`maxsmooth` and its applications in 21-cm cosmology

UNIVERSITY OF CAMBRIDGE

CAVENDISH ASTROPHYSICS
CM-WAVE EXPERIMENTAL RADIO COSMOLOGY
Science Technology Engineering Mathematics

maxsmooth: rapid maximally smooth function fitting with applications in Global 21-cm cosmology

H. T. J. Bevins[1★], W. J. Handley[1,2], A. Fialkov[2,3], E. de Lera Acedo[1,2], L. J. Greenhill[4], D. C. Price[4,5]

[1] Astrophysics Group, Cavendish Laboratory, J. J. Thomson Avenue, Cambridge, CB3 0HE, UK
[2] Kavli Institute for Cosmology, Madingley Road, Cambridge CB3 0HA, UK
[3] Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK
[4] Harvard-Smithsonian Center for Astrophysics, MS42, 60 Garden Street, Cambridge MA 02138 USA
[5] Centre for Astrophysics and Supercomputing, Swinburne University of Technology, Hawthorn VIC 3122 Australia

https://arxiv.org/abs/2007.14970

What are Maximally Smooth Functions?

- Originally proposed in Sathyanarayana Rao et al. (2015, 2017)

- Functions constrained by:

$$\frac{d^m y}{dx^m} \geq 0 \quad \text{or} \quad \frac{d^m y}{dx^m} \leq 0$$

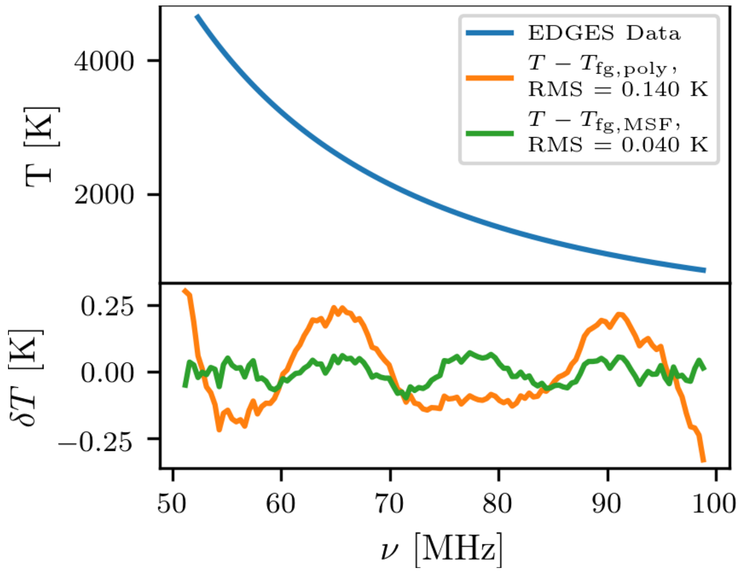
  with $m \geq 2$.

- Example function $y = y_0 \sum_{k=0}^{N} a_k \left(\frac{x}{x_0}\right)^k$

- Smooth properties ideal for modelling foregrounds

- Generalise to Derivative Constrained Functions (DCFs) with different constraints on derivatives

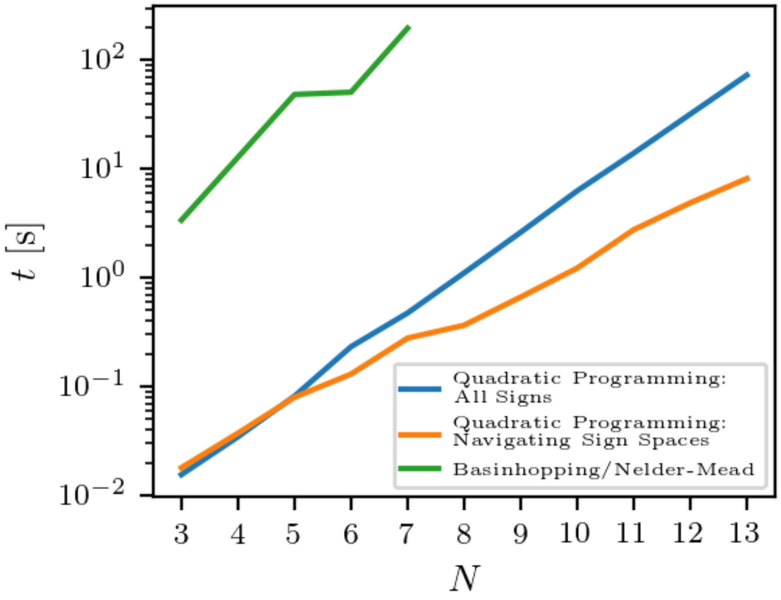
- Example using the EDGES data



What is `maxsmooth`?

- Robust and fast algorithm for fitting DCFs
- 2 orders of magnitude faster than historically used Basin-hopping/Nelder-Mead routines
- Based in quadratic programming and a division of the parameter space determined on the constraints (more details in the paper)

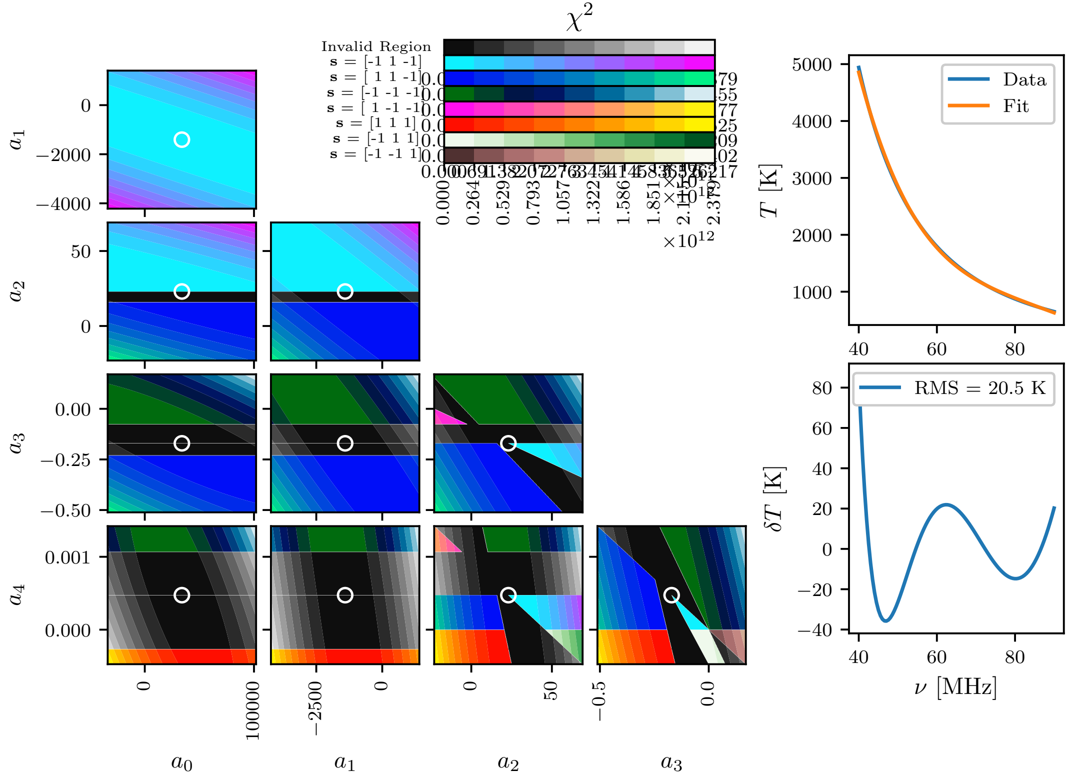
$$\frac{d^m y}{dx^m} \geq 0 \quad \text{or} \quad \frac{d^m y}{dx^m} \leq 0$$



What is `maxsmooth`?

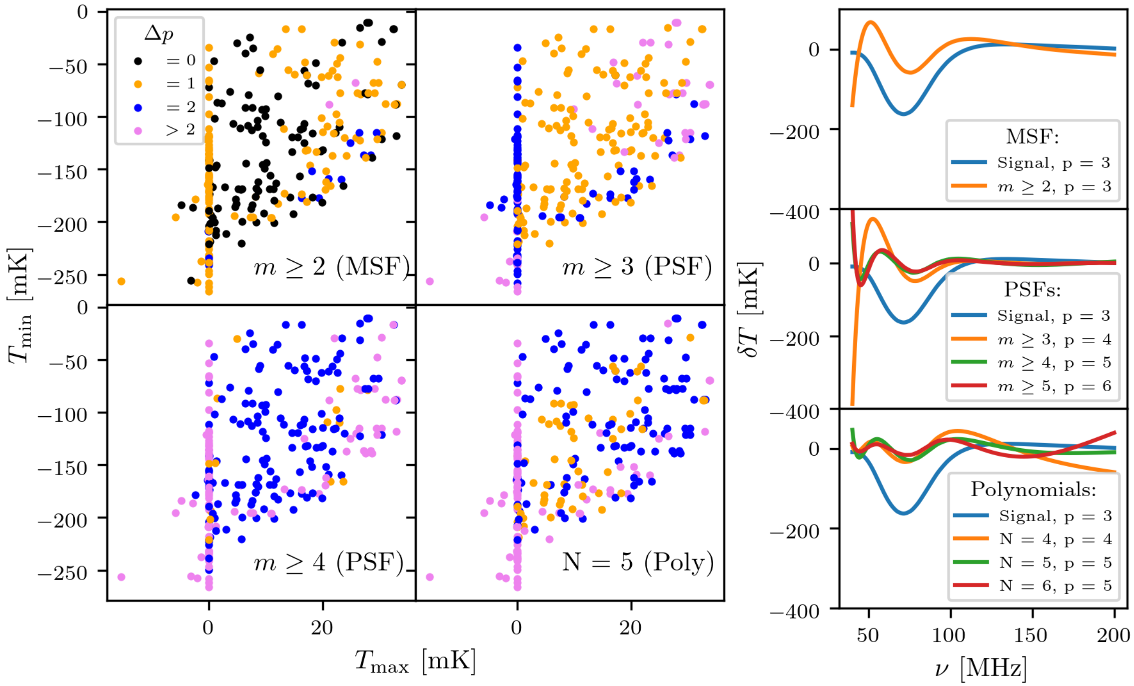
- Division of the parameter space
- For Basin-hopping each region acts as a basin but these aren't always explored
- `maxsmooth` allows you to explore the entire parameter space
- Available on Github: https://github.com/htjb/maxsmooth
- pip install `maxsmooth`

```python
import numpy as np
from maxsmooth.DCF import smooth

n_samples = 242
x = np.linspace(60, 200, n_samples)
y = 5e7*x**(-2.5) + np.random.normal(0, 0.2, n_samples)

N = 10

res = smooth(x, y, N)
```



DCFs and 21-cm Cosmology

- 264 mock signals and a $\nu^{-2.5}$ power law foreground.
- Models from Cohen et al. (2017) and used by Singh et al. (2018)
- Fit and remove different foreground models
- Count turning points, *p*, in residuals and compare with those in the signal
- *Δp=0* indicates a strong remnant of the signal in the residuals

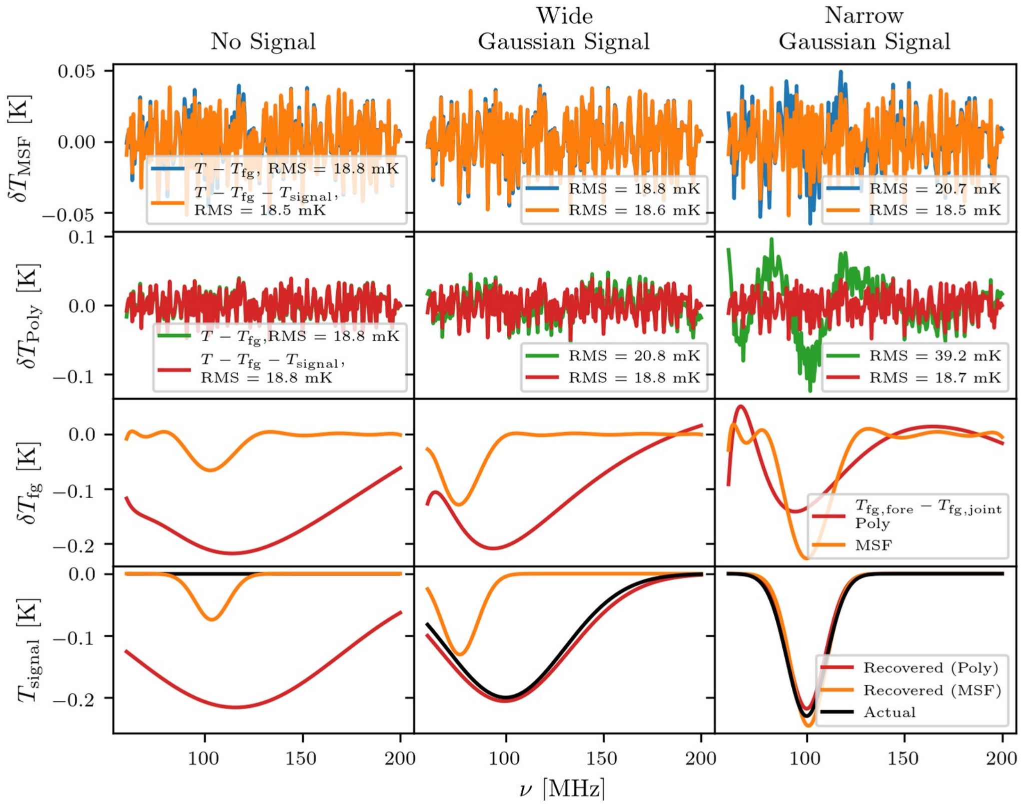


Limitations of DCFs

UNIVERSITY OF CAMBRIDGE

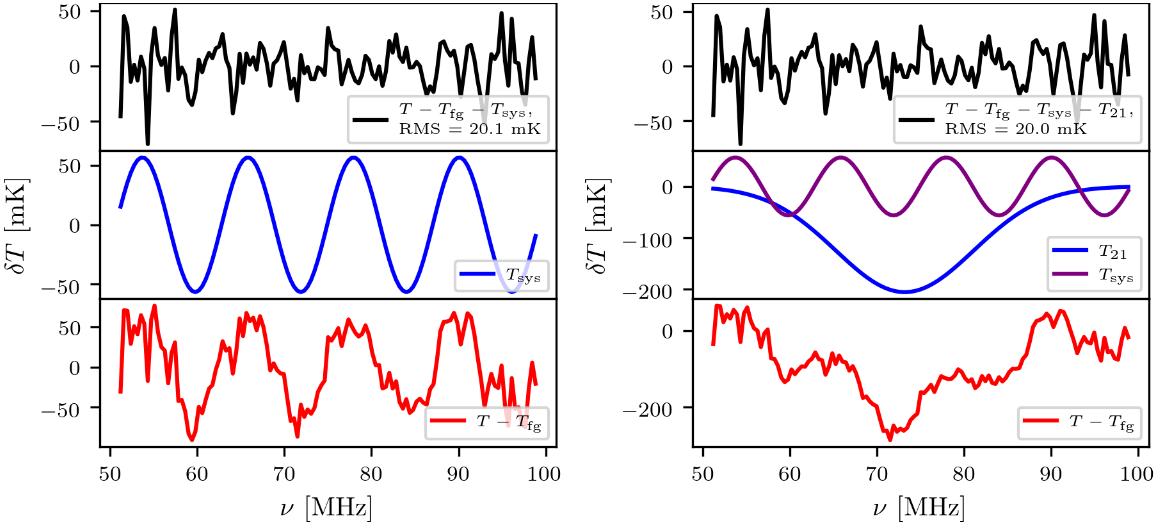
CAVENDISH ASTROPHYSICS
CM-WAVE EXPERIMENTAL RADIO COSMOLOGY
Science Technology Engineering Mathematics

- Three data sets with same foreground and noise but different signals
- Pure foreground and Joint fits of signal with MSFs and an unconstrained polynomial
- Find evidence that MSFs struggle with recovery of ‘smooth’ signals and identification of signals in the absence of any real signal
- Similar problems with polynomials
- Evidence is a strong indicator of presence of smooth signals



`maxsmooth`, DCFs and EDGES

- Testing `maxsmooth` on the EDGES data
- Previously analysed with MSFs by Singh and Subrahmanyan (2019)
- Pure foreground fit has log*(Z)=216.80±0.09*
- Found evidence for sinusoidal systematic, log*(Z)=302.99±0.08*
- We can recover a theoretically plausible Gaussian signal
- However the evidence is lower than for the systematic fit, log*(Z)=292.67±017*

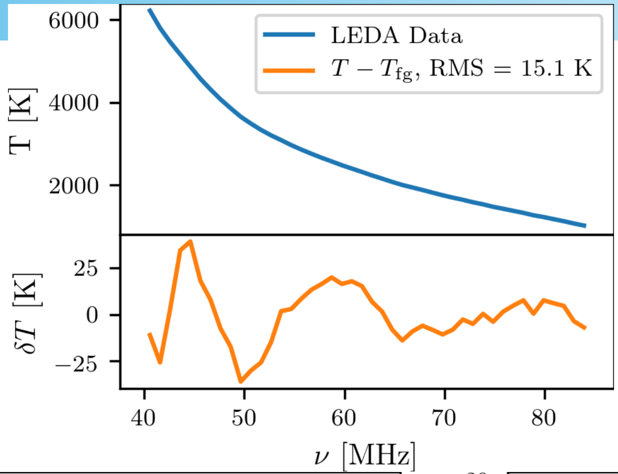


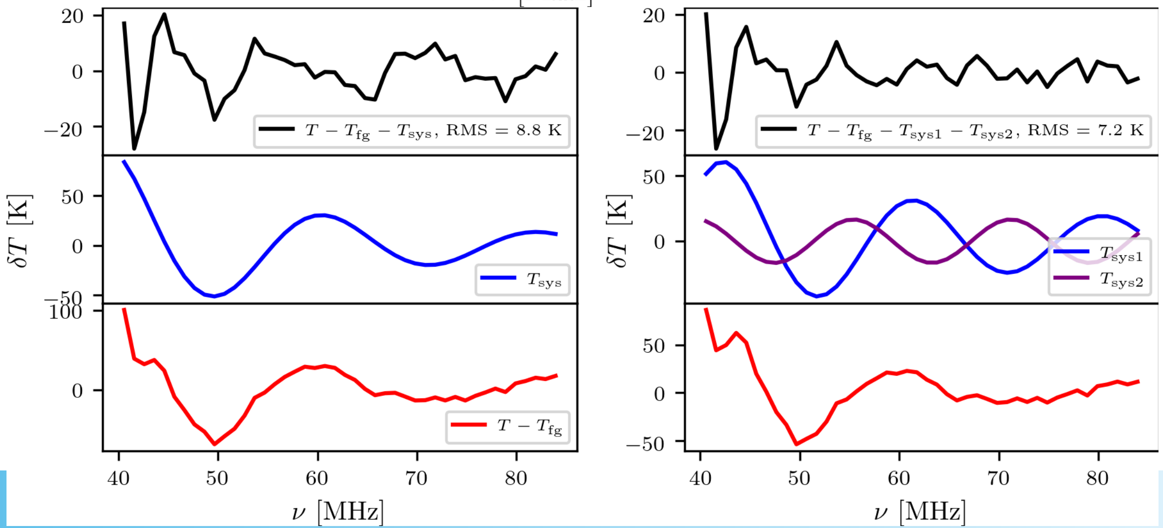
N = 11

$$y = \sum_{k=0}^{N} a_k (x - x_0)^k$$

`maxsmooth`, DCFs and LEDA

- Fit LEDA data, taken in 2016, for the first time with MSFs
- Large RMS and log*(Z)=-185.45±0.09*
- Perform Joint fit with a damped sinusoidal systematic and return log*(Z)=-175.50±0.19*
- Additional sinusoidal structure which we can fit and find log*(Z)=-168.34±0.19*
- Systematic described by the leading order terms in a damped Fourier series
- Similarities between the systematics here and in EDGES might highlight larger causes in Global 21-cm experimentation





Applications to REACH

Best Case Scenario:

- `maxsmooth` can be used as an independent foreground modelling technique alongside the existing Bayesian pipeline to confirm a detection.

Worst Case Scenario:

- `maxsmooth` can be used to confidently identify systematics in the data set leading to quicker identification and improvements to the REACH experiment.

Summary and Conclusions

- `maxsmooth` is a fast and robust tool for fitting MSFs and related functions.
- 2 times faster than historically used methods for fitting MSFs
- MSFs are shown to be better suited for modelling foregrounds than commonly used unconstrained polynomials.
- There are limitations to the usefulness of MSFs
- We can use MSFs to identify the presence of non-smooth systematics in 21-cm experiments like LEDA and EDGES

Paper: https://arxiv.org/abs/2007.14970 Github: https://github.com/htjb/maxsmooth