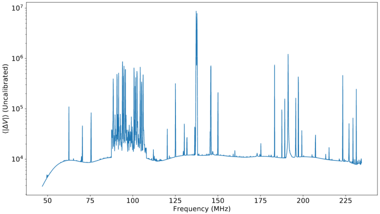
A Bayesian approach to RFI mitigation

Samuel AK Leeney (sakl2@cam.ac.uk)

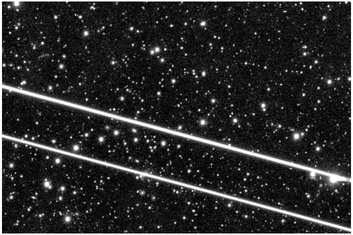
Developed in collaboration with Will Handley and Eloy de Lera Acedo

Astrostatistics and Astro-Machine Learning Workshop **2 June 2023**

What is RFI?



[photo credit: HERA Collaboration]



[photo credit: Elon Musk?]

Why take a Bayesian approach?

Many effective algorithms already exist...

- Cumulative sum [Baan et al., 2004]
- Single value decomposition [Offringa et al., 2012].
- Convolutional neural nets [Sun et al., 2022].

Our approach

- Can be used as part of a single step fitting process within Bayesian pipelines.
- Flagging and management performed simultaneously.

Bayes Theorem

$$\text{likelihood} \times \text{prior} = \text{posterior} \times \text{evidence} \tag{1}$$

$$P(\mathcal{D}|\theta) \times P(\theta) = P(\theta|\mathcal{D}) \times P(\mathcal{D}), \tag{2}$$

$$\mathcal{L} \times \pi = \mathcal{P} \times \mathcal{Z}, \tag{3}$$

RFI correcting likelihood

a) Generate new likelihood capable of modeling probability data point is corrupted.

$$P(\mathcal{D}_i|\theta) = \begin{cases} \mathcal{L}_i(\theta) & : \text{uncorrupted} \\ \Delta^{-1}[0 < \mathcal{D}_i < \Delta] & : \text{corrupted}, \end{cases} \tag{4}$$

RFI correcting likelihood

b) Incorporate prediction of a datum containing RFI into Boolean mask ϵ.

$$P(\mathcal{D}|\theta, \varepsilon) = \prod_i \mathcal{L}_i^{\varepsilon_i} \Delta^{\varepsilon_i - 1} \tag{5}$$

RFI correcting likelihood

c) Ascribe Bernoulli prior $P(\varepsilon_i)$ to $P(\mathcal{D}|\theta)$

$$P(\varepsilon_i) = p_i^{(1-\varepsilon_i)}(1-p_i)^{\varepsilon_i}. \tag{6}$$

RFI correcting likelihood

c) Marginalise over epsilon.

$$P(\mathcal{D}|\theta) = \sum_{\varepsilon \in \{0,1\}^N} P(\mathcal{D}, \varepsilon|\theta) \tag{7}$$

RFI correcting likelihood

d) Assume that the correct (maximum) mask will generate a likelihood that is orders of magnitude 'more likely' than all other masks.

$$P(\mathcal{D}|\theta, \varepsilon^{\mathrm{max}}) \gg P(\mathcal{D}|\theta, \varepsilon^{2}), \tag{8}$$

$$P(\mathcal{D}|\theta) \approx P(\mathcal{D}, \varepsilon^{\mathrm{max}}|\theta). \tag{9}$$

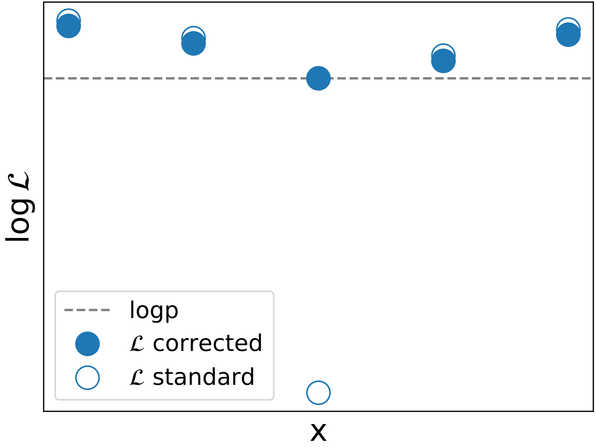
RFI correcting likelihood

e) Taking logs, the loglikelihood is

$$\log P(\mathcal{D}|\theta) = \sum_i [\log \mathcal{L}_i + \log(1 - p_i)]\varepsilon_i^{\max} + [\log p_i - \log \Delta](1 - \varepsilon_i^{\max}), \tag{10}$$

$$\log P(\mathcal{D}|\theta) = \begin{cases} \log \mathcal{L}_i + \log(1 - p_i), & [\log \mathcal{L}_i + \log(1 - p_i) \\ & > \log p_i - \log \Delta] \\ \log p_i - \log \Delta, & \text{otherwise.} \end{cases} \tag{11}$$

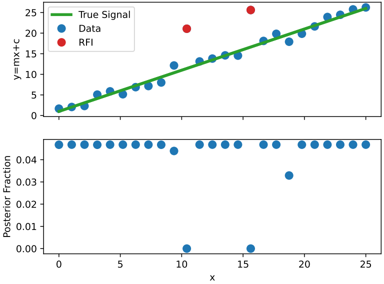
.



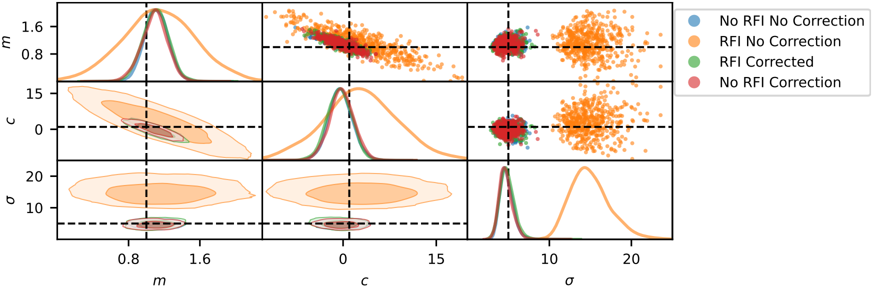
Testing on a simple toy model

Using numerical sampling techniques to make predictions with $\mathcal{L}$.

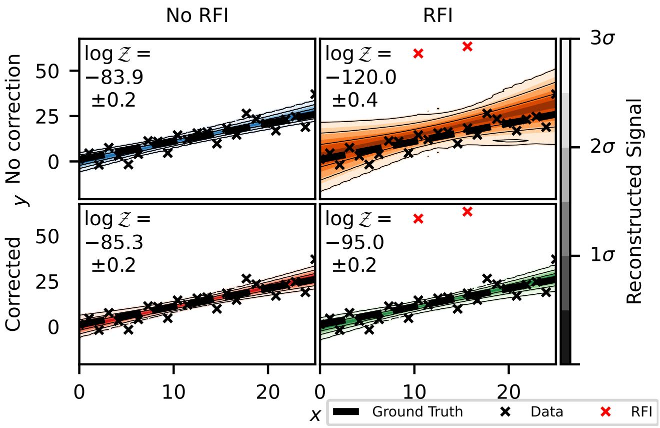
- ‘Belief’ in classification incorporated into model.
- Individual datum are not excised.
- The masks ‘opacity’ changes based confidence of classification.



Posterior plots (basic toy model)

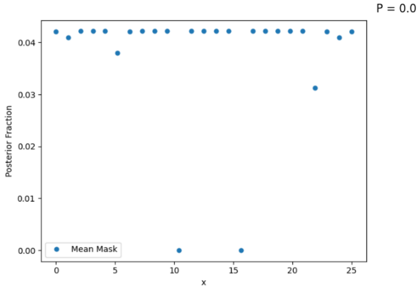


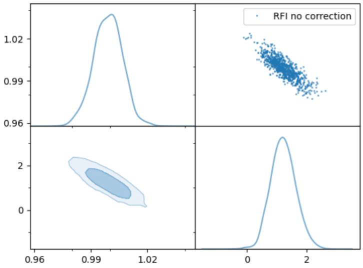
Testing on a simple toy model



Probability thresholding condition p

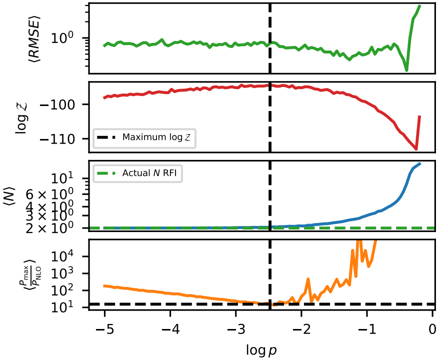
$$\log P(\mathcal{D}|\theta) = \sum_i [\log \mathcal{L}_i + \log(1 - p_i)]\varepsilon^{\max} + [\log p_i - \log \Delta](1 - \varepsilon_i^{\max}), \tag{12}$$



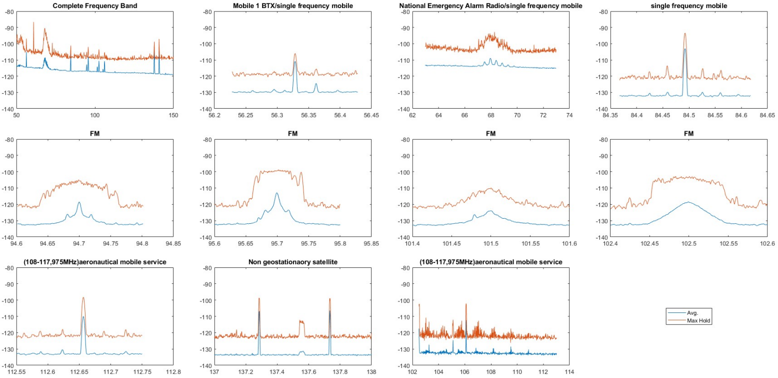


Selection Strategy for p

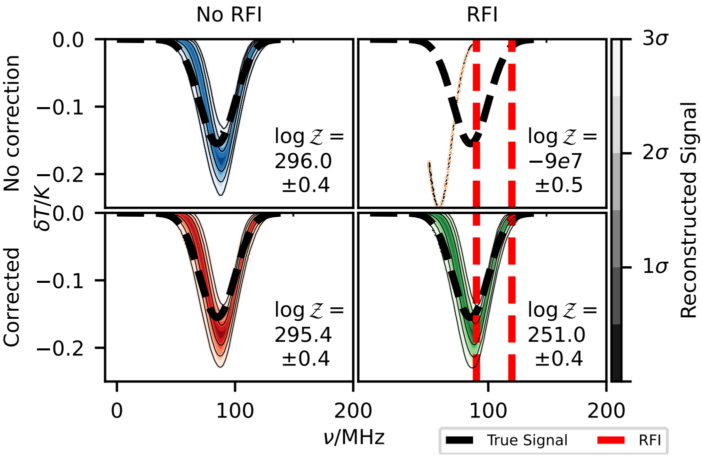
- 'Select p such that the Bayesian Evidence $\mathcal{Z}$ is maximised'



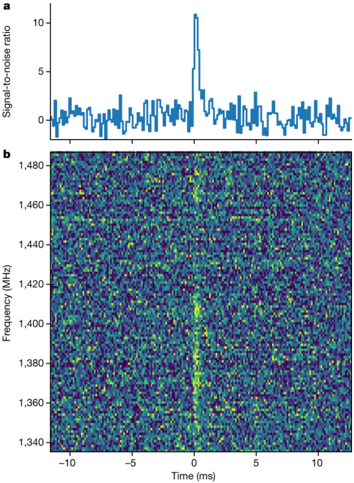
Use case: Global 21cm Cosmology



Use case: Global 21cm Cosmology



General Bayesian anomaly detection?



- Can we use this to detect as well as mitigate?
- Extension beyond just RFI? (see Dominic's talk!)
- Anomalies with more complex structure?

Implement with 2 lines of code

```python
def likelihood(theta):
    sig = theta[0]
    logL = -(f_noise - window)**2/sig**2/2 - np.log(2*np.pi*sig**2)/2
    return logL, []
```

```python
def likelihood(theta):
    sig = theta[0]
    logL = -(f_noise - window)**2/sig**2/2 - np.log(2*np.pi*sig**2)/2 + np.log(1-p)
    emax = logL > logp - np.log(delta)
    logPmax = np.where(emax, logL, logp - np.log(delta)).sum()
    return logPmax, []
```

Tutorial: https://github.com/samleeney

Conclusions

So far ...

- These work serves as a proof of concept that RFI can be mitigated in a truly Bayesian sense.
- RFI can be mitigated as part of a single step fitting process, alongside the Bayesian Evidence and parameter estimations.
- Effective on a toy model and on simulated data for a global 21cm experiment.

Future Works?

- Test on real data.
- Develop for time integrated data (see Dominic's talk!).
- Examine in case where data bins may be correlated.
- Benchmark.

References

Baan, W., Fridman, P., and Millenaar, R. (2004).
Radio frequency interference mitigation at the westerbork synthesis radio telescope: Algorithms, test observations, and system implementation.
The Astronomical Journal, 128(2):933.

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

Sun, H., Deng, H., Wang, F., Mei, Y., Xu, T., Smirnov, O., Deng, L., and Wei, S. (2022).
A robust rfi identification for radio interferometry based on a convolutional neural network.
Monthly Notices of the Royal Astronomical Society,