Flagging VLBI data

Slide deck

JIVE VLBI school 2025

Jack F. Radcliffe & Célestin Herbé-George

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

EUROPEAN vlbi NETWORK

JIVE Joint Institute for VLBI ERIC

ACME Astrophysics Centre for Multimessenger studies in Europe

Funded by the European Union

Note that this is heavily based upon the Development with Africa with Radio Astronomy content.

Reminder: sources of errors i.e. what attacks our data

Atmosphere
- Ionosphere
- Troposphere
- Water vapour

Antenna / feed
- System temperature
- Primary beam
- Pointing
- Antenna location

LNA / conversion chain
- Clock
- Gain, phase, delay
- Frequency response

Digitiser / Correlator
- Auto-leveling
- Baseline errors

Radio Frequency Interference (RFI)

Data Inspection and Flagging

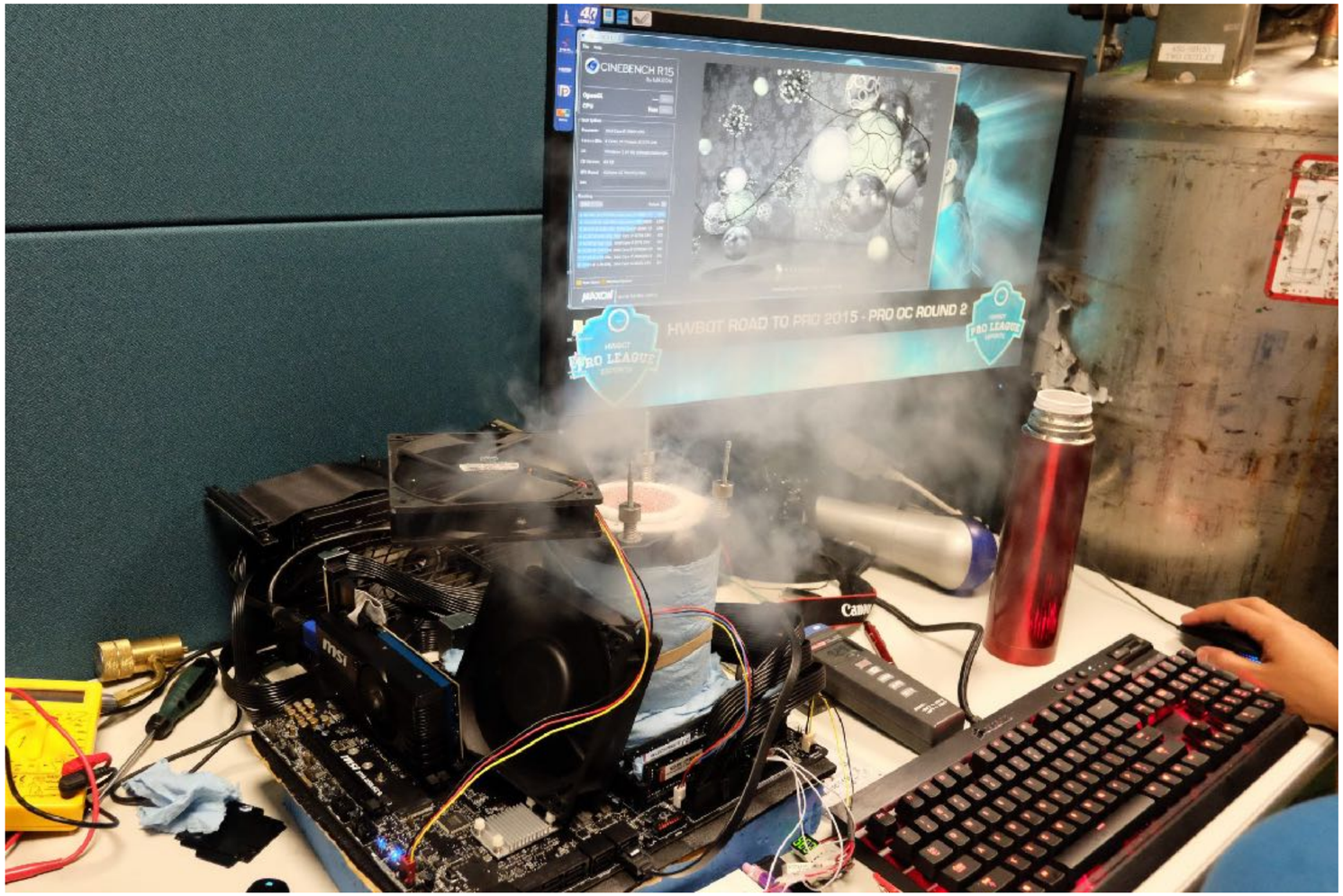
- Broken Elements

Why data editing?

- Broken Elements
- Antenna Shadowing

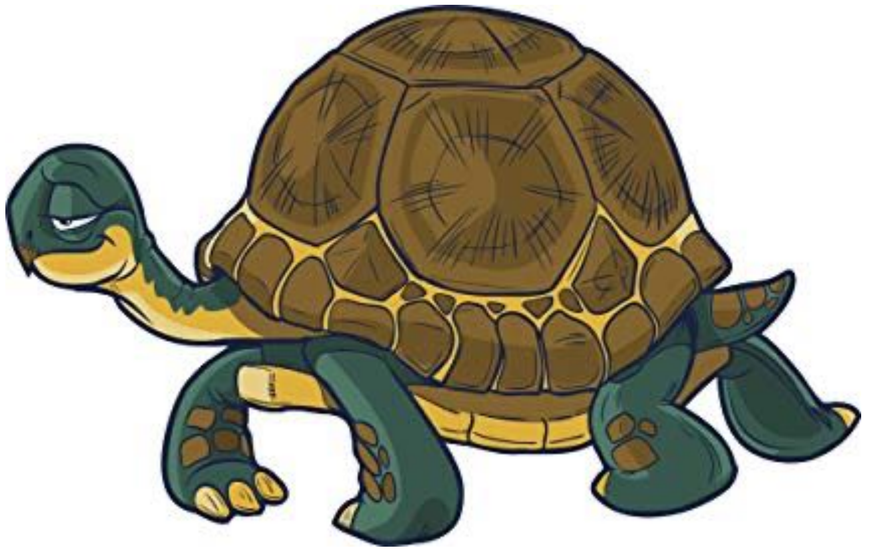
Why data editing?

- Broken Elements
- Antenna Shadowing
- Correlator malfunctions



Why data editing?

- Broken Elements
- Antenna Shadowing
- Correlator malfunctions
- Initial pointing delay



Why data editing?

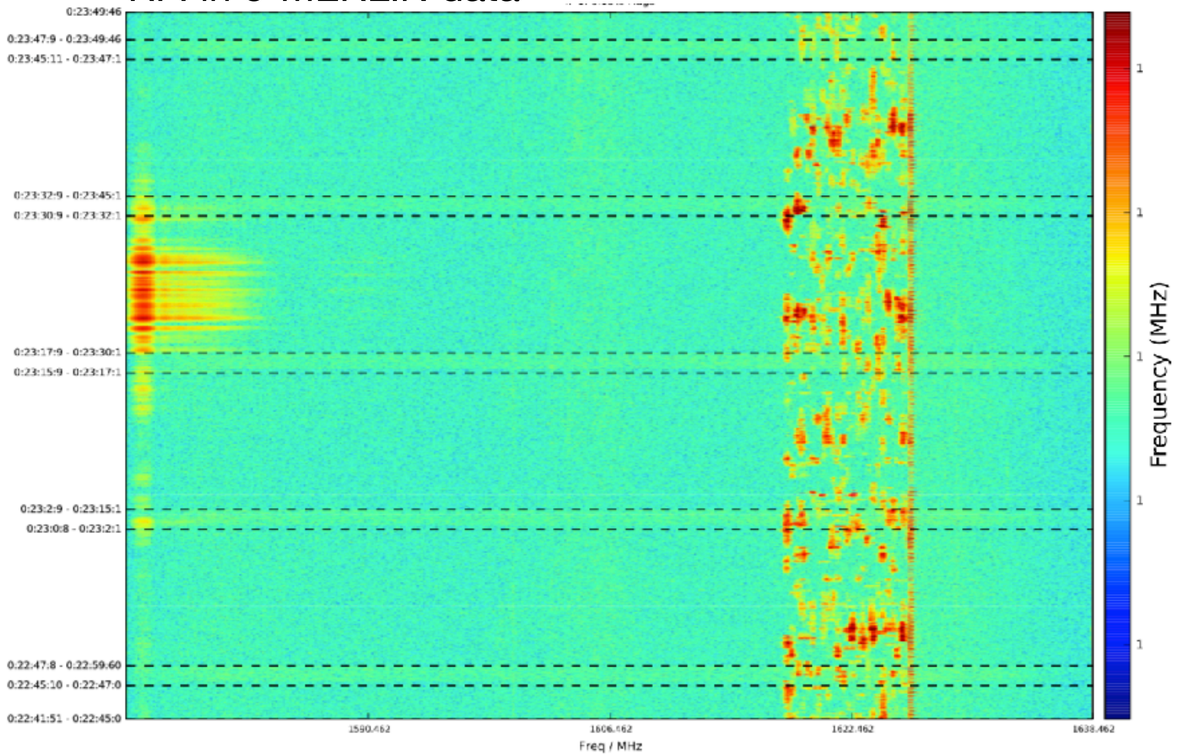
- Broken Elements
- Antenna Shadowing
- Correlator malfunctions
- Initial pointing delay

- Bandpass issues
- Low elevation for some antennas
- Correlated noise on some baselines

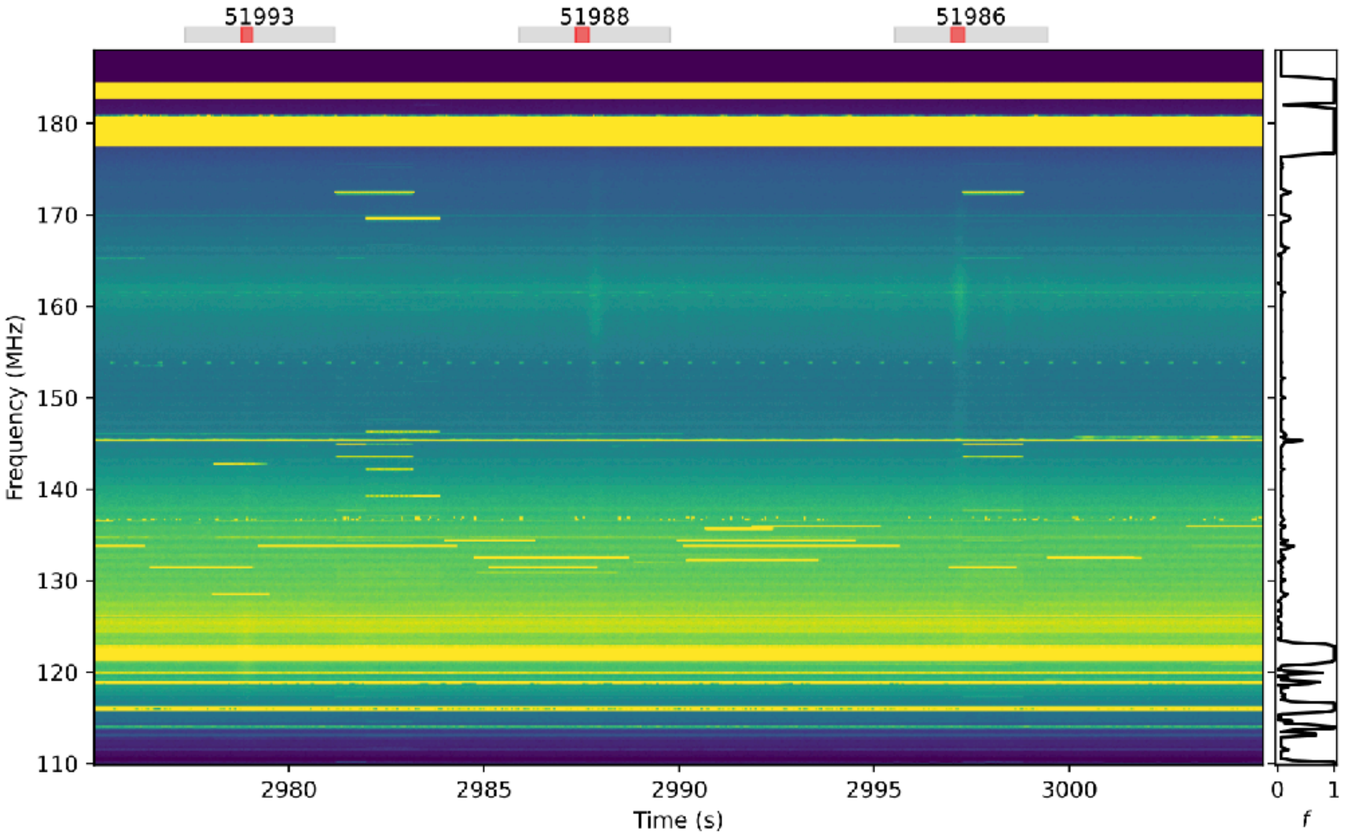
Radio frequency interference (RFI) – a radio astronomers worst enemy (most of the time)

- Your mobile phone is brighter than any radio source in the sky…
- Causes errors, so this needs to be removed from our data!

RFI in e-MERLIN data

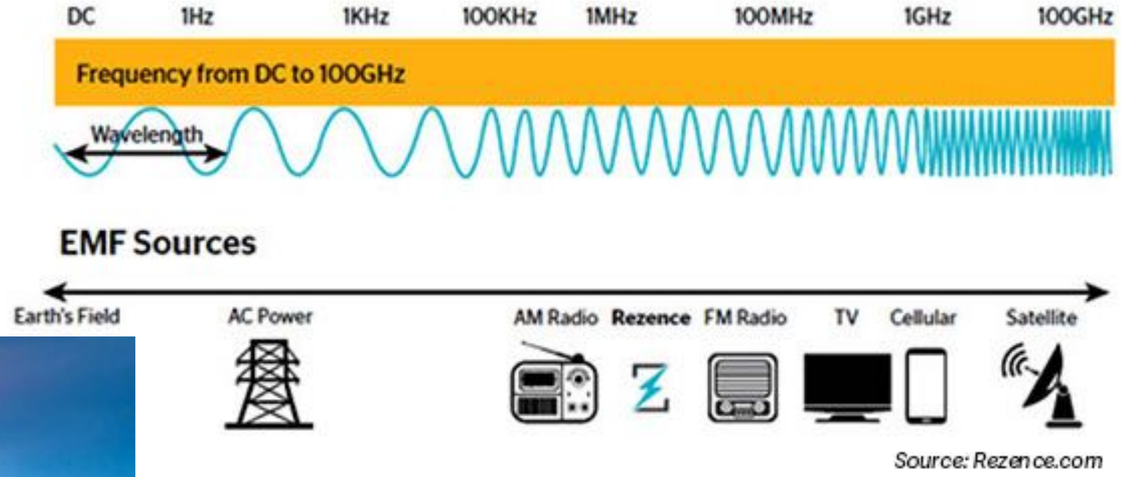


Starlink in LOFAR data (di Vruno+23)



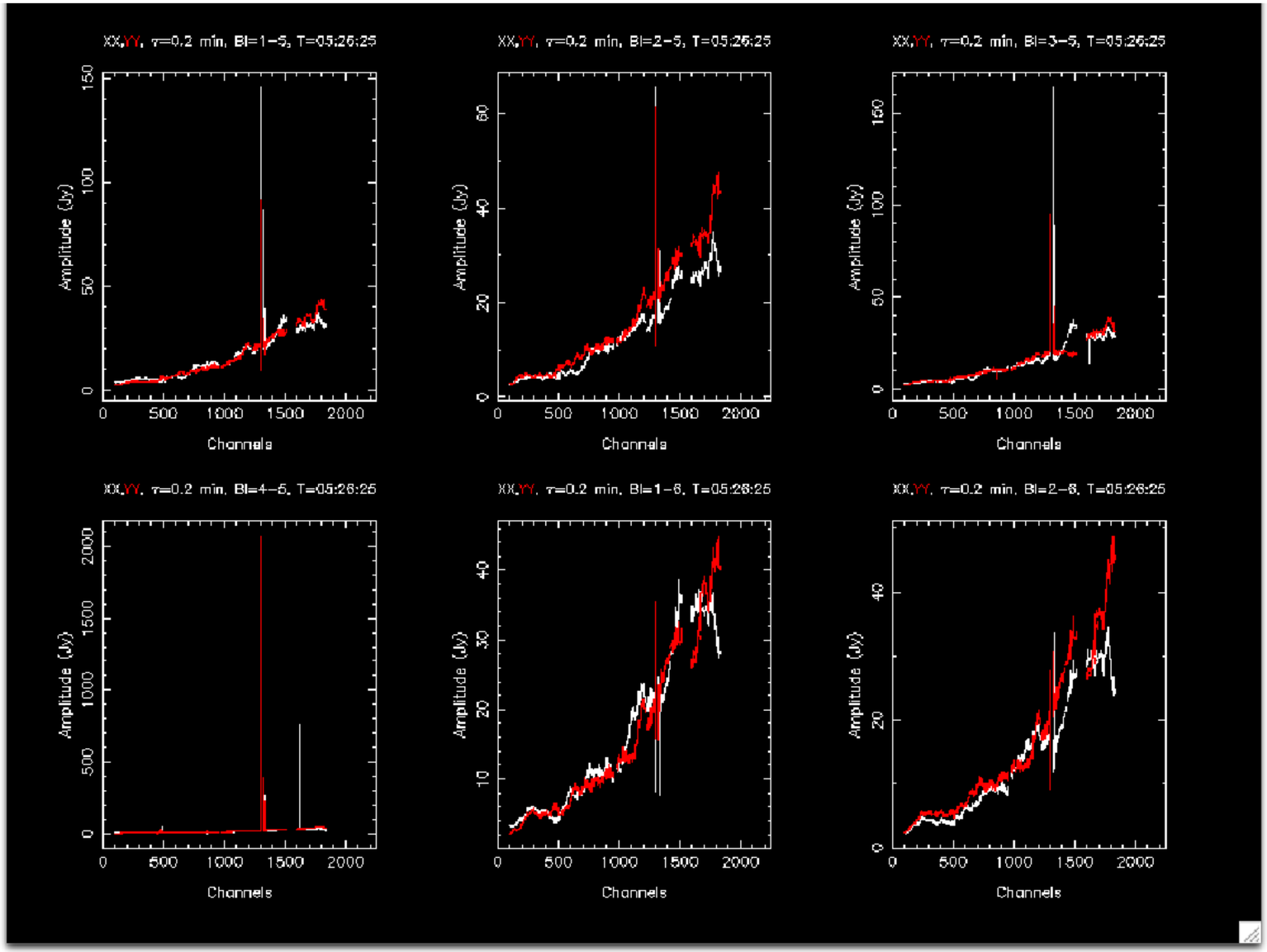
Radio frequency interference (RFI) – a radio astronomers worst enemy (most of the time)

- Discrete bands

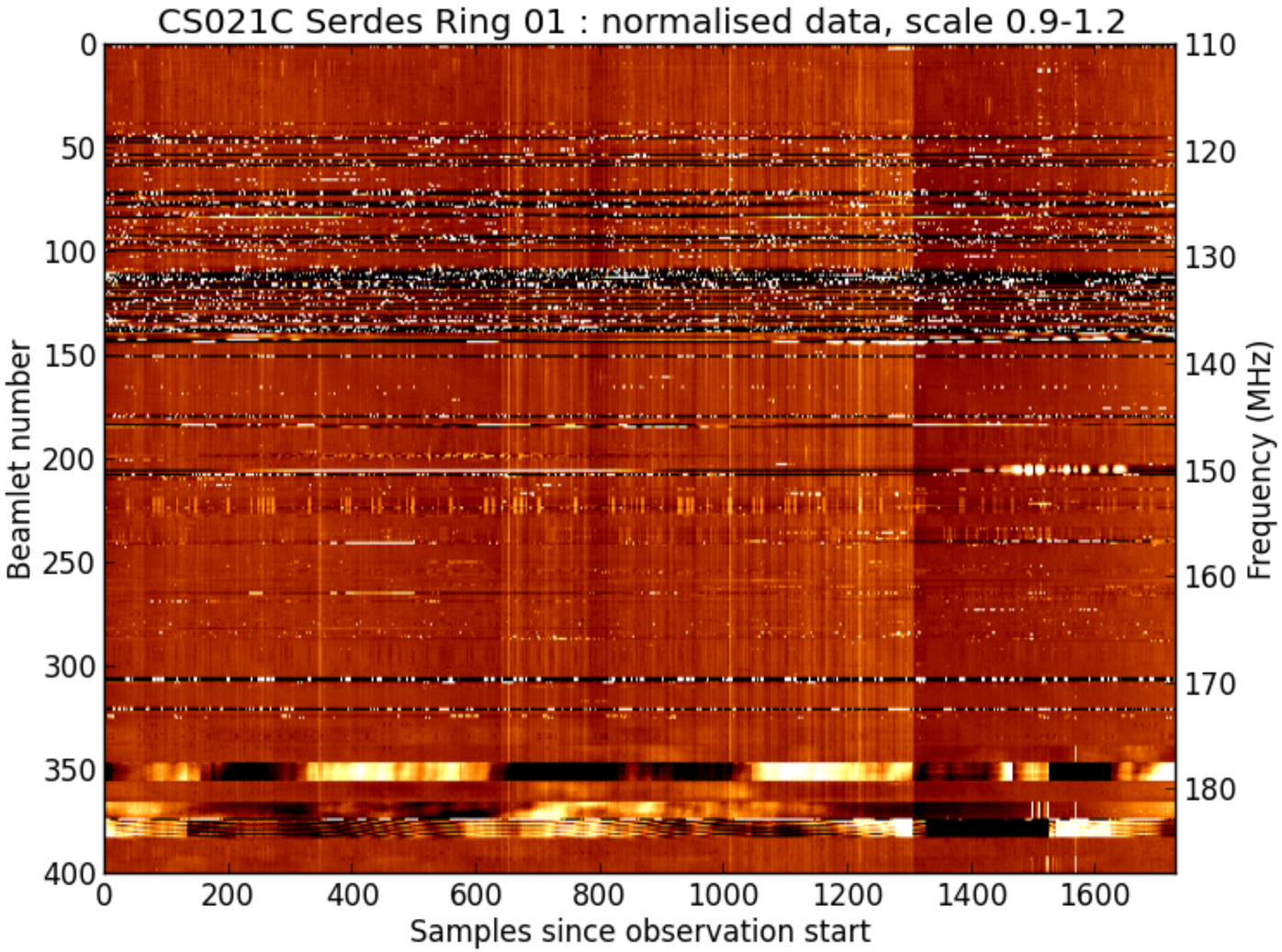


- Broadband

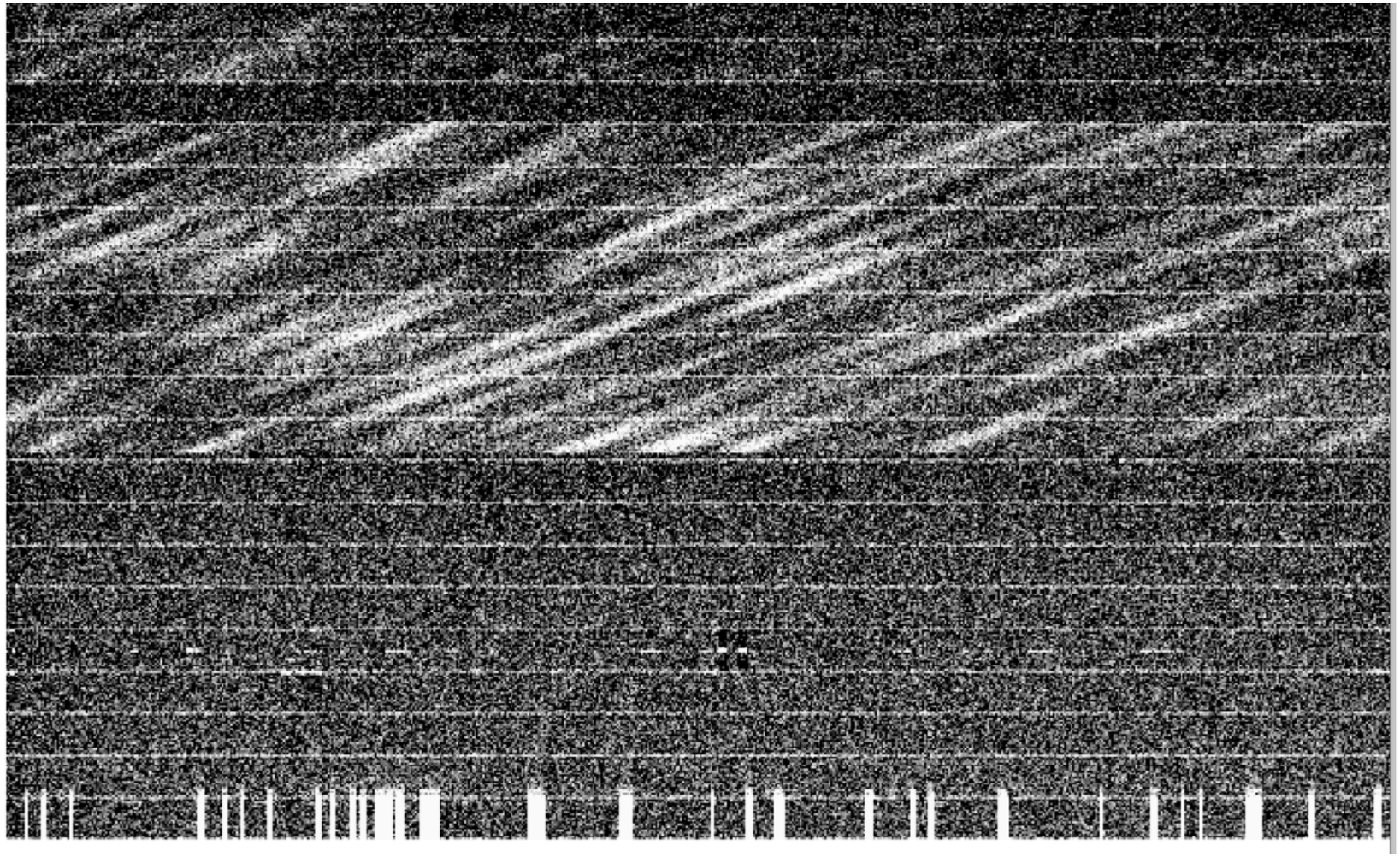
What can RFI look like?



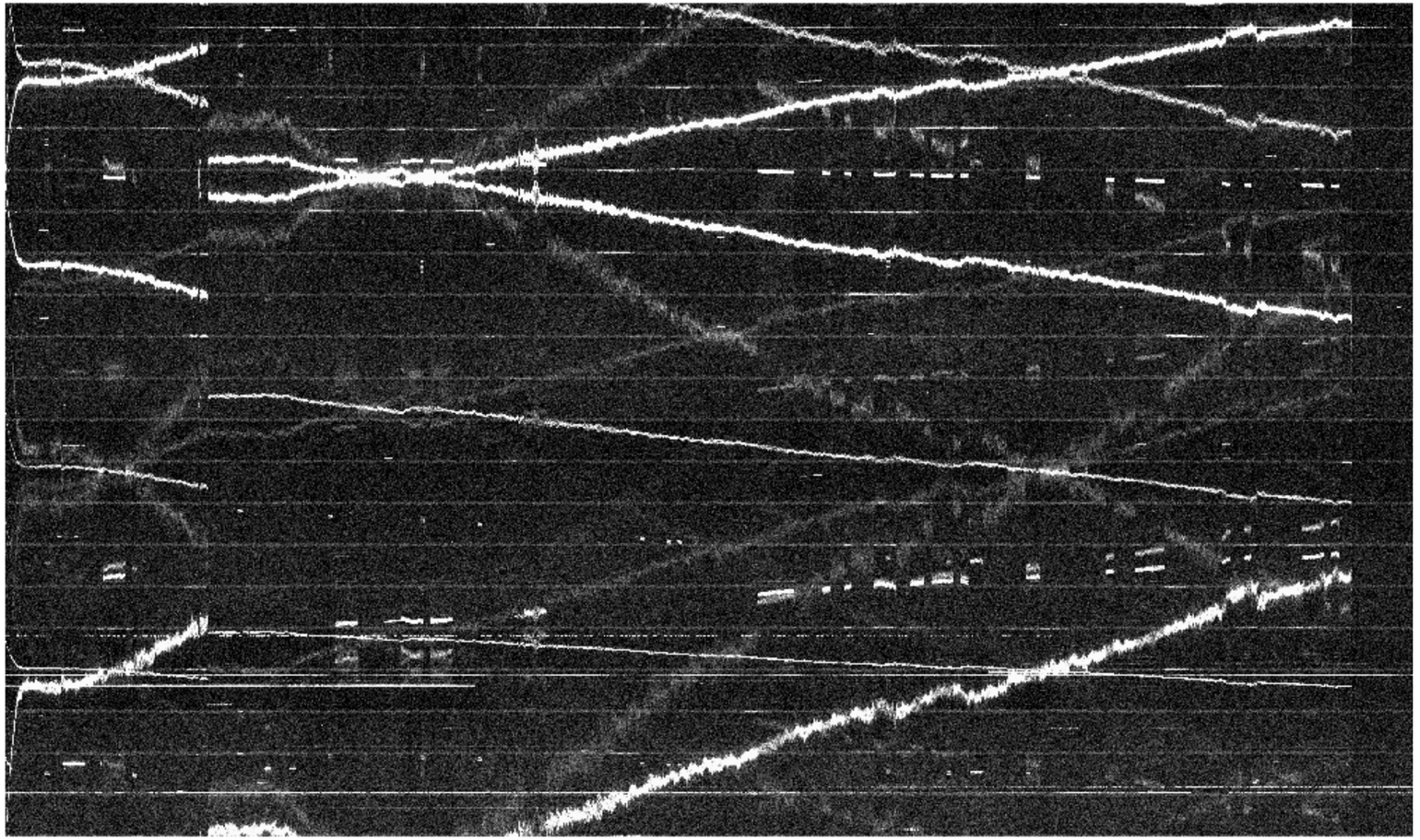
What can RFI look like?



What can RFI look like?

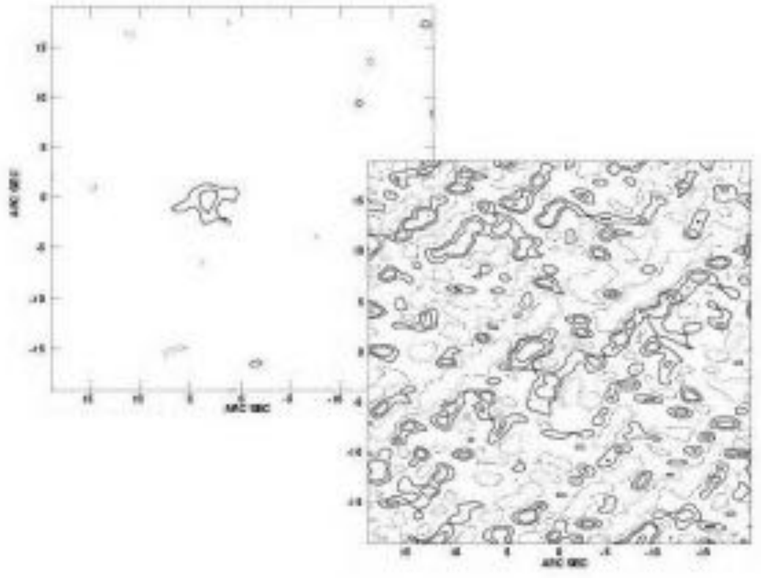


What can RFI look like?



What if you do not flag?

- Probably will not be able to derive sensible calibration solutions (e.g. delay or phase solutions in particular)
- If you can make an image, it will be noiser and (depending on the brightness of your target), you will not be able to see your target
- Remember RFI is not constant in time. So you have to flag all calibrators and targets seperately.



How can we fix this in one slide…

- Broken elements → **remove antennas**
- Correlator malfunctions → **remove timesteps**
- Shadowing → **remove antennas in time range**
- Initial pointing delay → **remove first timesteps**
- Bandpass issues → **remove channels**

- Low elevation → **remove antennas with low elevation**
- Correlated noise on some baselines (e.g. LOFAR split stations) → **Flag baselines**
- RFI→ **remove antennas, timestep, frequencies or baselines...**

How can we fix this in one slide…

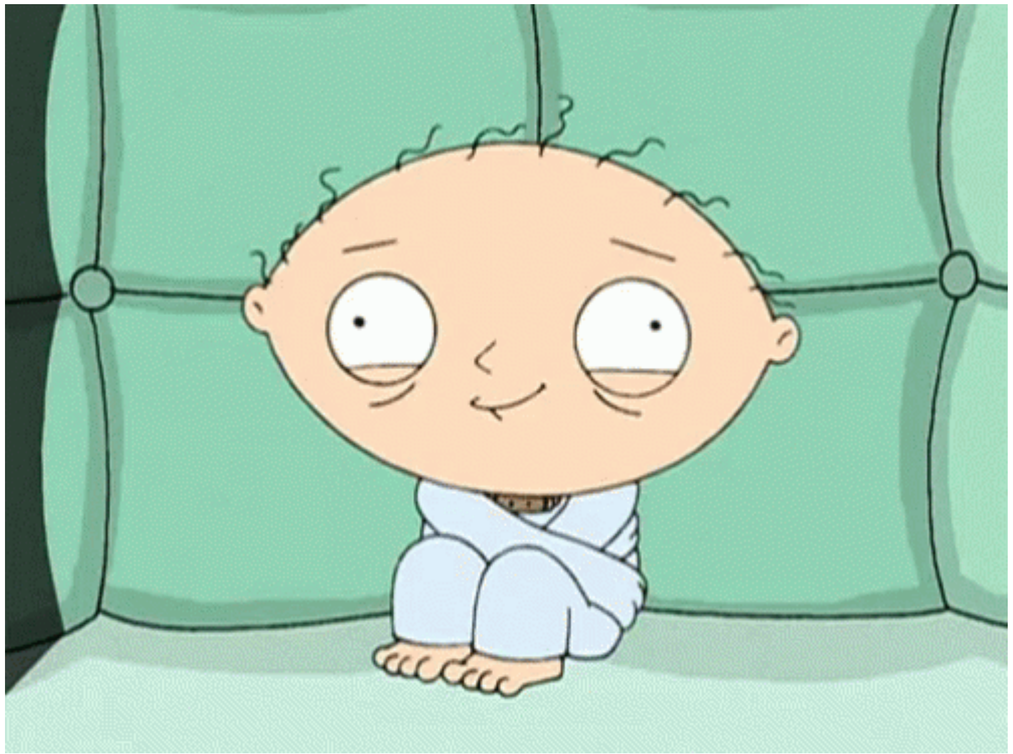
- Broken elements → **remove antennas**
- Correlator malfunctions → **remove timesteps**
- Shadowing → **remove antennas in time range**
- Initial pointing delay → **remove first timesteps**
- Bandpass issues → **remove channels**

- Low elevation → **remove antennas with low elevation**
- Correlated noise on some baselines (e.g. LOFAR split stations) → **Flag baselines**
- RFI→ **remove antennas, timestep, frequencies or baselines...**

Flagging data does not mean deleting

Modern ways to flag can not be manual

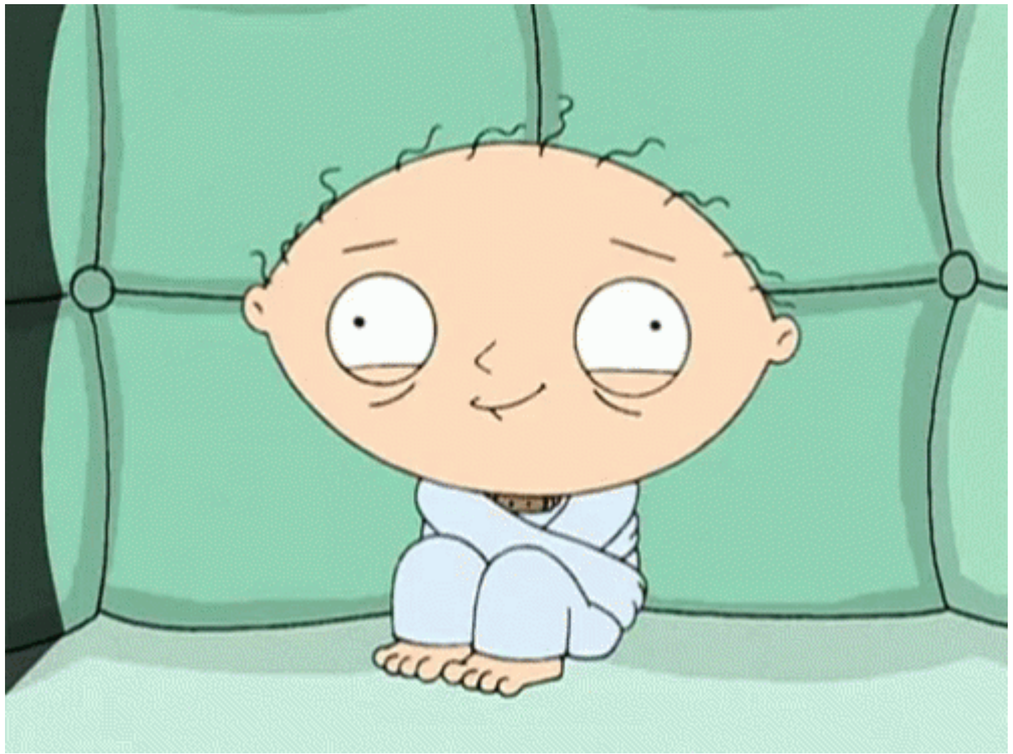
- For example, LOFAR has 44 stations (including international stations). That means there are N(N-1)/2 = **946 baselines**



- Far too many to inspect manually

Modern ways to flag can not be manual

- For example, LOFAR has 44 stations (including international stations). That means there are N(N-1)/2 = **946 baselines**



- Far too many to inspect manually

Other more automated/modern ways to flag

- **Two classes of RFI excision methods:**
  - **Detection: find & throw away affected data**
  - **Filtering or subtracting: estimate RFI contribution and restore affected data**

Other more automated/modern ways to flag

- **Two classes of RFI excision methods:**
  - **Detection: find & throw away affected data**
  - **Filtering or subtracting: estimate RFI contribution and restore affected data**

- Detection methods (“flagging”) commonly used
  - Some specialized pipelines for surveys or instruments

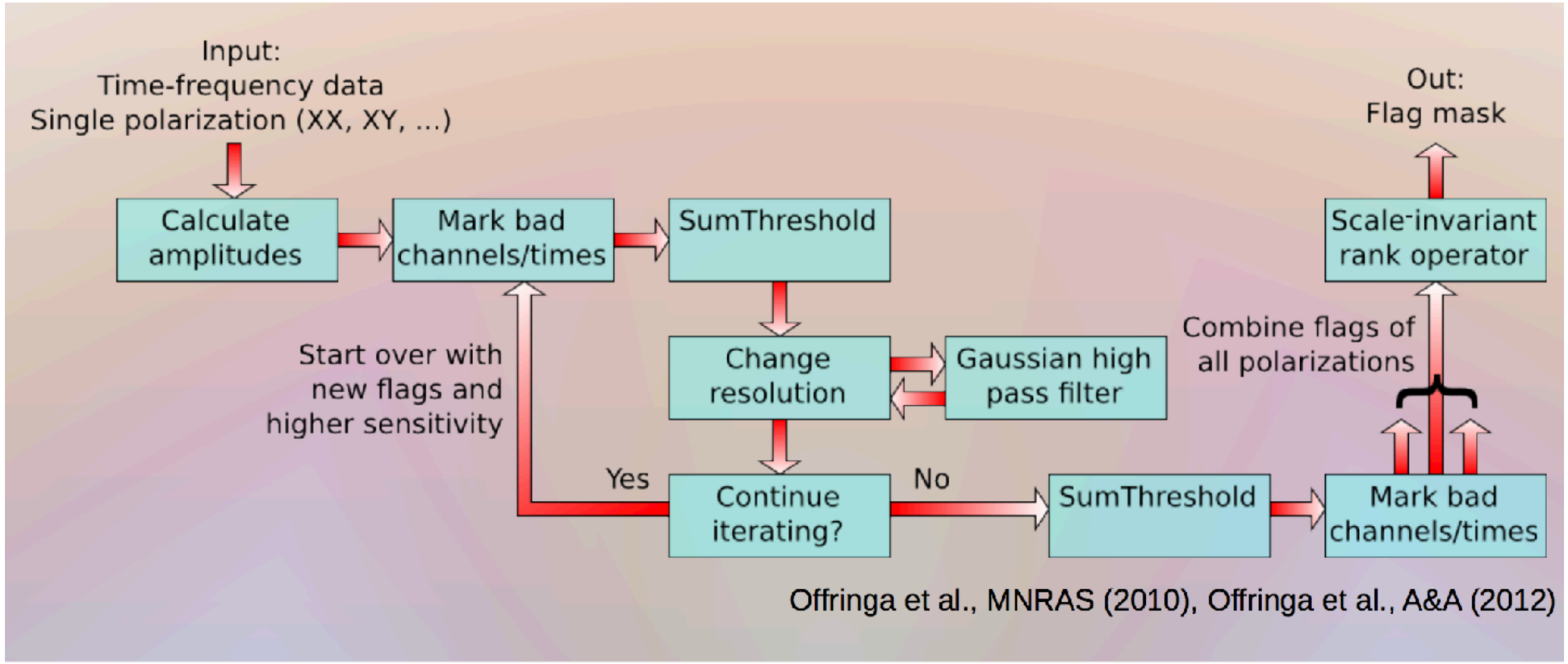
Other more automated/modern ways to flag

- **Two classes of RFI excision methods:**
  - **Detection: find & throw away affected data**
  - **Filtering or subtracting: estimate RFI contribution and restore affected data**

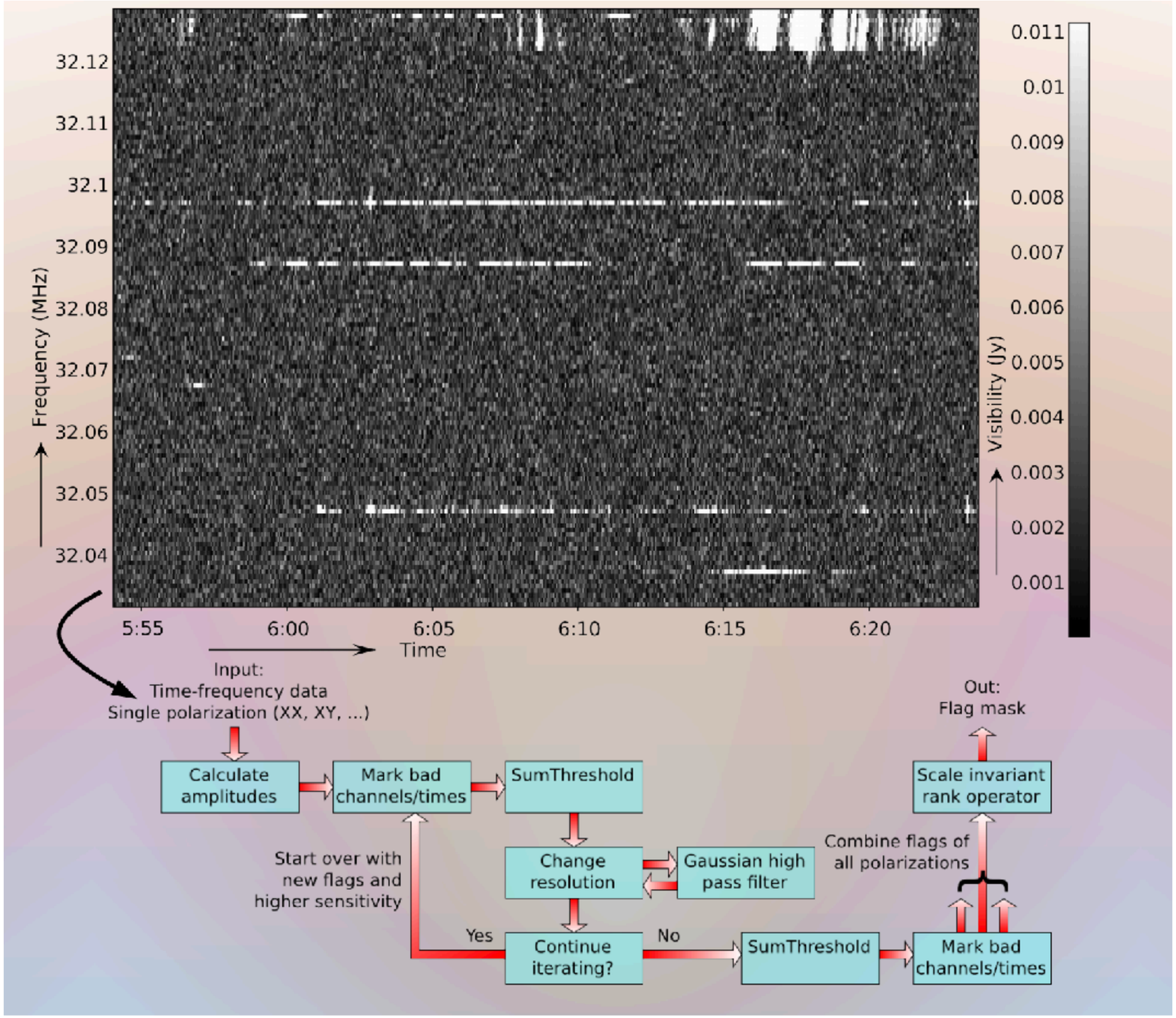
- Detection methods (“flagging”) commonly used
  - Some specialized pipelines for surveys or instruments

- Filtering RFI is harder
  - Resulting data quality is not well understood
  - Requires more resources
  - Lack of full (automated) filtering pipelines

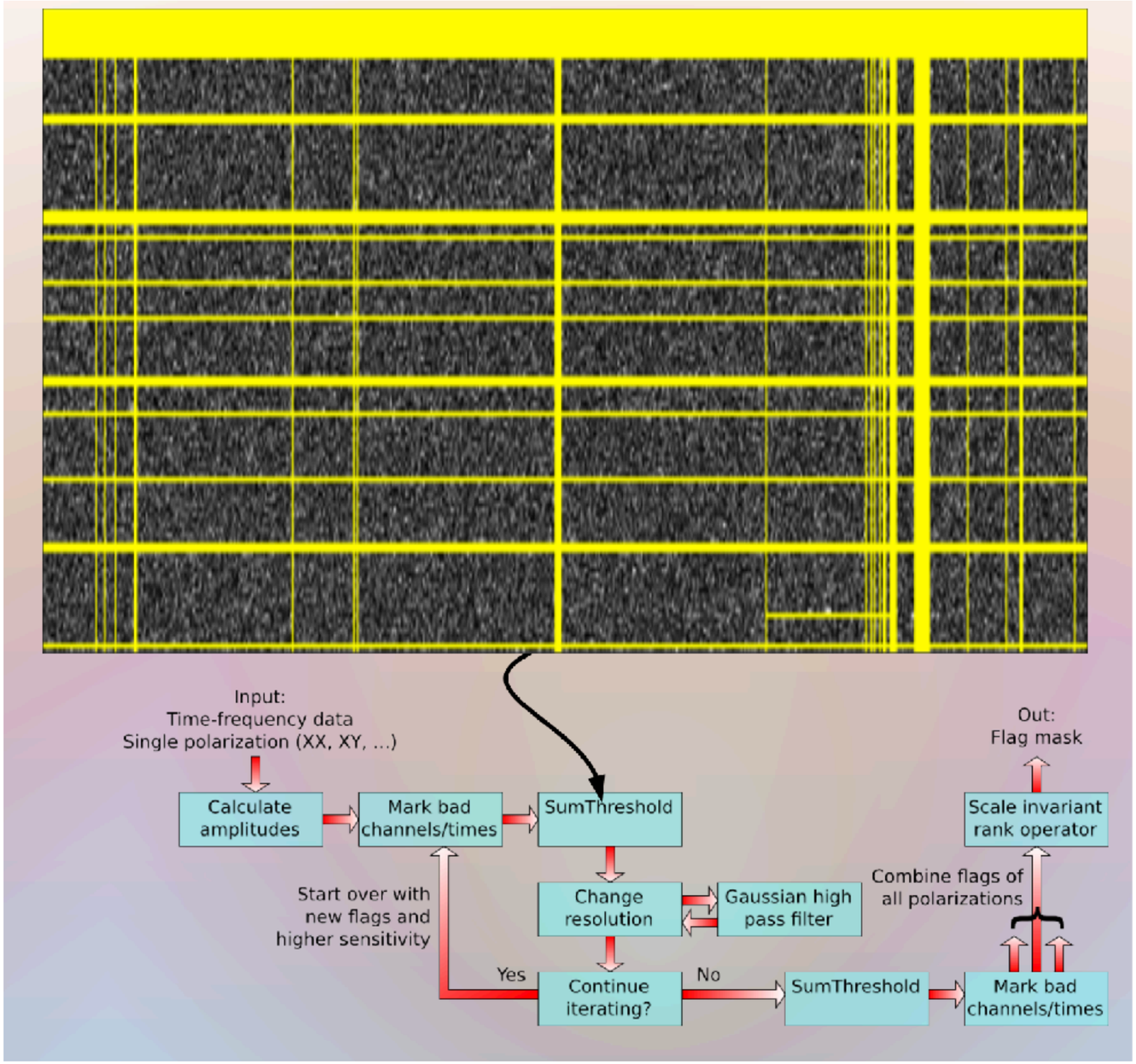
AOFlagger


Input:
Time-frequency data
Single polarization (XX, XY, ...)
Calculate amplitudes
Mark bad channels/times
SumThreshold
Change resolution
Gaussian high pass filter
Start over with new flags and higher sensitivity
Yes
Continue iterating?
No
SumThreshold
Mark bad channels/times
Combine flags of all polarizations
Scale-invariant rank operator
Out:
Flag mask
Offringa et al., MNRAS (2010), Offringa et al., A&A (2012)

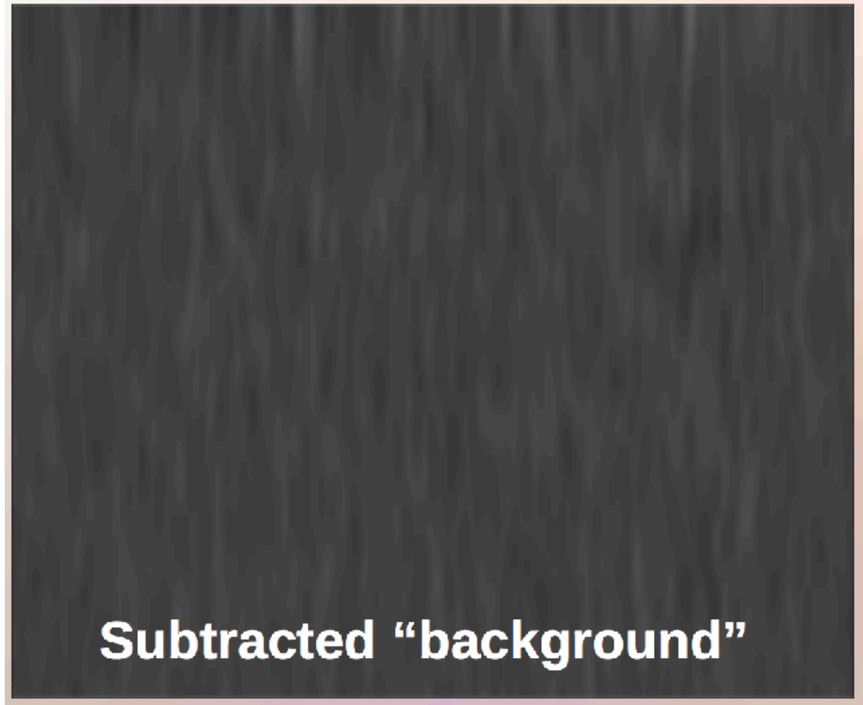
AOFlagger

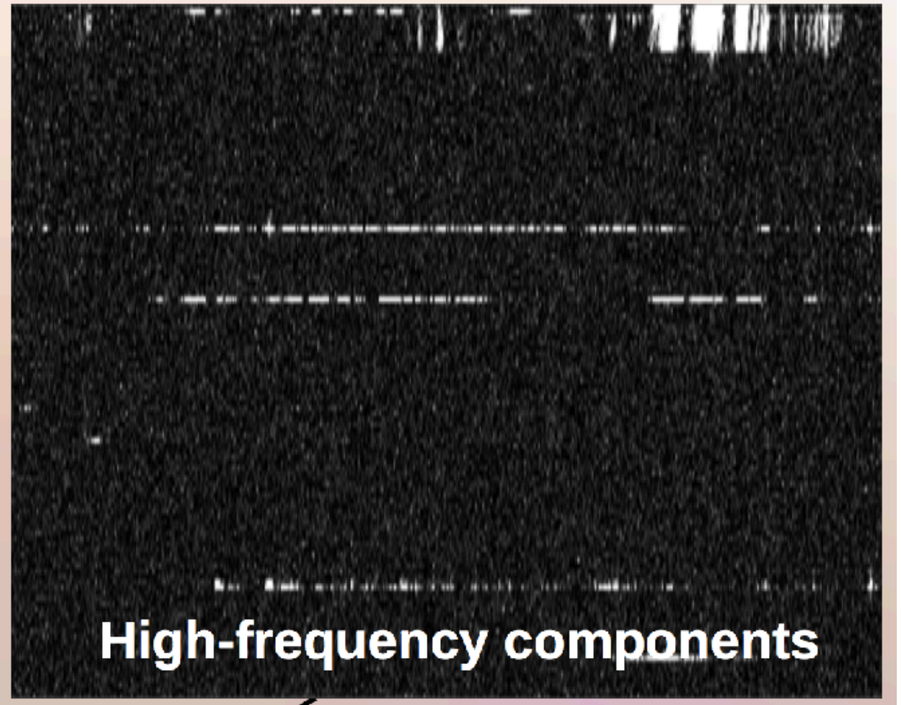

32.12
32.11
32.1
32.09
32.08
32.07
32.06
32.05
32.04
Frequency (MHz)
5:55
6:00
6:05
6:10
6:15
6:20
Time
0.011
0.01
0.009
0.008
0.007
0.006
0.005
0.004
0.003
0.002
0.001
Visibility (Jy)
Input:
Time-frequency data
Single polarization (XX, XY, ...)
Calculate amplitudes
Mark bad channels/times
SumThreshold
Change resolution
Gaussian high pass filter
Start over with new flags and higher sensitivity
Yes
Continue iterating?
No
SumThreshold
Mark bad channels/times
Combine flags of all polarizations
Scale invariant rank operator
Out:
Flag mask

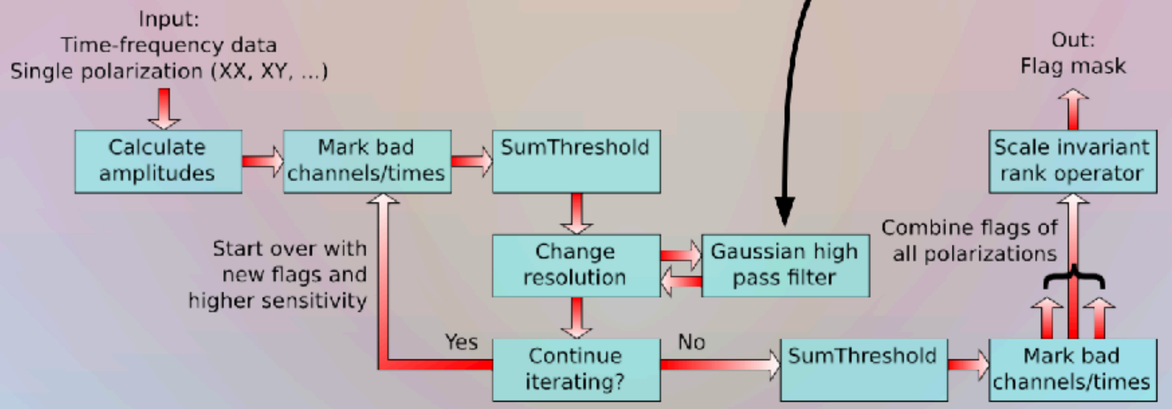
AOFlagger


Input:
Time-frequency data
Single polarization (XX, XY, ...)
Calculate amplitudes
Mark bad channels/times
SumThreshold
Change resolution
Gaussian high pass filter
Start over with new flags and higher sensitivity
Yes
Continue iterating?
No
SumThreshold
Mark bad channels/times
Combine flags of all polarizations
Scale invariant rank operator
Out:
Flag mask

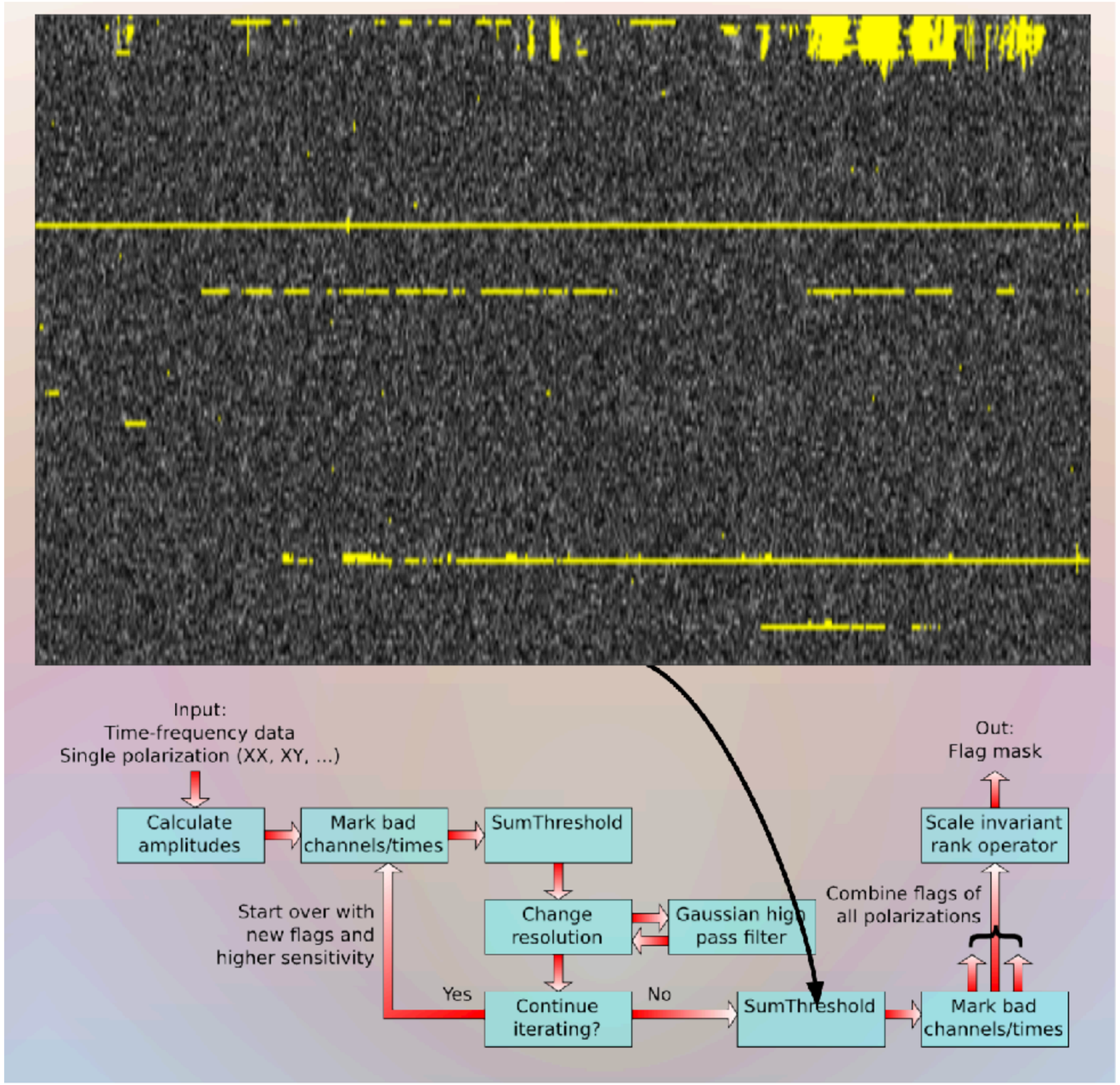
AOFlagger


Subtracted “background”

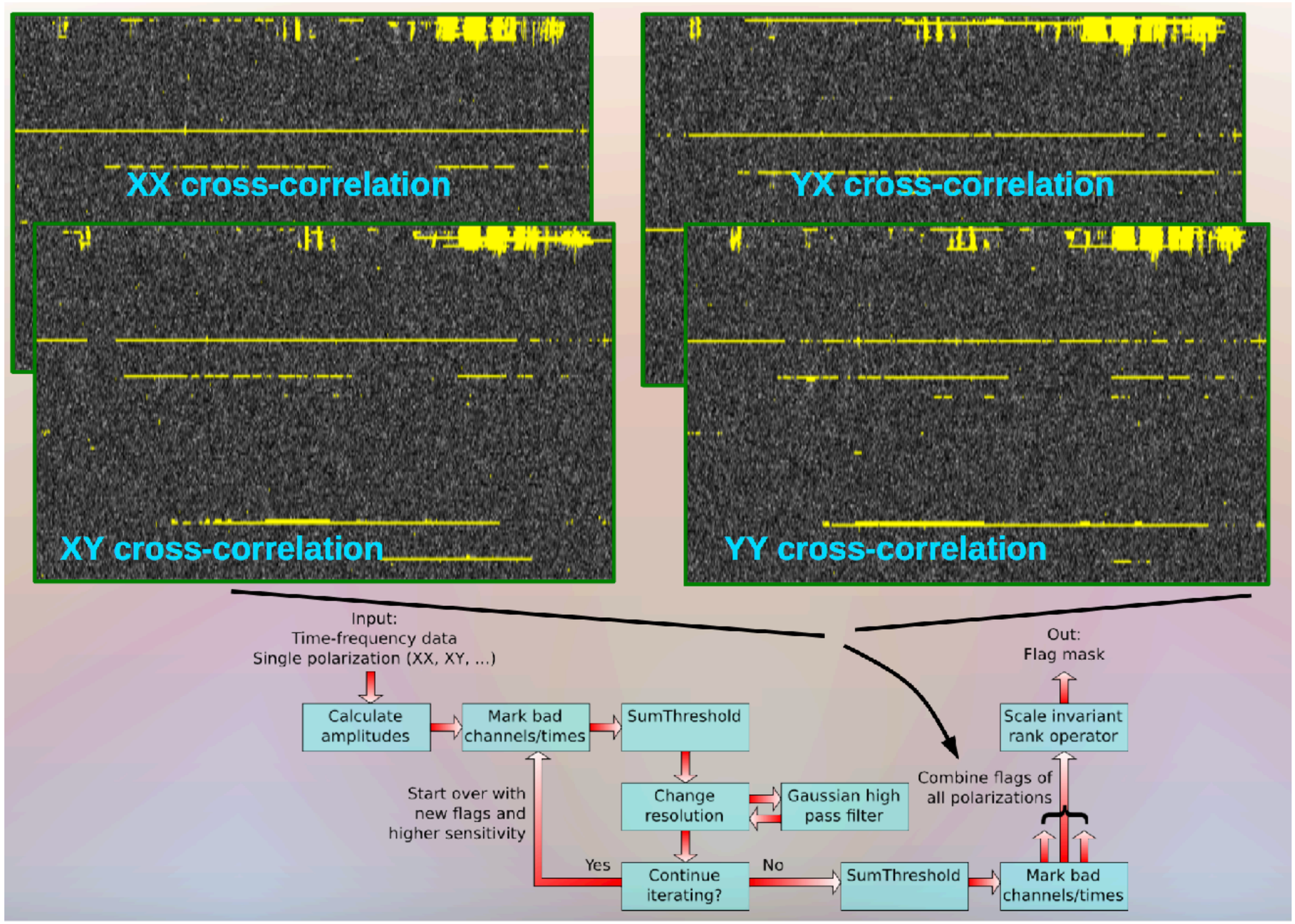

High-frequency components


Input:
Time-frequency data
Single polarization (XX, XY, ...)
Calculate amplitudes
Mark bad channels/times
SumThreshold
Change resolution
Gaussian high pass filter
Start over with new flags and higher sensitivity
Yes
Continue iterating?
No
SumThreshold
Mark bad channels/times
Combine flags of all polarizations
Scale invariant rank operator
Out:
Flag mask

AOFlagger

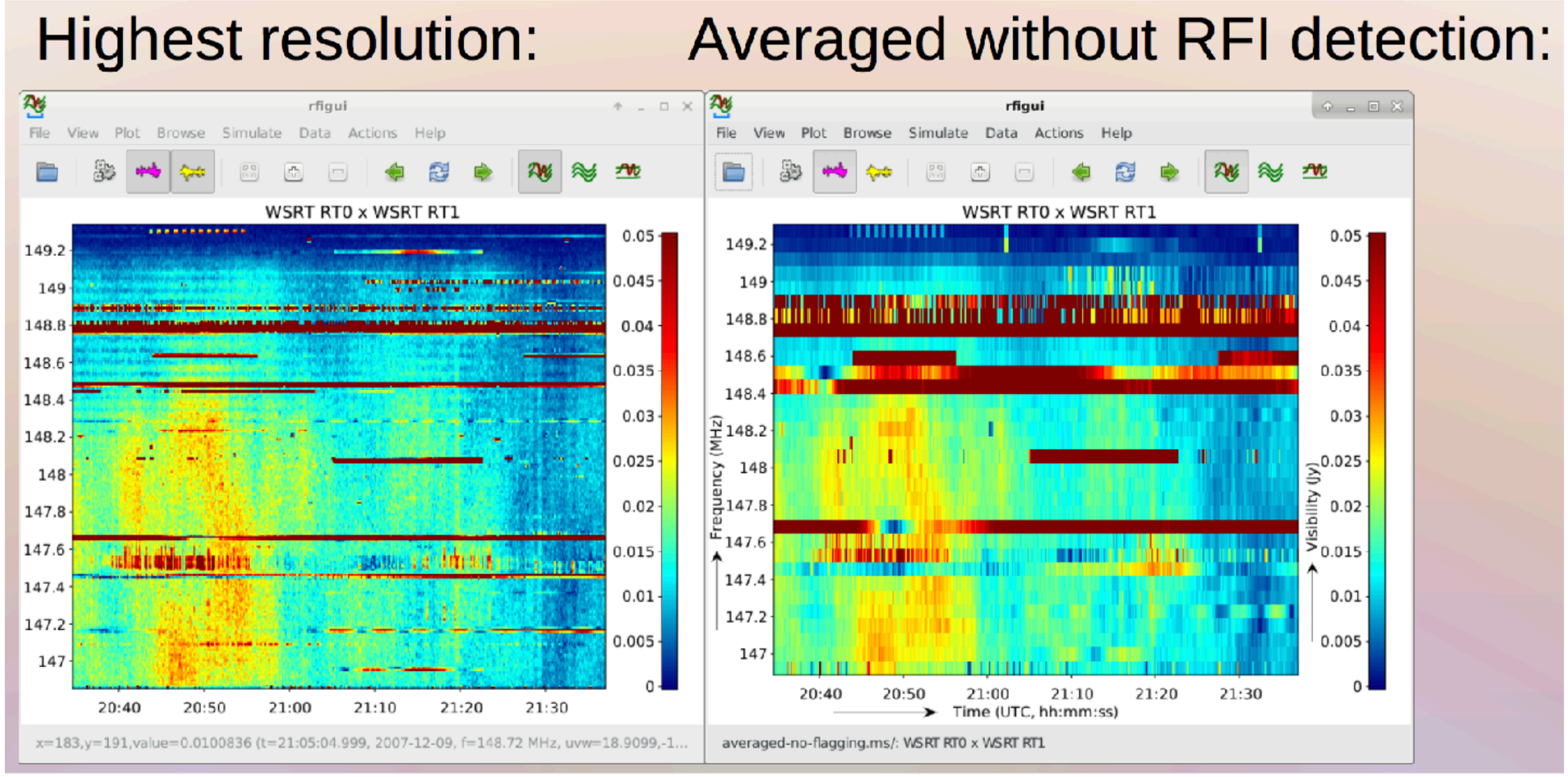

Input:
Time-frequency data
Single polarization (XX, XY, ...)
Calculate amplitudes
Mark bad channels/times
SumThreshold
Change resolution
Gaussian high pass filter
Start over with new flags and higher sensitivity
Yes
Continue iterating?
No
SumThreshold
Mark bad channels/times
Combine flags of all polarizations
Scale invariant rank operator
Out:
Flag mask

AOFlagger

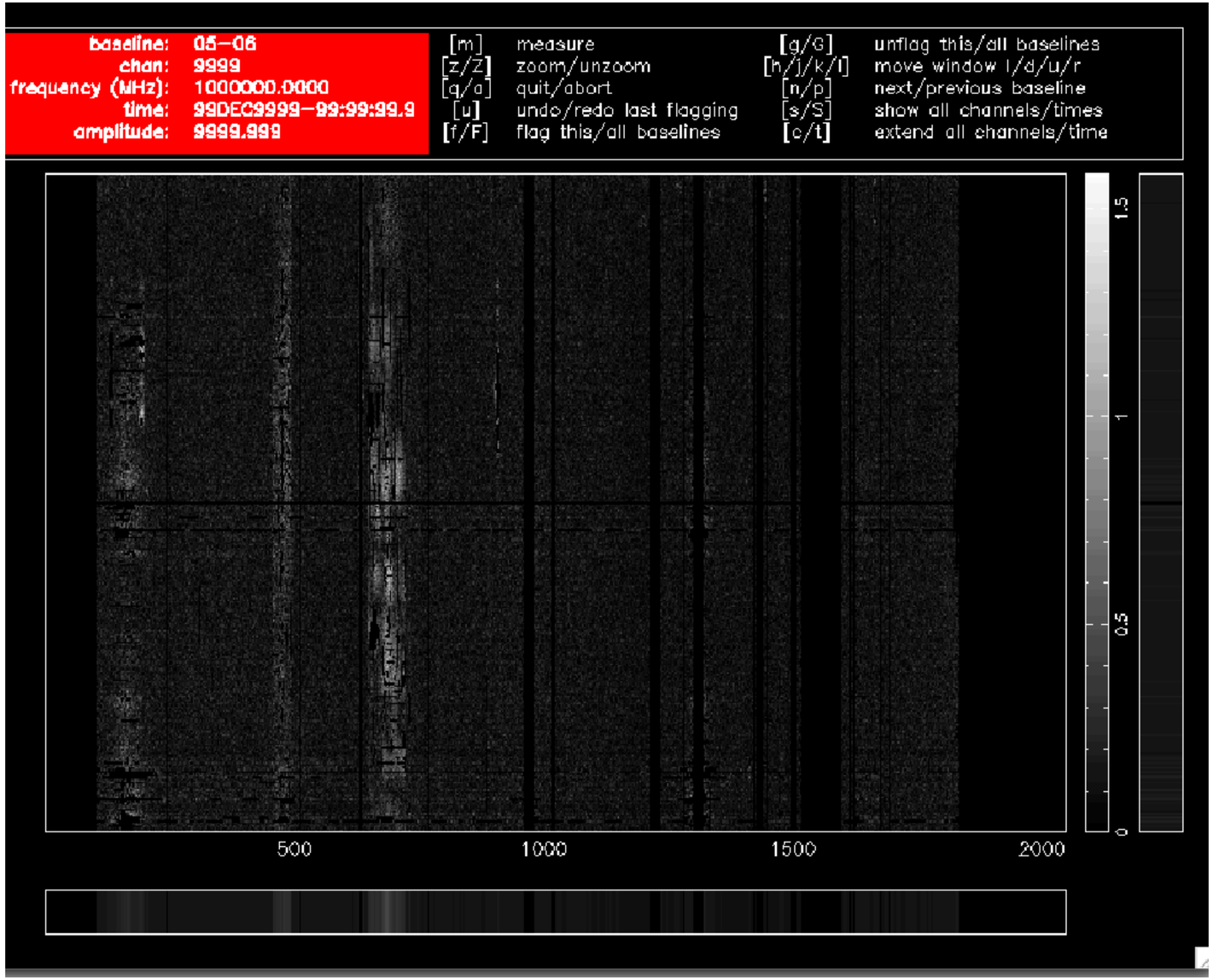

XX cross-correlation
YX cross-correlation
XY cross-correlation
YY cross-correlation
Input:
Time-frequency data
Single polarization (XX, XY, ...)
Calculate amplitudes
Mark bad channels/times
SumThreshold
Change resolution
Gaussian high pass filter
Start over with new flags and higher sensitivity
Yes
Continue iterating?
No
SumThreshold
Mark bad channels/times
Combine flags of all polarizations
Scale invariant rank operator
Out:
Flag mask

AOFlagger

- Always flag (first) at highest possible resolution



Flagging is often iterative with calibration



Conclusion

- First step in data processing is data inspection
- Then flagging (making sure you back up your data!)
- Flagging in VLBI can be quite manual but most other flagging routines are automative
- Remember: Better to have less data that is good than more data that has bad components.