of the day, the flags fall to background level (but notably, not always to zero).

2.1.3 Orbital communications

The *Global Frequency Database*[2] (GFD) lists an (unconfirmed) AM signal at 134.4 MHz with callsign 'RADAR' in North-East Johannesburg[3]. This may be the source of the 136.05±0.45 MHz signal present in 1/3 rd of the season . It could also be 'spillage' of signal from ORBCOMM.

ORBCOMM Inc.'s constellation of 29 LEO communication satellites is a well-known contaminant of the low-frequency sky, dominating over any astronomical signal at 137–138 MHz (although each satellite emits within a 20 kHz band). For this reason we have built flags at 137.35±0.85 MHz into the compression pipeline.

The largest contaminant without flags built-in to the pipeline is communications from the International Space Station (ISS). The 149.75±0.55 MHz transmissions are semi-regular in time; they 'beep'.

Onboard the ISS are HAM radio devices. Some countries have also launched satellites with these onboard, one of the purposes of which is to provide HAM radio operators something in space to communicate with. These devices are licensed to operate at 145.2 and 145.8 MHz, and SATFA lists the whole 144-146 MHz band as reserved for 'Amateur–Satellite' communications. We detect RFI at 145.85±0.45 MHz, although strong signal across ∼10% of the season that occurs 9pm-1am argues against human operation.

2.1.4 Mobile phones and VHF TV

A weak RFI signal at 141.45±0.35 MHz is within the 'mobile 1 BTX' and aeronautical mobile band in SATFA, but other than this single listing I do not have a strong case for the signal.

VHF TV is broadcast over specifically-spaced video and audio frequencies. The strong signals at 183.2±0.5 MHz and 189.15±0.35 MHz have almost identical gradients for the percentage of flagging as a function of time of night. These frequencies correspond exactly to Channel 5 of South African System I 625-line VHF TV signals for video and audio transmission, respectively. Similarly, the weaker signals at 175.15±0.35 and 181.15±0.15 MHz correspond to Channel 4's video and audio transmission, respectively, but they do not share the same temporal properties.

2.1.5 Unidentified sources

There are 5 RFI frequencies in the averaged data that I cannot identify the sources of: weak emissions (flagged < 5% of the season) at 114.05±0.85, 116.55±0.35, 130.25±0.55 and 186.25±0.35 MHz, and one strong emission at 182.15±0.35 MHz. The variation of each source with time is shown in Figure 5. The 186.25±0.35 MHz has a sharp turn-off around 9:45pm each night, suggesting that it is some kind of automated device.

---

[2] http://qrg.globaltuners.com/
[3] There is an active HAM radio operator/survivalist in South Africa with a hobby for what he refers to as 'rapid deployment of amateur radio stations (RaDAR)', which involves him training people to set up radio broadcasting stations in remote locations. He might be the origin of this callsign, and causing a serious detriment to radio astronomy in the Karoo. It could be worth getting in touch with him and letting him know that he may be interfering with our science. See https://zs6bne.wordpress.com/.
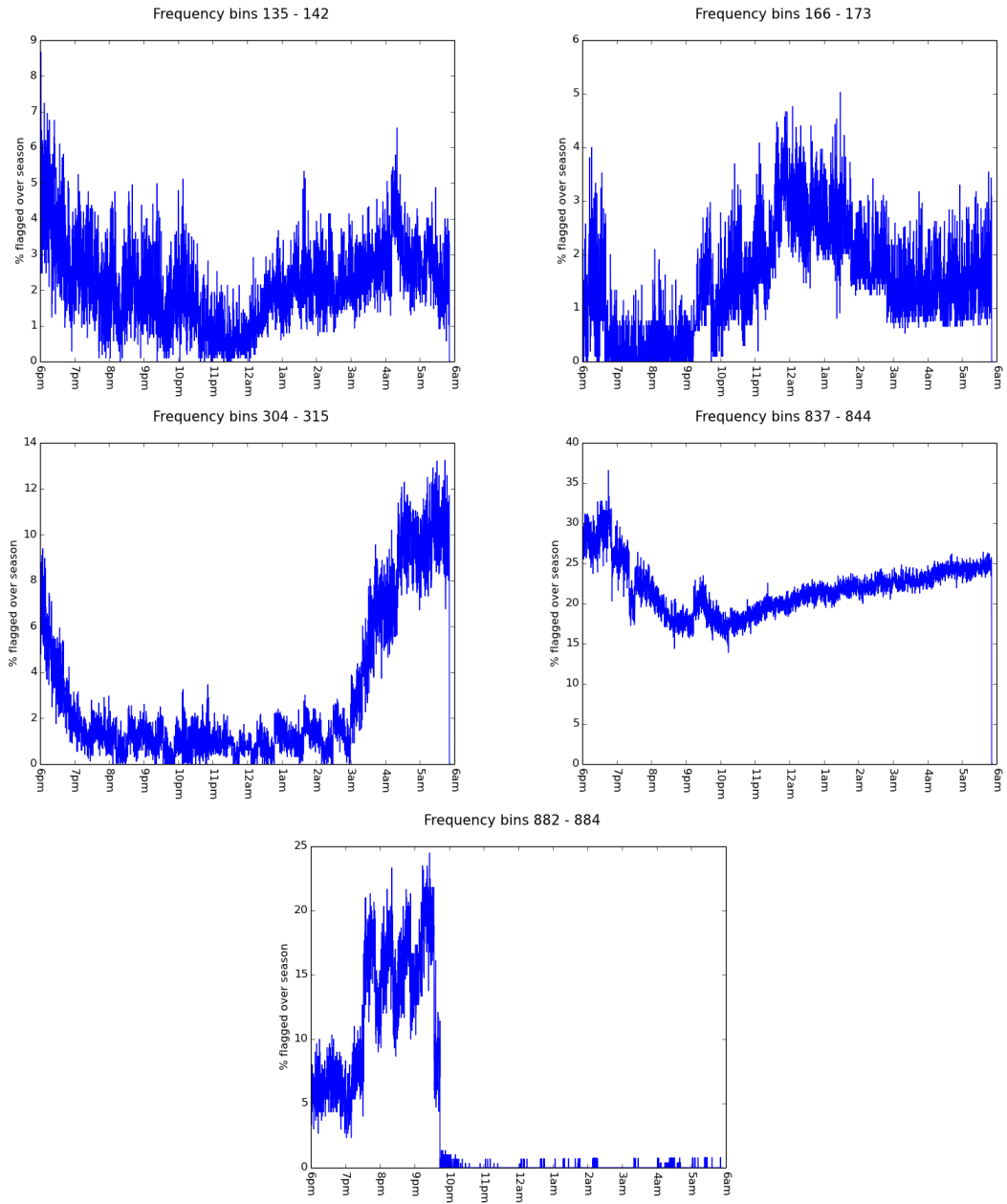
Figure 5: The temporal profile of the 5 RFI frequencies with unidentified causes. *Top, left to right:* 114.05±0.85 and 116.55±0.35 MHz. *Middle, left to right:* 130.25±0.55 and 182.15±0.35 MHz. *Bottom:* 186.25±0.35 MHz. The 182.15±0.35 MHz frequency is flagged a large amount of the time, making it our most-offending unidentified source.
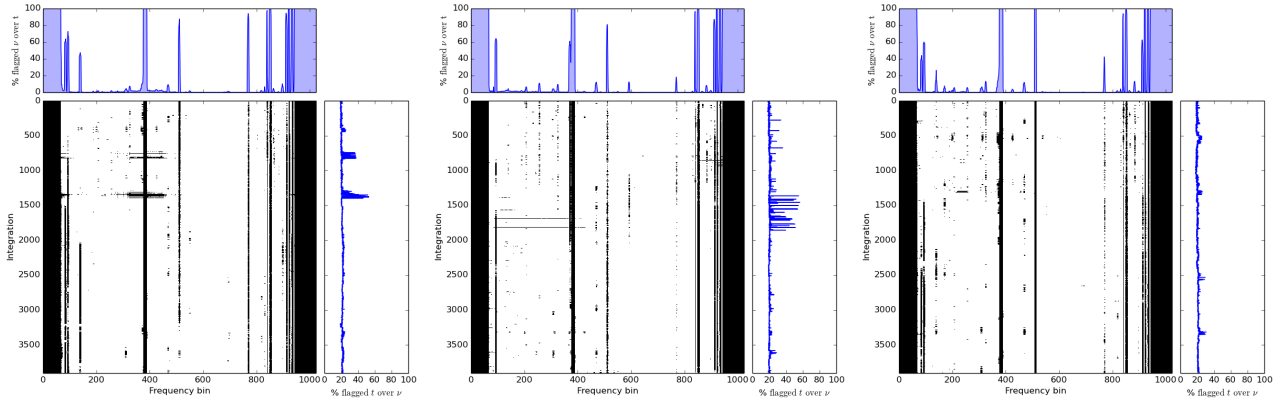
Figure 6: *Left to right:* waterfalls of flags for nights 2456732, 2456958 and 2457038. These three nights, along with 2456965, are >20.2% flagged; > 2σ above the average flagging amount per night.
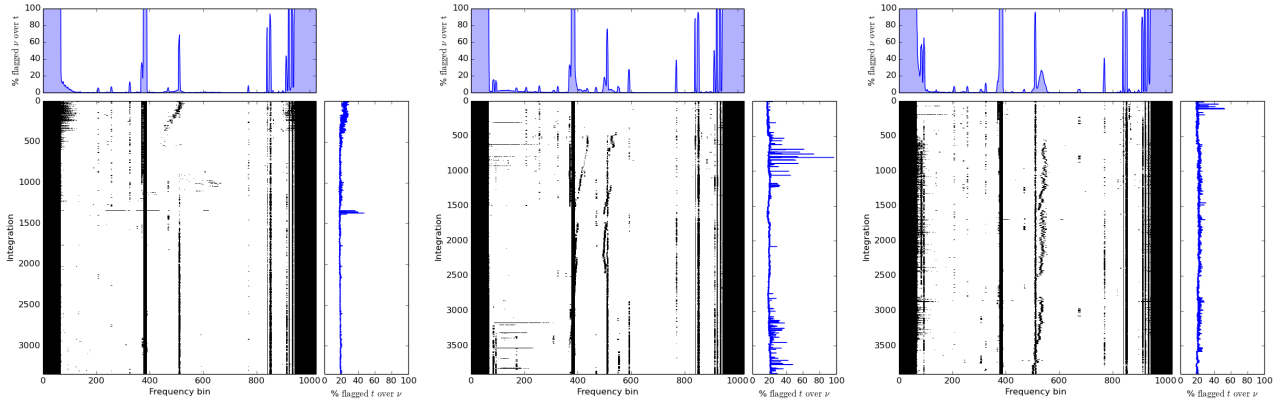

Figure 7: *Left to right:* waterfalls of flags for nights 2456898, 2456924 and 2456965. These three nights exhibit a strange behaviour of RFI that changes in frequency and time. JD 2456965 is by far the worst, and during this night as well as 2456898, we see a broadband 'comb' of flagged frequencies near the band edges

3   Individual Properties

The previous section addresses season-long problems with RFI. In this section I discuss the properties of individual nights.

Using the flags per night, I was able to take a look at the total number of flags as a percentage of the waterfall (i.e. $N_{\text{flags}}/(3920 \times 1024)$). The average flagging per night is 19.2±0.5%, which is dominated by the permanent flagging of ORBCOMM and band edges. Four nights deviate from the average by a $\geq 2\sigma$ excess: JDs 2456965, 2456732, 2456958 and 2457038. Their flag waterfalls are shown in Figure 6 (2456732, 2456958 and 2457038) and Figure 7 (2456965). While the strange nature of night 2456965 is discussed below, the three others follow the pattern of having strong contamination from FM and aircraft communication bands, but also have broadband 'pulses' up to about 20 minutes in length. The source of these broadband pulses is not well understood, although it is clear that ORBCOMM tends to spill outside of its allocated band on occasion.

JD 2456965 is easily the worst offender, and it exhibits a strange signal that wanders in frequency and time close to the ISS band. An event of note on this date (23rd November 2014) was a Soyuz FG launch that docked with the ISS – could this signal be a signature of their transmissions[4]? Similar signals are seen on 2456898 (28th August 2014; although only at the beginning of the night) and 2456924 (23rd September 2014). There is no listed orbital or suborbital activity for 2456898. There was an US ICBM test off of the coast of Virginia on 2456924, but this is probably not the cause of the RFI. The flag waterfalls for these nights are shown in Figure 7.

Another property that the flag waterfalls in Figures 6 and 7 highlight is the presence of broadband RFI signals, typically present at frequencies lower than the ORBCOMM band. However, while we flag at the low-end of the band (which has higher noise levels to begin with), is is likely that such broadband pulses dominate the band at those times, and that we fail to flag all of the integrations. Our flagging script *xrfi_simple.py* does contain a thresholding option for flagging the entire integration given some arbitrary number of frequencies flagged during that integration: some experimentation will be required to decide if that threshold should change.

---

[4]The internet also suggests... less plausible explanations: `https://www.youtube.com/watch?t=11&v=VtZx8iPO4zs`.

4 Conclusions

What does all this mean for HERA? I feel like I have left more questions than answers in this memo. However, there are some actions that can be taken:

- Steps to reduce and ideally eliminate the VHF TV transmissions in the area would be very helpful, since these are clearly interfering with our measurements.

- The ISS 149.75±0.55 MHz band should be permanently flagged within the compression pipeline.

- Pursuing re-routing of flight paths will not do much to help: we see aircraft signals for the duration of their flight, not just when they're over the Karoo.

- A lower threshold for identifying broadband RFI should be experimented with (discussed a little more below).

I should note that after compression, our frequency resolution decreases dramatically. If a low-resolution frequency bin contains a frequency that has been flagged at the high-resolution flagging stage, the low-resolution bin is also flagged. This caveat means that analyses that work with HERA data in a compressed format, as opposed to a higher frequency resolution stage, will have different views of the amount of RFI in a given visibility.

A new, lower-frequency feed is currently under development by the HERA analog group. This would nominally allow measurements to be taken in the range 50–250 MHz, allowing science observations of the Dark Ages and the post-reionization Universe. It should be noted that at the lowAOflest frequencies FM radio will be a constant harassment to these measurements. At the higher frequencies, VHF TV will be the primary contaminant, but should be much easier to remove as it is both narrow-band and, as noted above, shut-down-able.

There are future directions to pursue in RFI flagging and analysis. An RFI-flagger that could be informed of the expected frequency structure of visibilities (there are higher noise levels on the low-frequency end of the band) would limit false-positive flags. An advanced example of such an algorithm is AOFLAGGER [5], but given the relatively low amount of RFI on out site (as opposed to Western Europe), this particular solution may not be necessary. Either way, before building a more adaptive flagger into the HERA reduction pipeline, we first need to learn more about the RFI we have already identified. It is outside the scope of this memo, but a power spectrum (or simply a histogram) of flagged visibilities would be an extremely useful diagnostic tool. How low-power does a signal need to be to not be flagged? It is not absurd to think that we could be missing a component of ambient RFI during flagging. How broad is the spread of RFI energies? Is there a more intelligent way to combine flagged and unflagged data during compression? These are still open questions that I have not yet had the time to try to answer, but would make excellent summer-student projects (for example).

References

[1] Z. S. Ali, A. R. Parsons, H. Zheng, J. C. Pober, A. Liu, J. E. Aguirre, R. F. Bradley, G. Bernardi, C. L. Carilli, C. Cheng, D. R. DeBoer, M. R. Dexter, J. Grobbelaar, J. Horrell, D. C. Jacobs, P. Klima, D. H. E. MacMahon, M. Maree, D. F. Moore, N. Razavi, I. I. Stefan, W. P. Walbrugh, and A. Walker. PAPER-64 Constraints on Reionization: The 21cm Power Spectrum at z=8.4. *ArXiv e-prints*, February 2015.

[2] South African Civil Aviation Authority. Flight calibration dates of vor's in the republic of south africa. Technical Report 25.6, Aeronautical Information Circular, `http://209.203.9.244/resource%20center/AIC'S/25.6.pdf`, June 2008.

[3] S. R. Furlanetto, S. P. Oh, and F. H. Briggs. Cosmology at low frequencies: The 21 cm transition and the high-redshift Universe. *Physics Review*, 433:181–301, October 2006.

[4] D. C. Jacobs, J. C. Pober, A. R. Parsons, J. E. Aguirre, Z. S. Ali, J. Bowman, R. F. Bradley, C. L. Carilli, D. R. DeBoer, M. R. Dexter, N. E. Gugliucci, P. Klima, A. Liu, D. H. E. MacMahon, J. R. Manley, D. F. Moore, I. I. Stefan, and W. P. Walbrugh. Multiredshift Limits on the 21 cm Power Spectrum from PAPER. *ApJ*, 801:51, March 2015.

[5] A. R. Offringa, J. J. van de Gronde, and J. B. T. M. Roerdink. A morphological algorithm for improving radio-frequency interference detection. *A&A*, 539:A95, March 2012.

[6] A. R. Parsons, A. Liu, J. E. Aguirre, Z. S. Ali, R. F. Bradley, C. L. Carilli, D. R. DeBoer, M. R. Dexter, N. E. Gugliucci, D. C. Jacobs, P. Klima, D. H. E. MacMahon, J. R. Manley, D. F. Moore, J. C. Pober, I. I. Stefan, and W. P. Walbrugh. New Limits on 21 cm Epoch of Reionization from PAPER-32 Consistent with an X-Ray Heated Intergalactic Medium at z = 7.7. *ApJ*, 788:106, June 2014.

[7] Staatskoerant. South african table of frequency allocations. Technical Report 31264, ICASA, `http://thornton.co.za/resources/31264_890_complete-1.pdf`, July 2008.