Watching the Cosmic Dawn from the Moon: how and why

Geraint Harker

University of Colorado

and

Lunar University Network for Astrophysics Research

Problem and approach

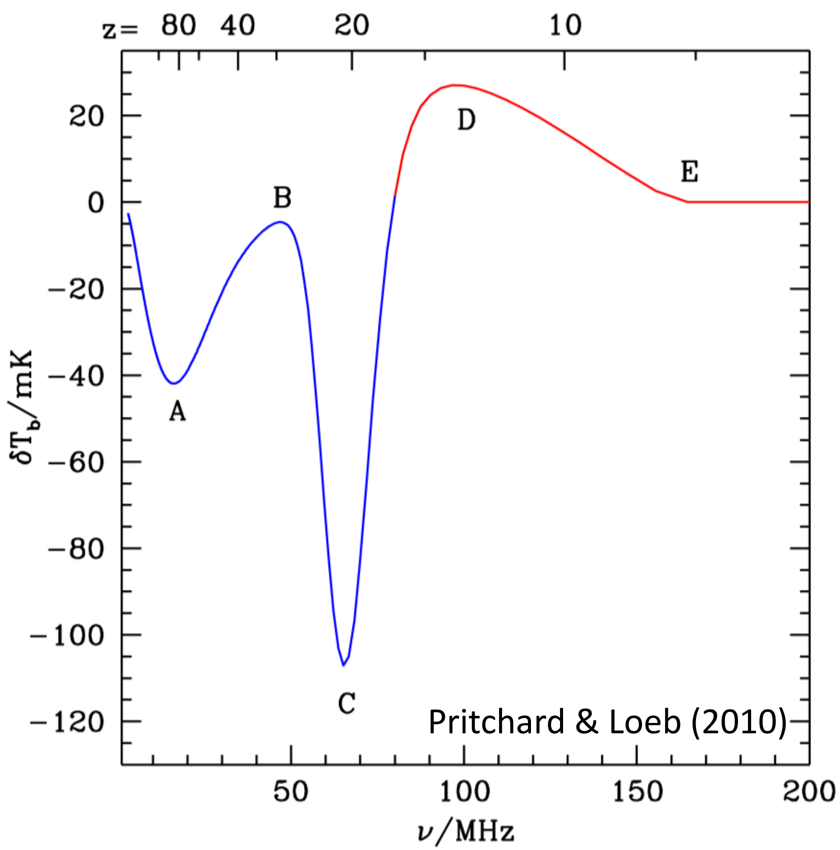
- The signal just described is buried in strong foregrounds and filtered through an instrument.
- We wish to extract properties of the signal despite that.
- Our approach is to parametrize the signal, foregrounds and instrument, then infer the values of those parameters (and errors on them).
- Our main inference tool is a Markov Chain Monte Carlo (MCMC) code, but there are options!

Why the Moon?

- Lunar environment simplifies this process:
  - Stable environment
  - Fewer complications from requirement to remove strong RFI (but must still deal with solar bursts etc.)
  - May avoid otherwise insurmountable problems from low-level RFI.
  - No ionosphere to deal with.
  - Generic advantage of space: whole sky is accessible. Important when our beam encompasses ≈1/8 of the sky at once.
- One of the aims of this analysis is to quantify these advantages.

Parametrization of the signal

- Steve has talked about the physics of the spin-flip background
- Abhi has outlined the experiment; I need a parametrization of that.
- I’m adopting the turning points parametrization to try to capture the essential behaviour Steve’s described.
- Jordan will explain how those turning points relate to the physics.



Bayesian inference

Signal, foregrounds, instrument…

Parameters

Calibrated spectra

Data

$$\boxed{\Pr(\boldsymbol{\Theta}|\mathbf{D},H)} = \frac{\overbrace{\boxed{\Pr(\mathbf{D}|\boldsymbol{\Theta},H)}}^{\text{Likelihood, } \mathcal{L}(\Theta)}\ \overbrace{\boxed{\Pr(\boldsymbol{\Theta}|H)}}^{\text{Prior, } \pi(\Theta)}}{\underbrace{\boxed{\Pr(\mathbf{D}|H)}}_{\text{Evidence, } \mathcal{Z}}}$$

Posterior

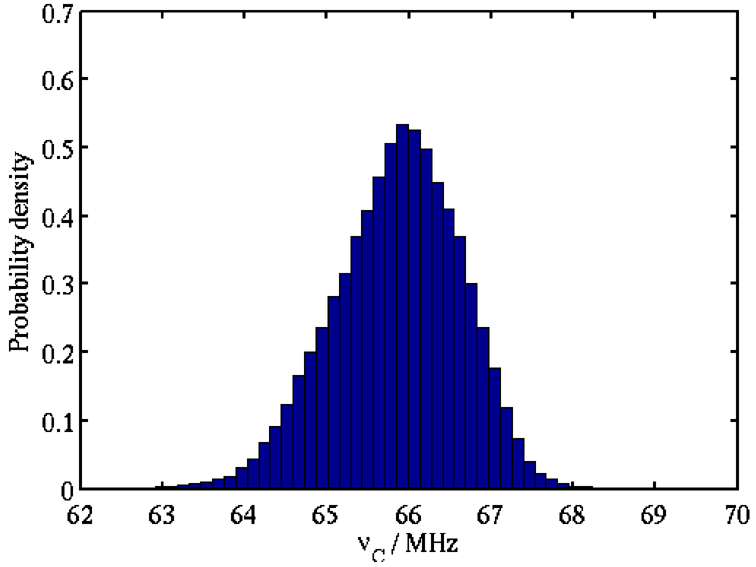
Hypothesis or model

E.g. ‘signal well described by turning points’

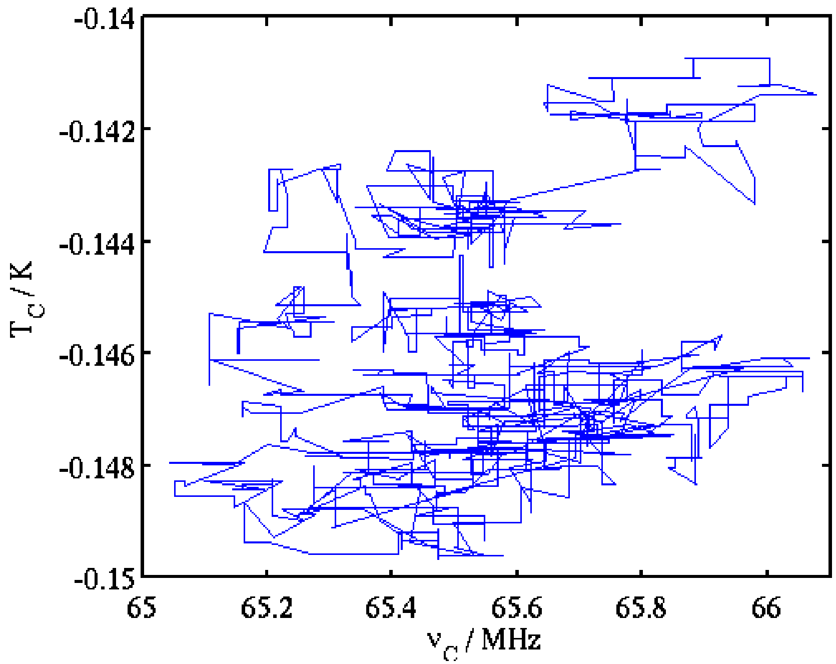
$$\mathcal{Z} = \int \pi(\Theta)\mathcal{L}(\Theta)d^N\Theta$$

Model comparison!

Outline of inference using MCMC



A Markov Chain Monte Carlo simulation allows us to draw unbiased, random samples from the posterior probability distribution of the parameters we're trying to find.

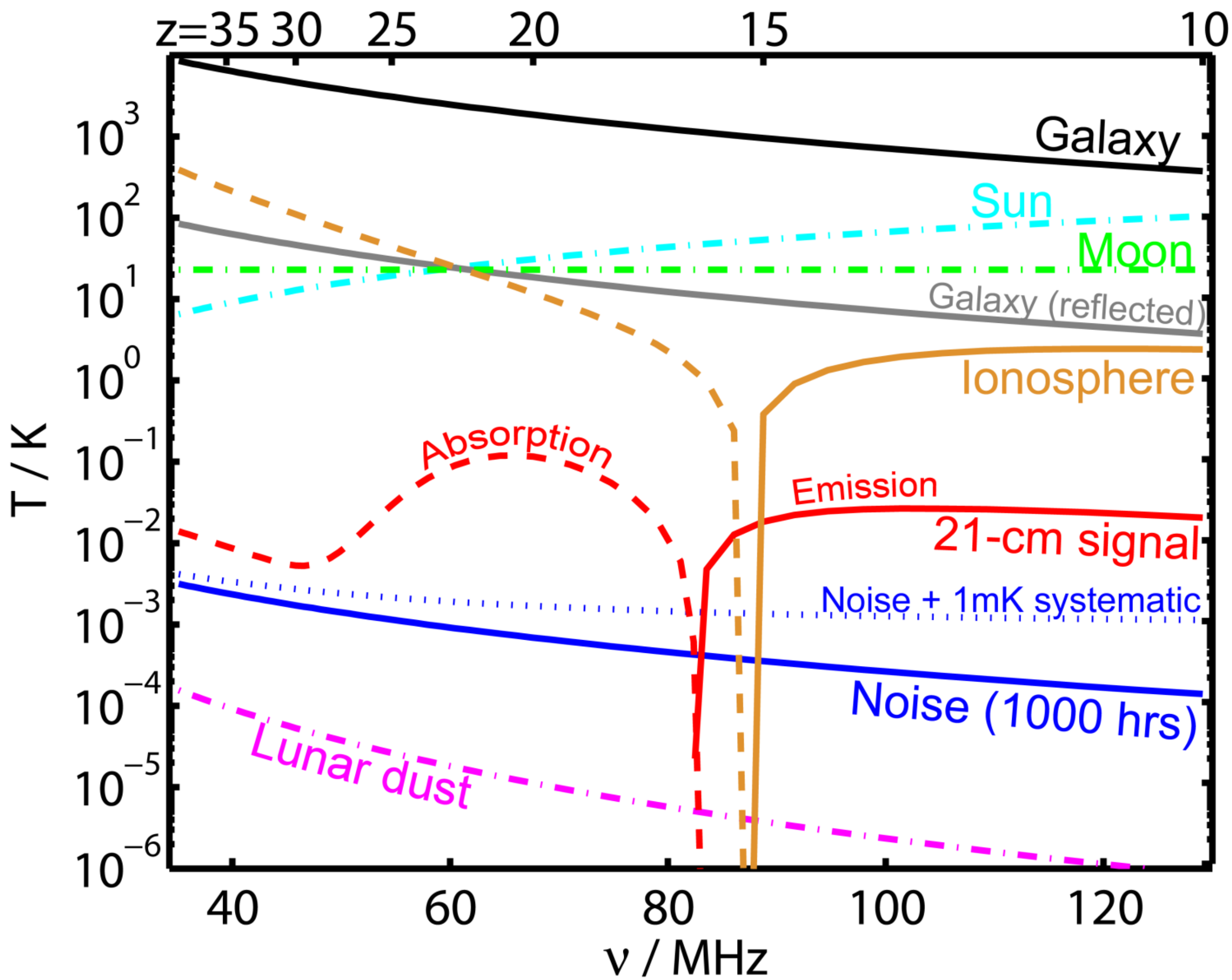


The path taken by part of the Markov Chain through a two-dimensional slice of parameter space.

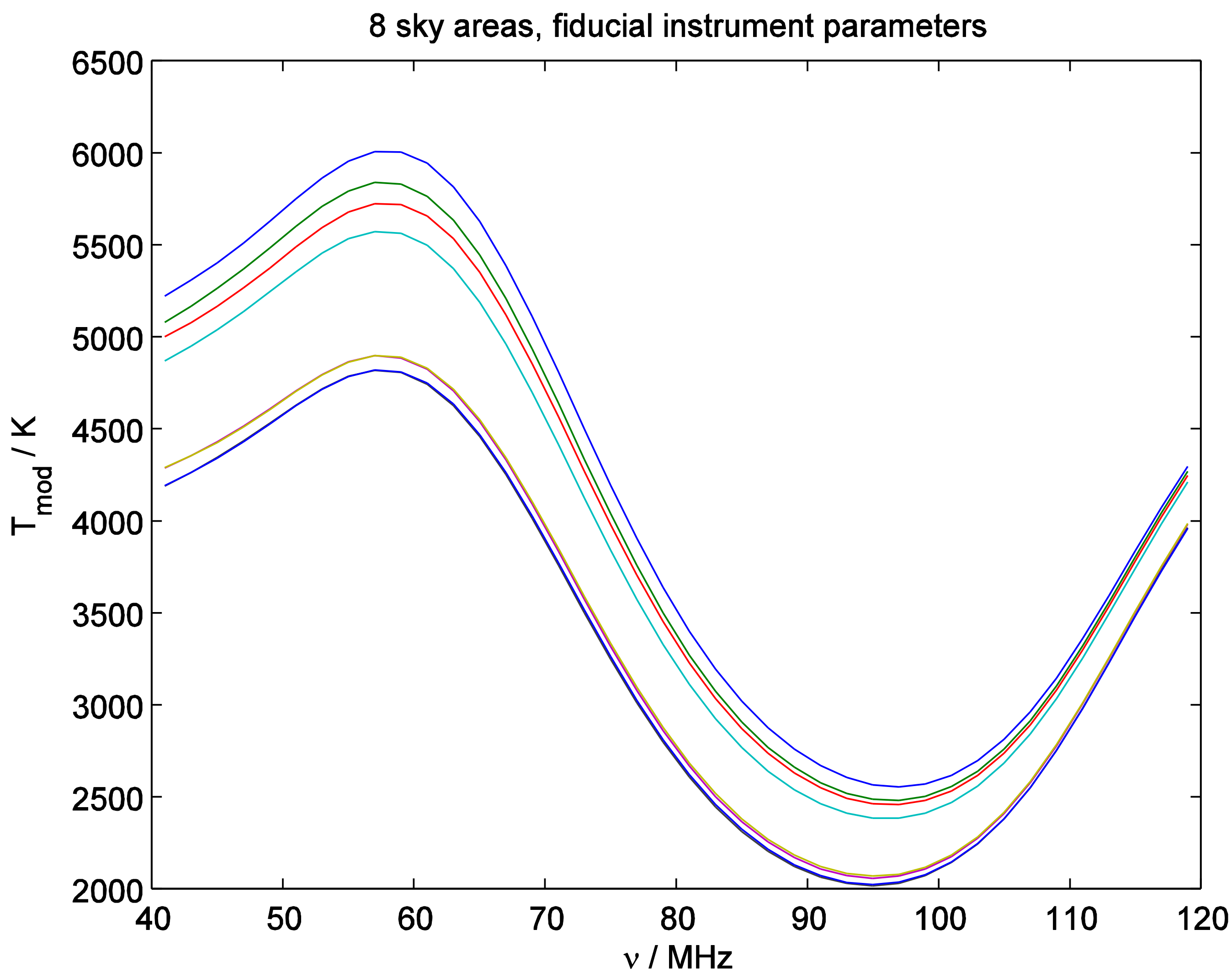
Form of the data input to the MCMC code

- Large beam -> small number of independent sky regions.
- We combine time-ordered data into integrated spectra in these regions.
  - There could be a lot hidden in this step!
- Frequency range 40-120 MHz, chosen to encompass the interesting part of the signal (turning points B, C and D).
- Aiming for thermal noise per 2 MHz channel of order mK.

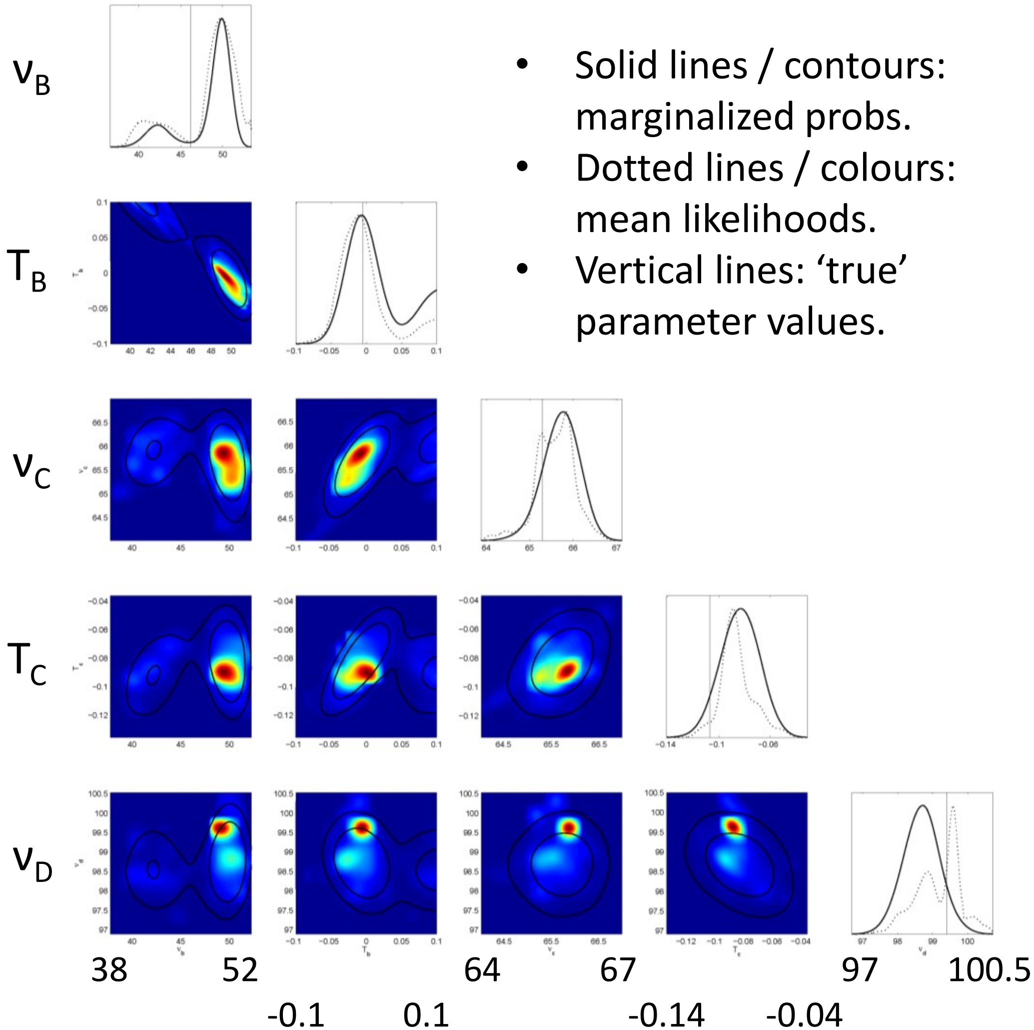
Contributions to the signal



Calibrated spectra



Form of the results

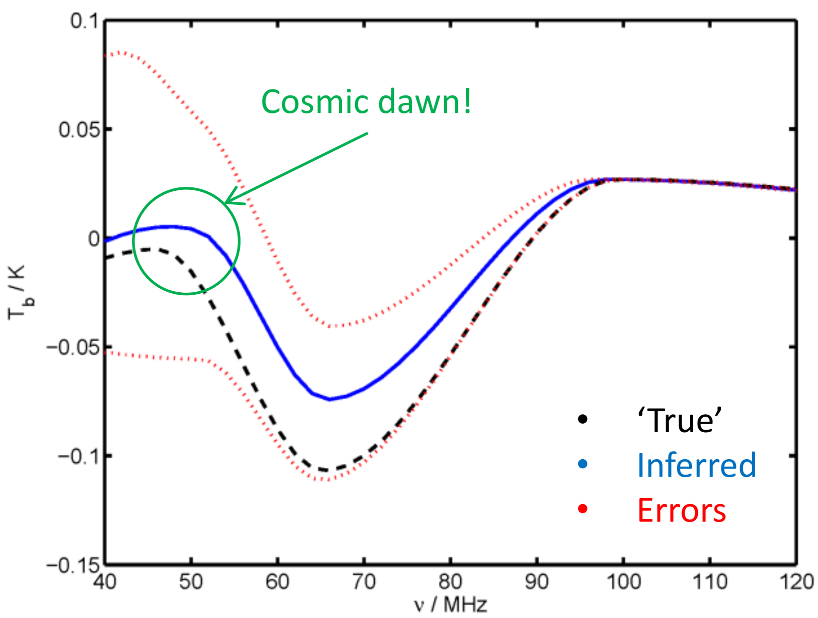


- Solid lines / contours: marginalized probs.
- Dotted lines / colours: mean likelihoods.
- Vertical lines: ‘true’ parameter values.

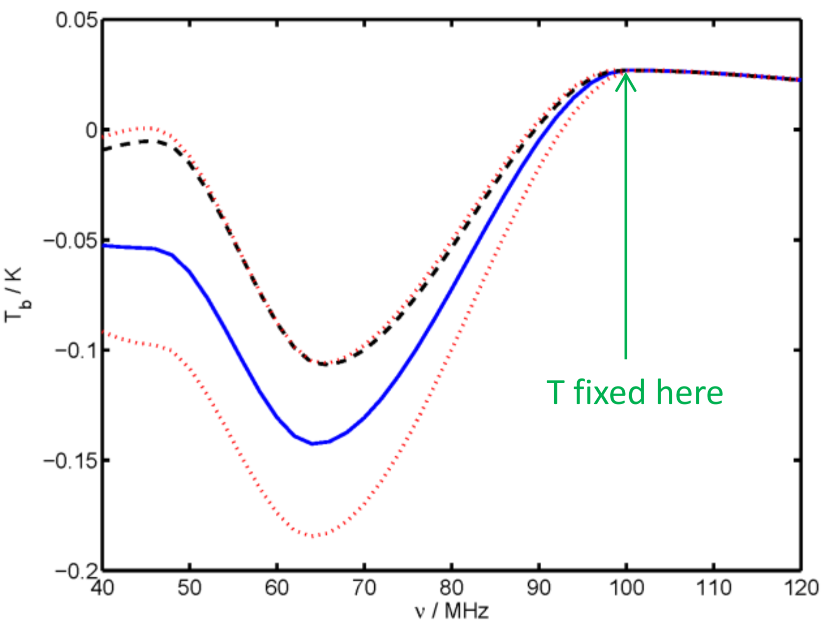
- 1D and 2D projections of a ≤86 dim. parameter space (so far).
- T_D isn’t shown as it’s held fixed.
- We can see marginalized constraints on, and correlations between, any combination of the parameters.

Translating this into a 21-cm spectrum – extracting the shape with 95% confidence regions

Eight sky regions

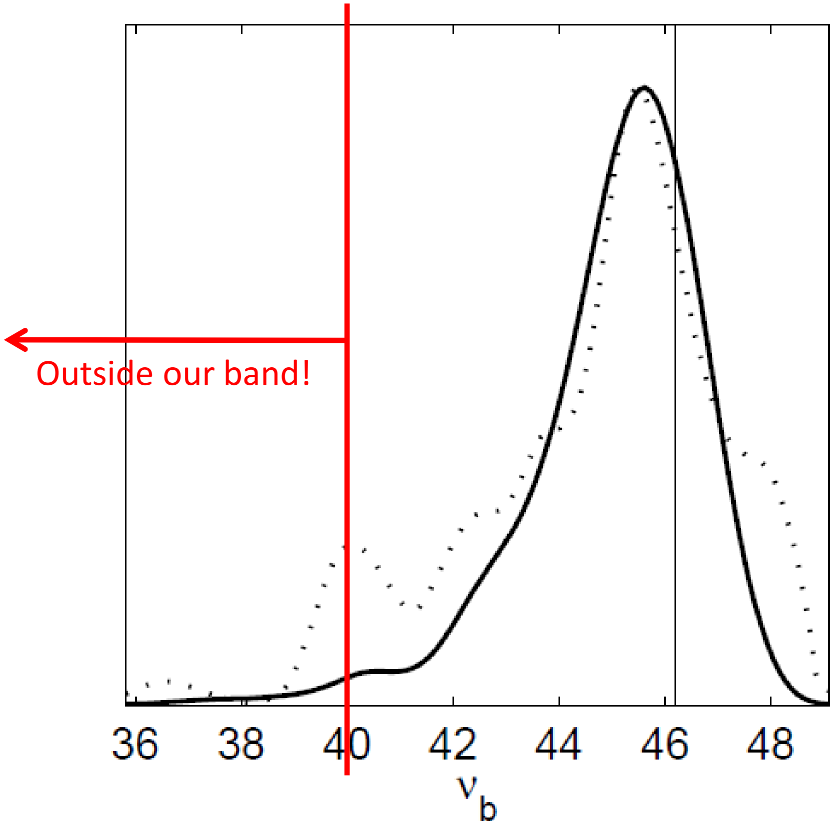


Two sky regions

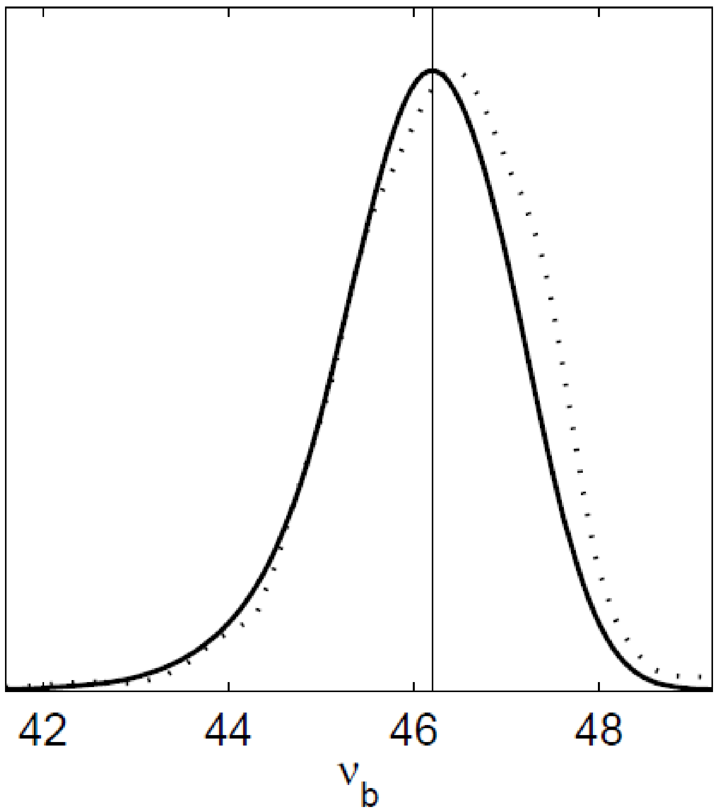


The ionosphere, an unavoidable problem from the ground: simplified toy model

With



Without



Conclusions

- To extract the 21-cm signal from DARE data we aim to use robust, well-established inference tools like MCMC.
  - To be extended with tools for rigorous model selection.
- Even for idealized models, these show the clear and quantitative benefits of doing 21-cm science from the lunar farside.
  - Whole sky accessible
  - No ionosphere!
  - Also: no RFI, stable environment
- Current work:
  - A more realistic, *variable* ionosphere.
  - Temperature dependence of instrumental parameters.
  - Dealing with the time-ordered data.