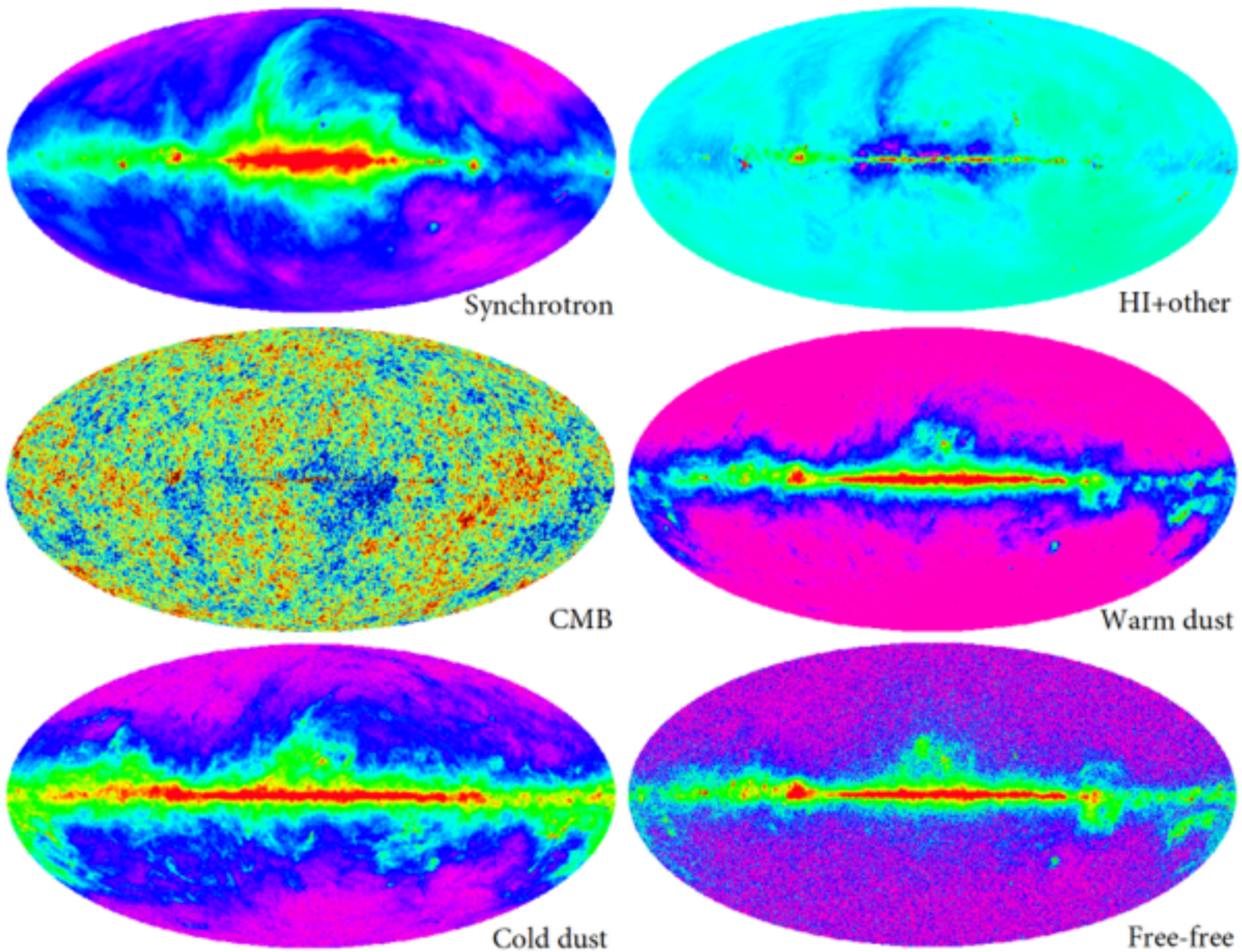
The extended Global Sky Model (eGSM)

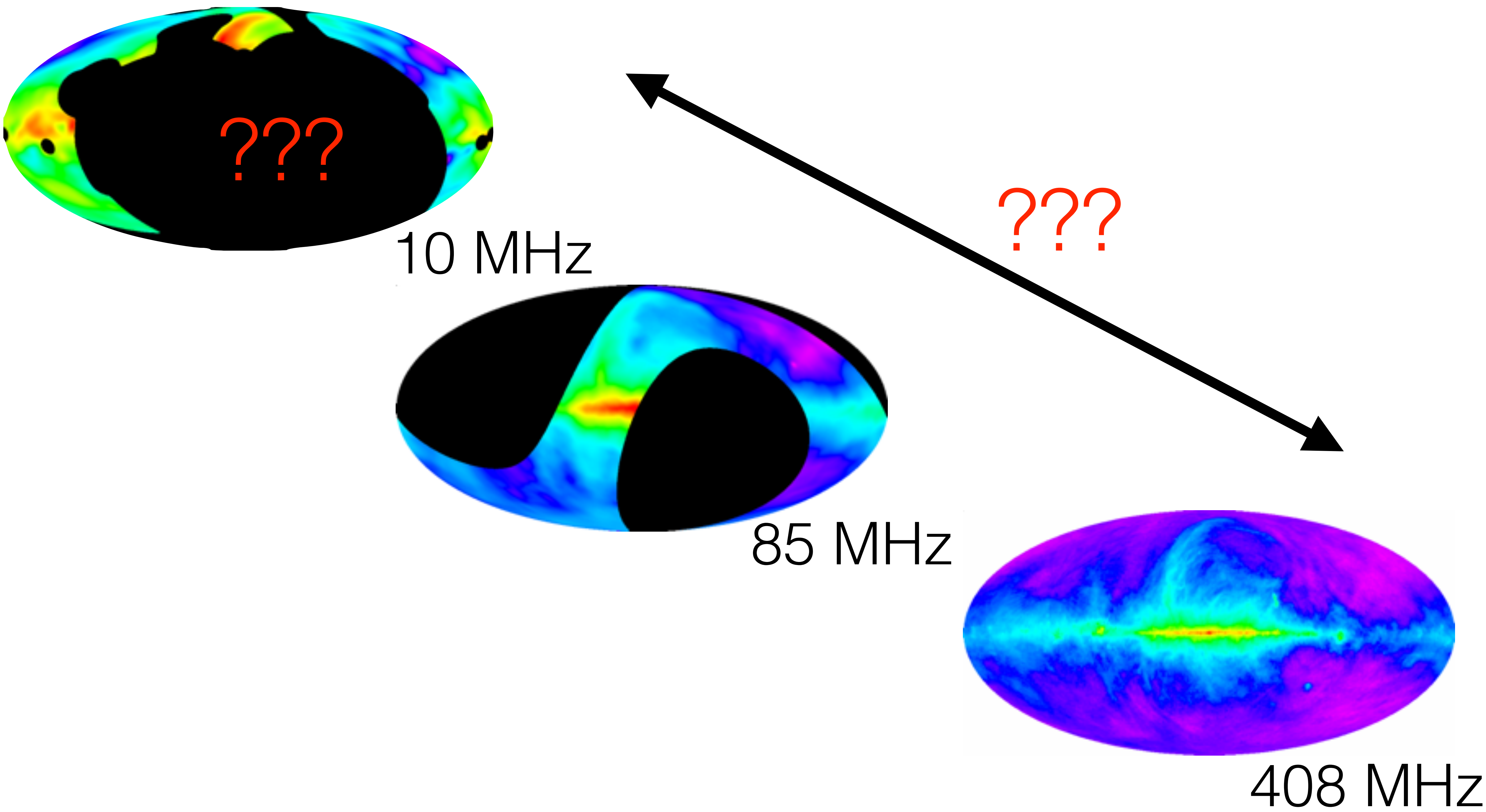


Adrian Liu, UC Berkeley/McGill

The extended Global Sky Model (eGSM) project

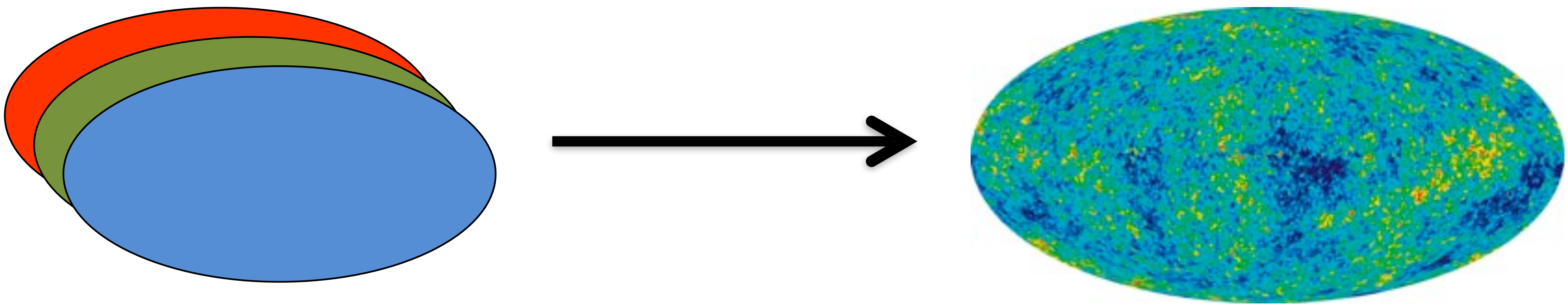
AL, UC Berkeley/McGill
Doyeon “Avery” Kim, UC Berkeley
Eric Switzer, NASA Goddard
Haoxuan “Jeff” Zheng, MIT/Intel

What does the sky look like in all directions at “all” frequencies?

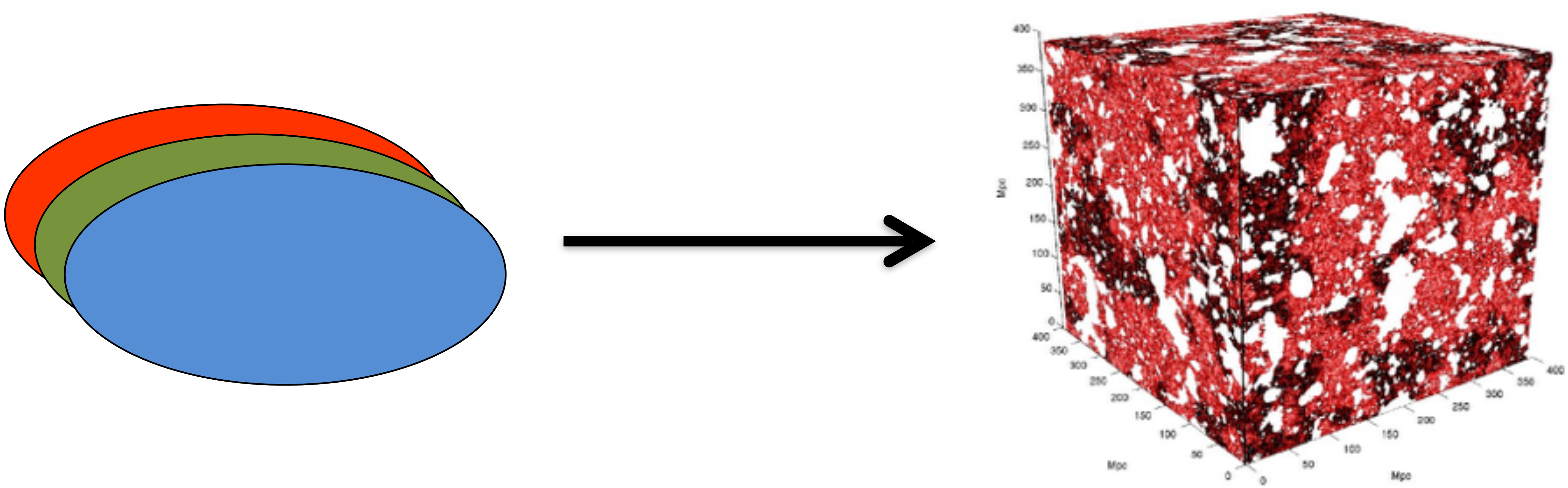


My motivation

Cosmic Microwave Background



21cm Cosmology



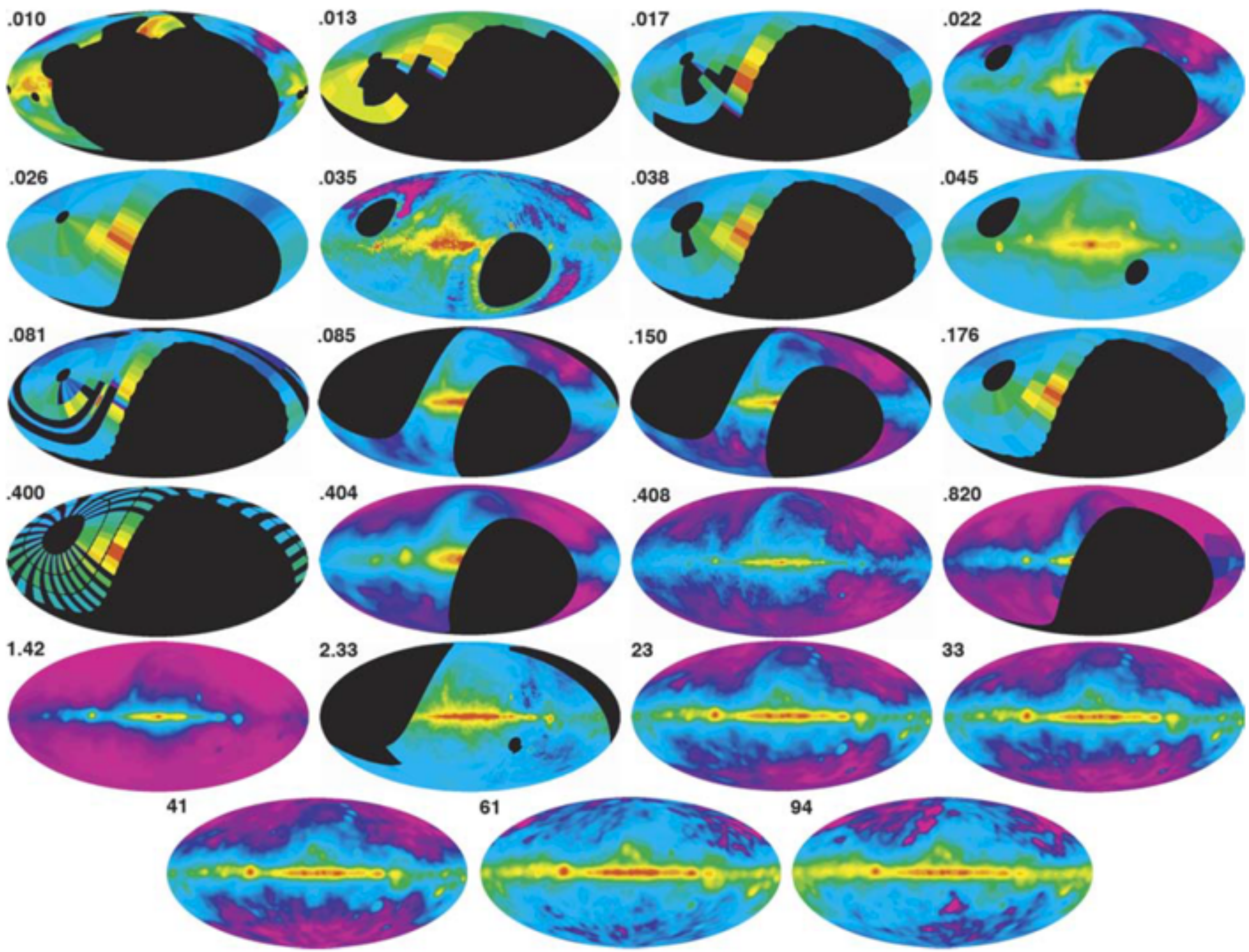
(See **AL**,et al. 2013 PRD 87, 043002
for more details)

How does one model the sky?

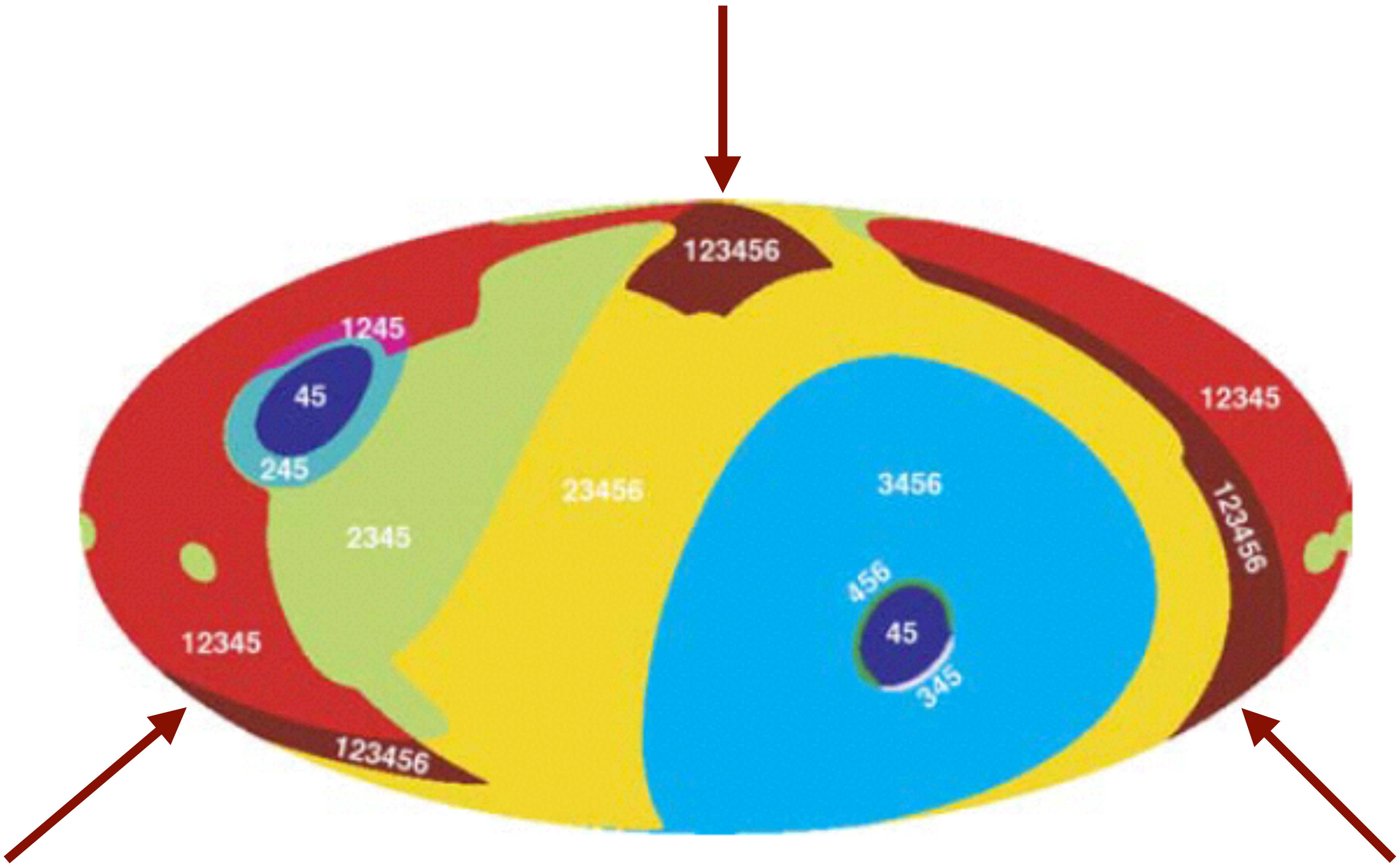
Global Sky Model v1

(de Oliveira-Costa et al. 2008, MNRAS 388, 247)

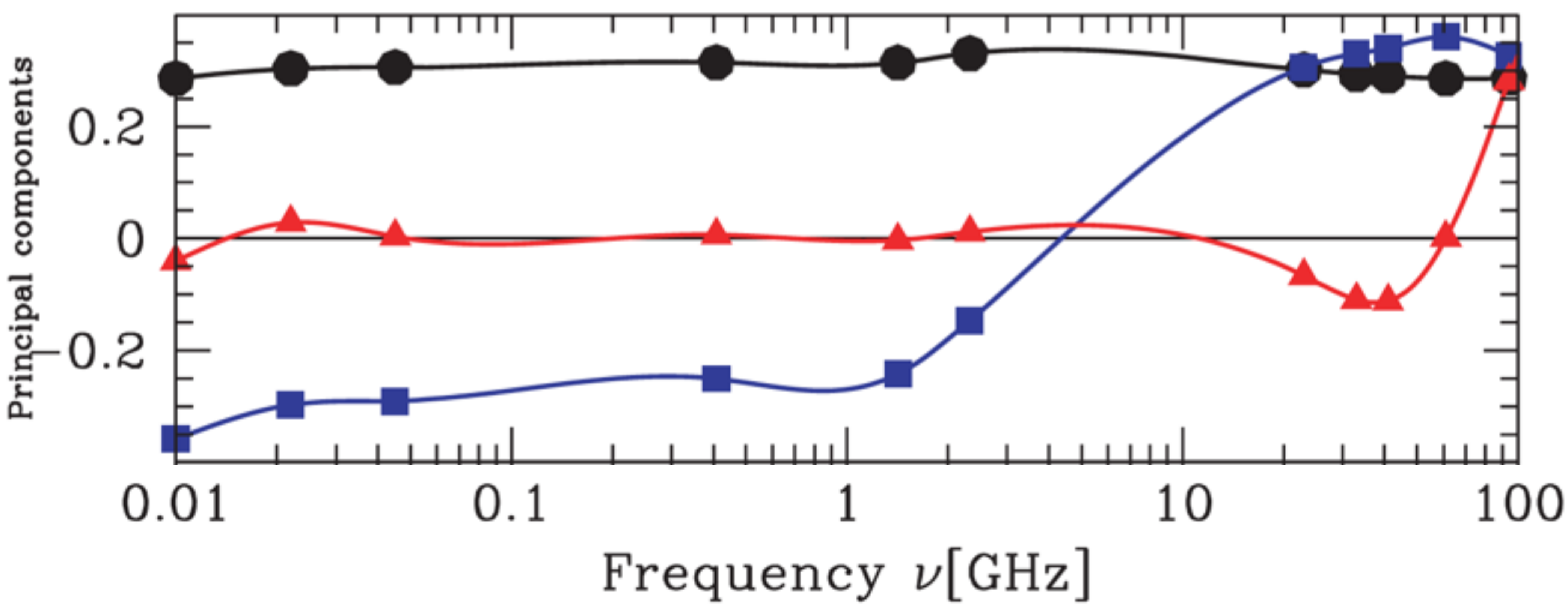
Take a wide selection of survey data…



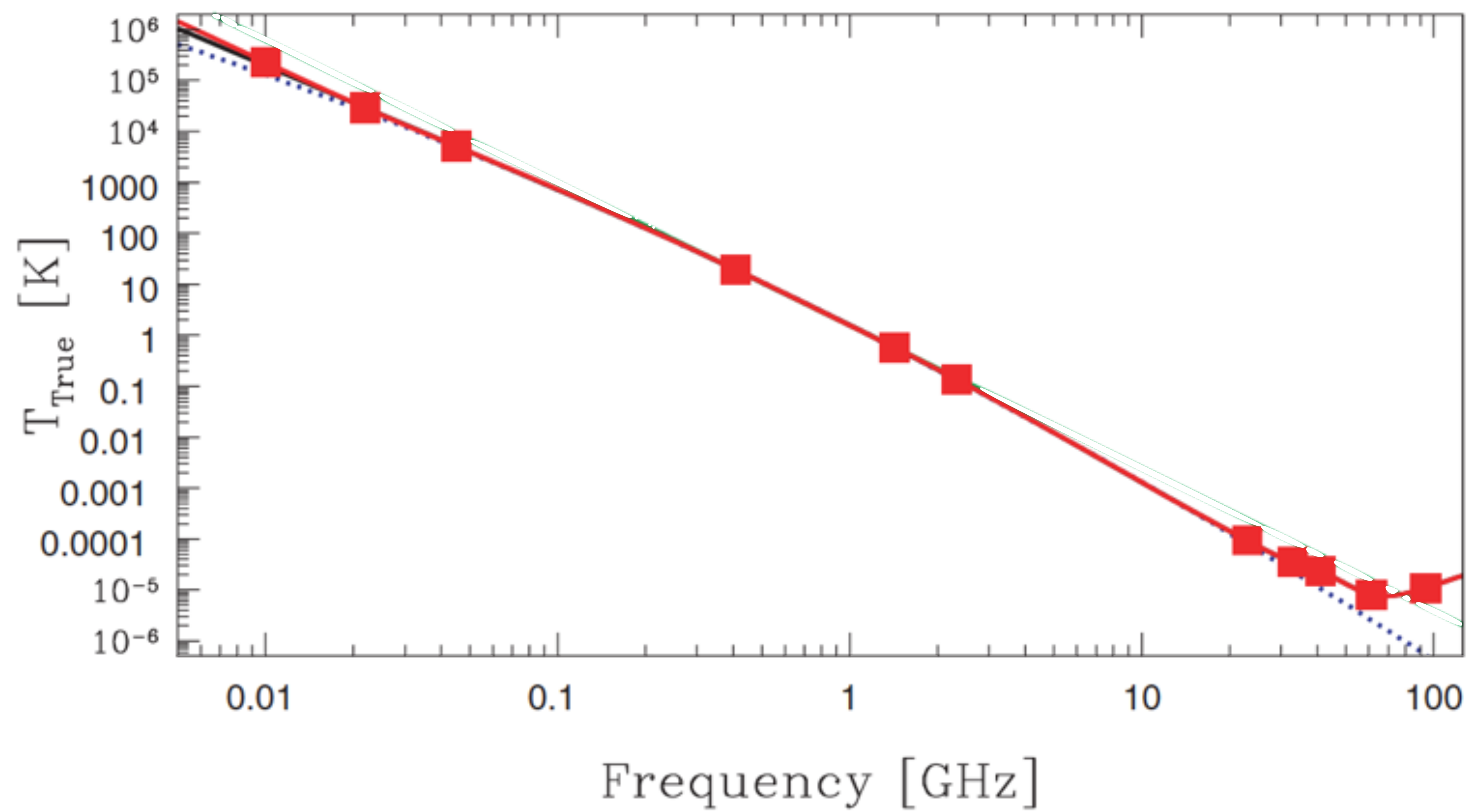
…identify common regions…



…which are then used to train three principal component spectral templates…



…that are used to fit the spectra in every pixel of the sky…



…that are used to fit the spectra in every pixel of the sky…

$$T(\hat{r}, \nu) = m_1(\hat{r})c_1(\nu) + m_2(\hat{r})c_2(\nu) + \ldots$$

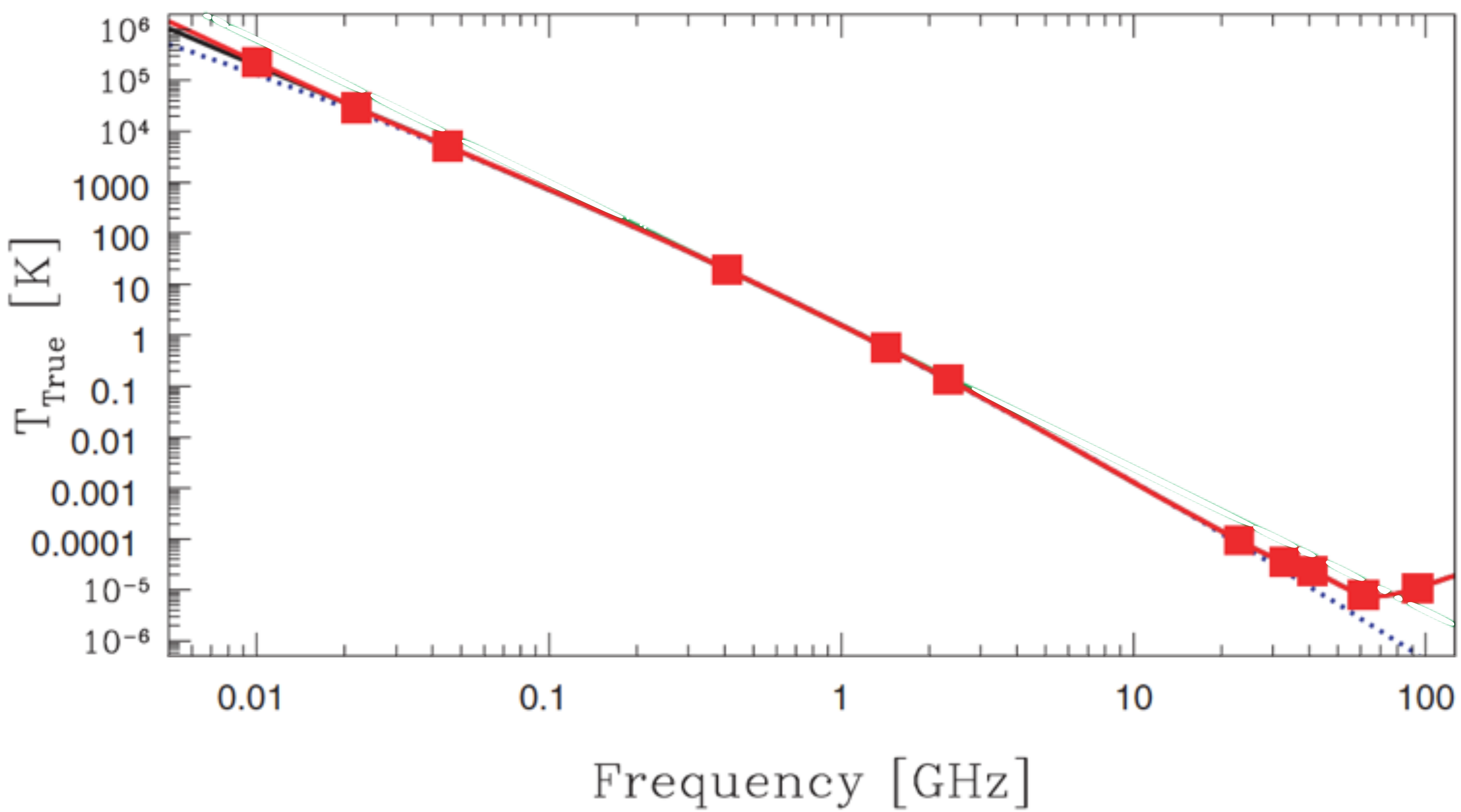
…that are used to fit the spectra in every pixel of the sky…

Given these

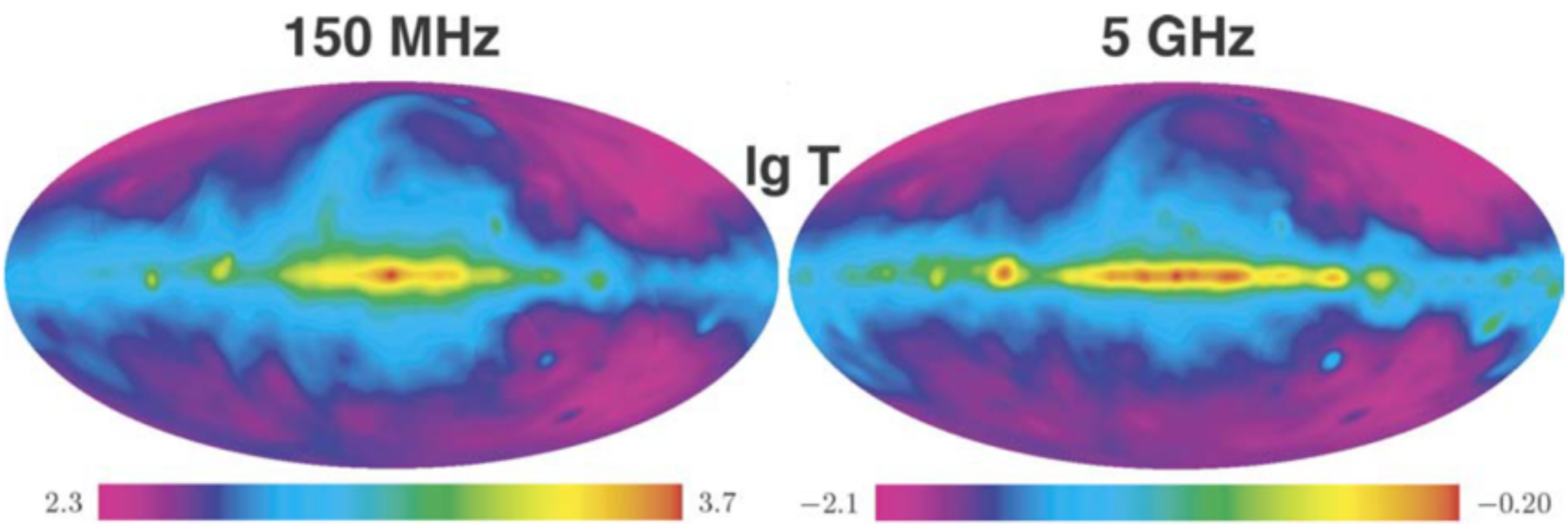
$$T(\hat{r}, \nu) = m_1(\hat{r})c_1(\nu) + m_2(\hat{r})c_2(\nu) + \ldots$$

Solve for these

…that are used to fit the spectra in every pixel of the sky…



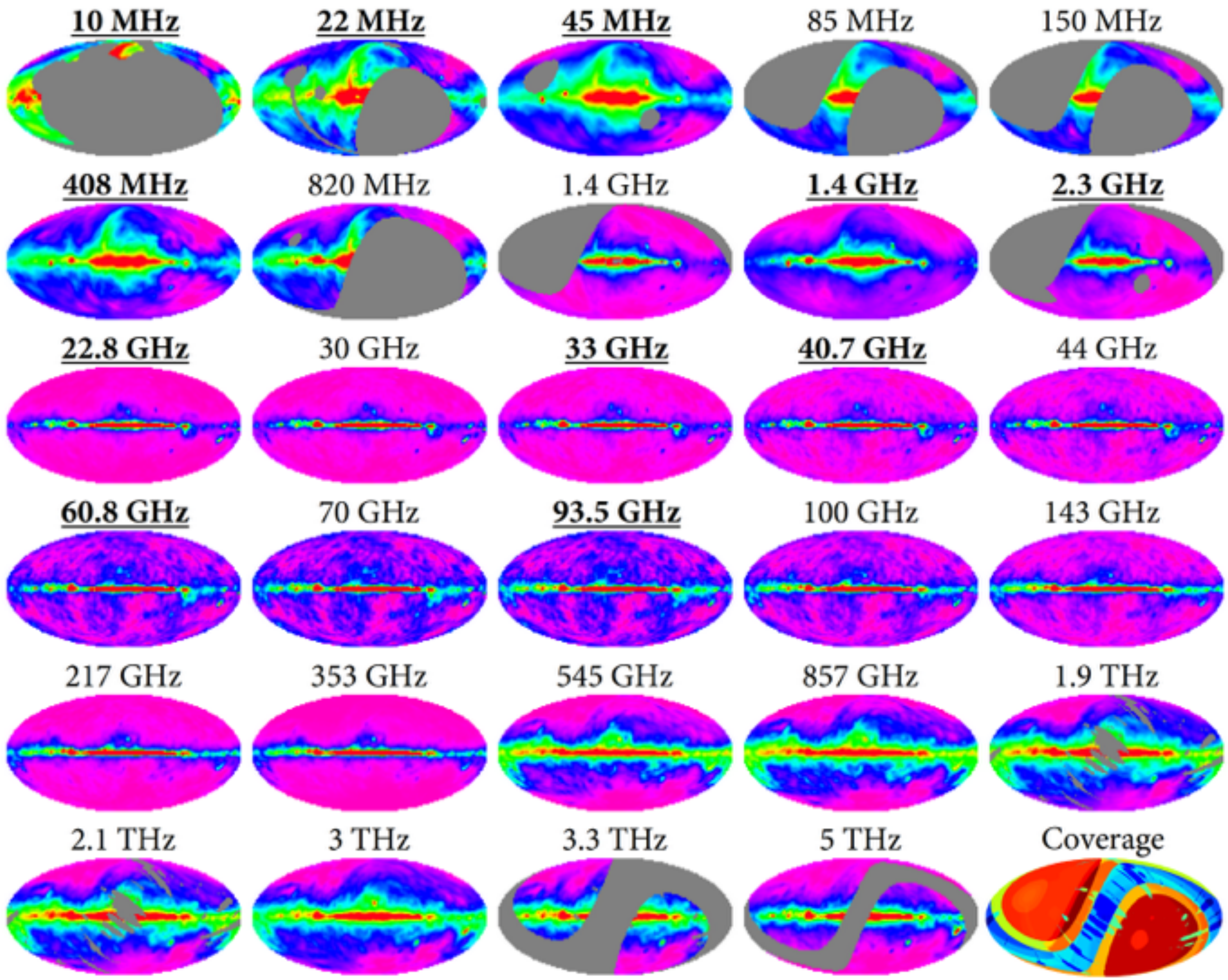
…and are interpolated to produces maps of the sky at "any" frequency



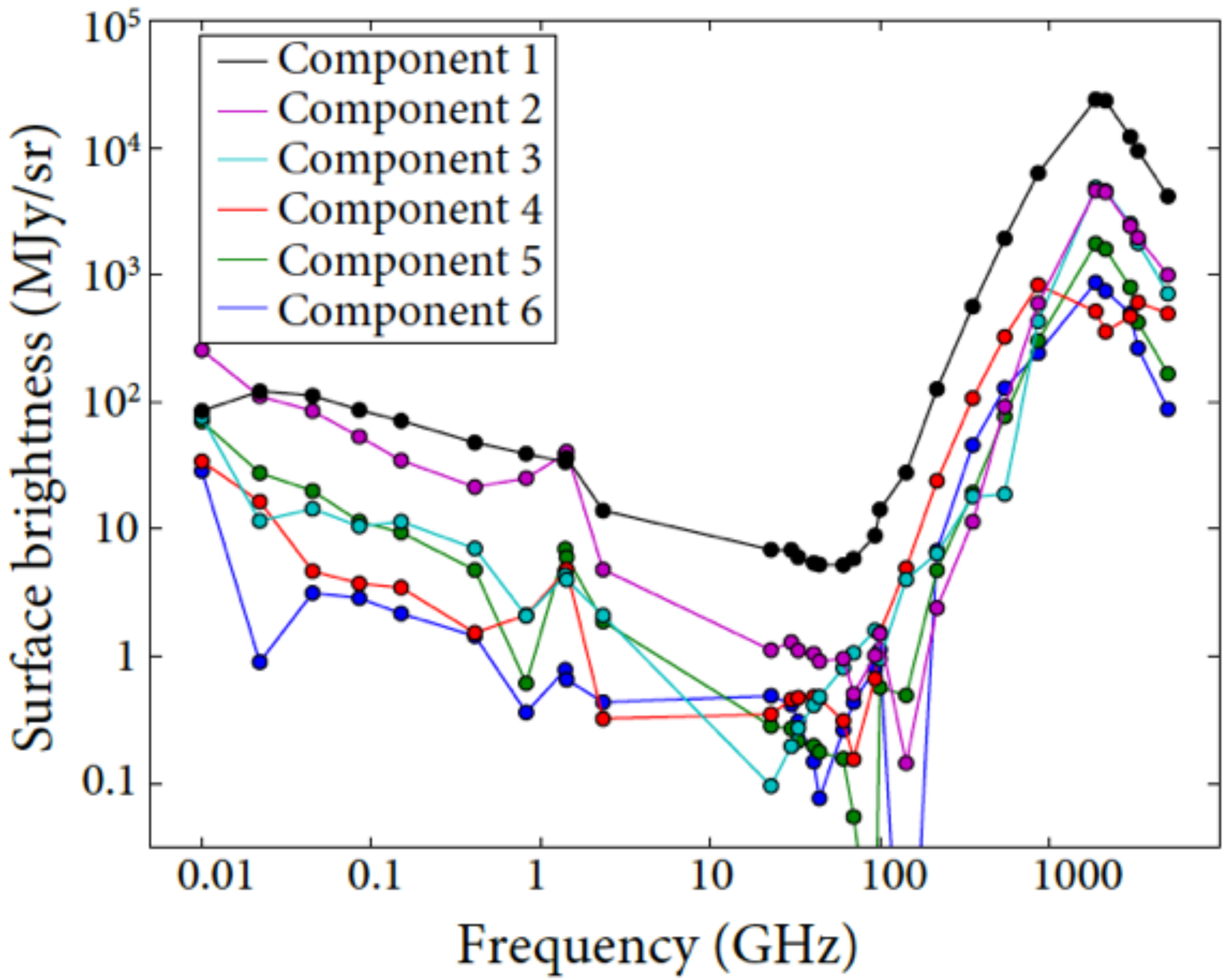
Global Sky Model v2

(Zheng… Kim, AL… et al. 2017, MNRAS 464, 3486)

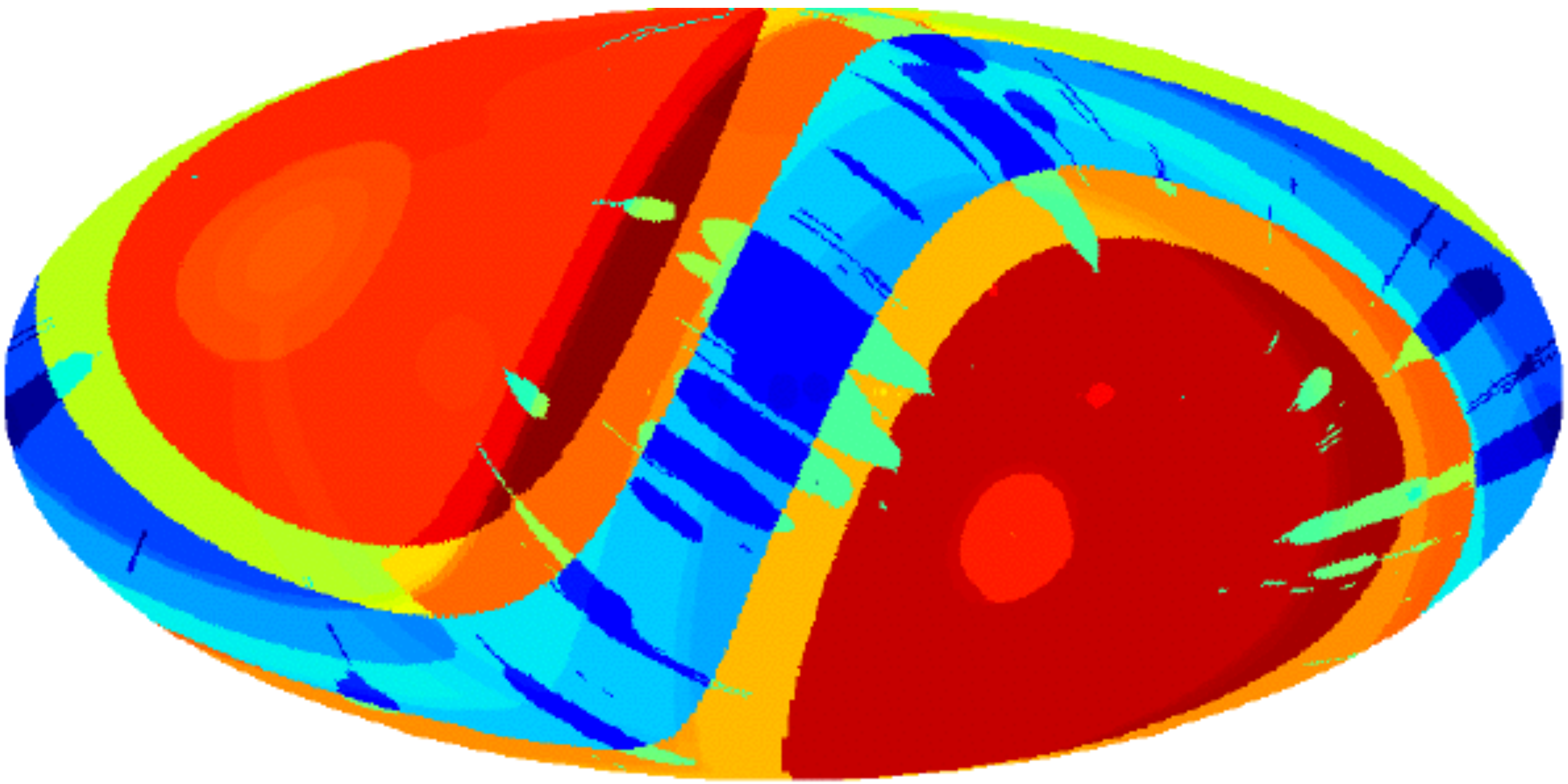
Take an even wider selection of updated maps…



…now using six spectral components…



…and iteratively fitting for spectral and spatial information across the whole sky



…and iteratively fitting for spectral and spatial information across the whole sky

Given these

$$T(\hat{r}, \nu) = m_1(\hat{r}) c_1(\nu) + m_2(\hat{r}) c_2(\nu) + \ldots$$

Solve for these

…and iteratively fitting for spectral and spatial information across the whole sky

Solve for these

$$T(\hat{r},\nu) = m_1(\hat{r})c_1(\nu) + m_2(\hat{r})c_2(\nu) + \ldots$$

Given these

…and iteratively fitting for spectral and spatial information across the whole sky

Given these

$$T(\hat{r}, \nu) = m_1(\hat{r})c_1(\nu) + m_2(\hat{r})c_2(\nu) + \ldots$$

Solve for these

…to derive even even better fits to the data.

Global Sky Model v3

(Kim, AL, Switzer 2017, in prep.)

Why three components?
Why six components?

Too few components: inadequate fits to data
Too many components: overfitting of data

Computing the Bayesian Evidence provides a way to determine the optimal number of principal components to fit

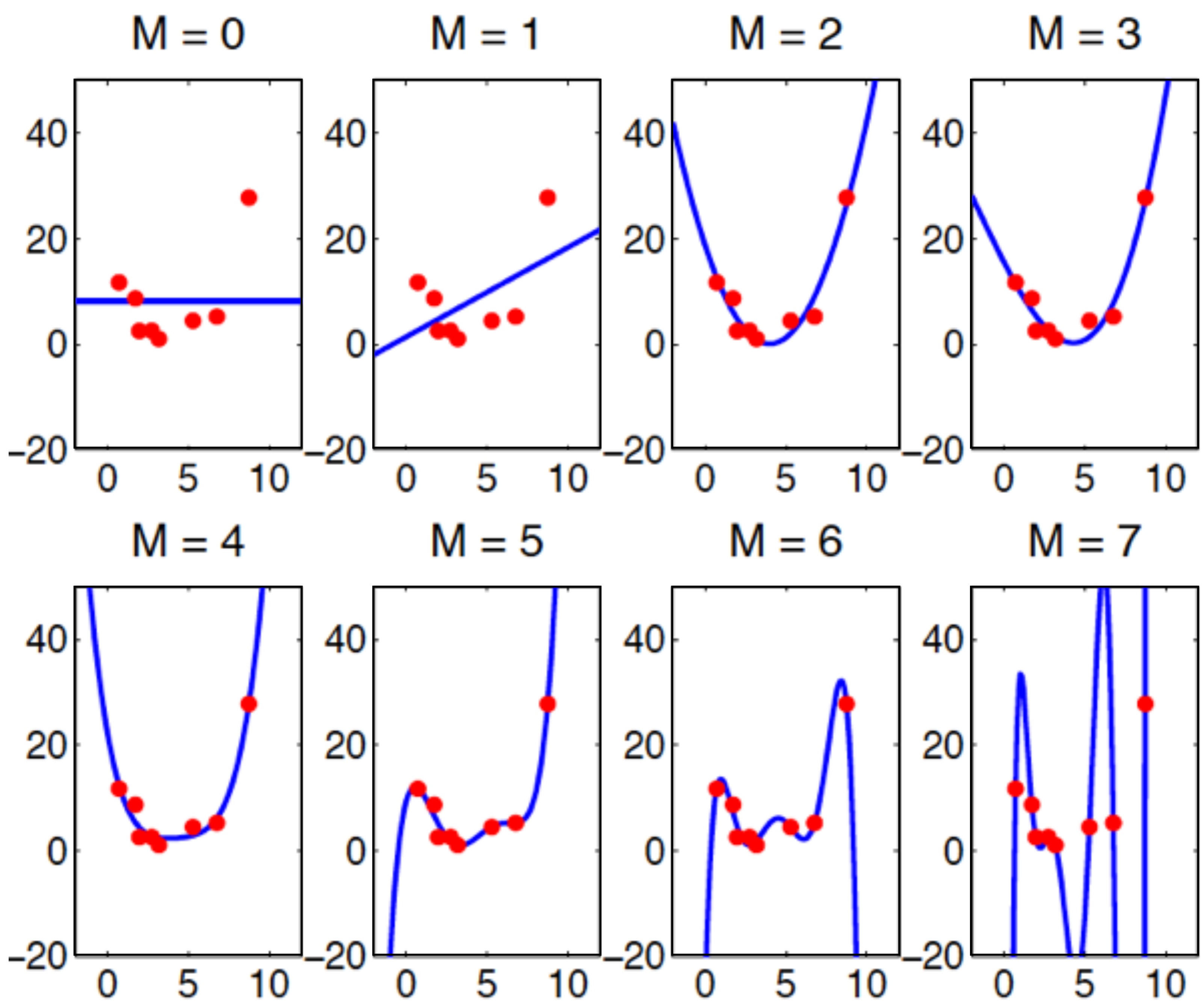
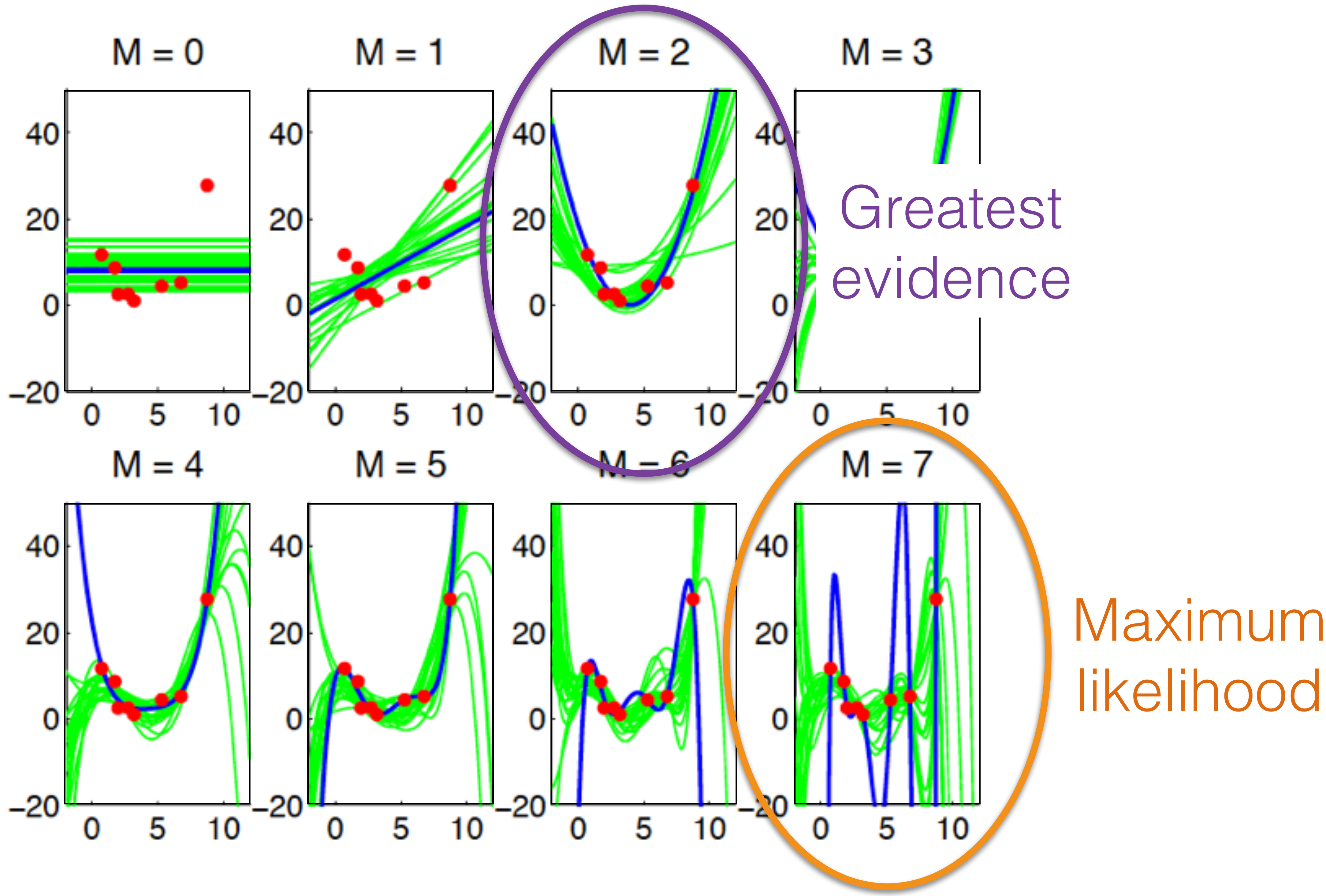


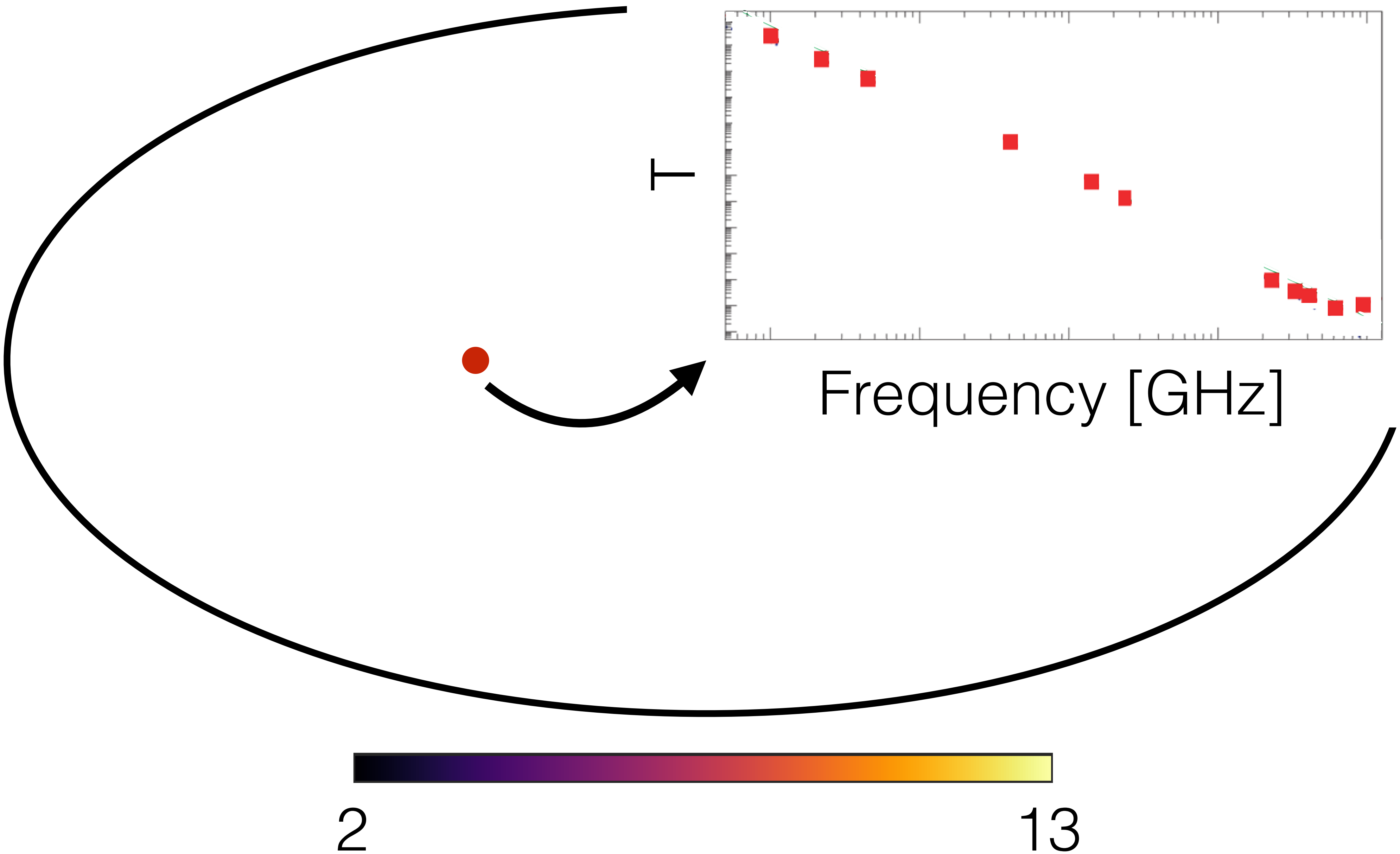
Image credit:
Zoubin Ghahramani

Computing the Bayesian Evidence provides a way to determine the optimal number of principal components to fit

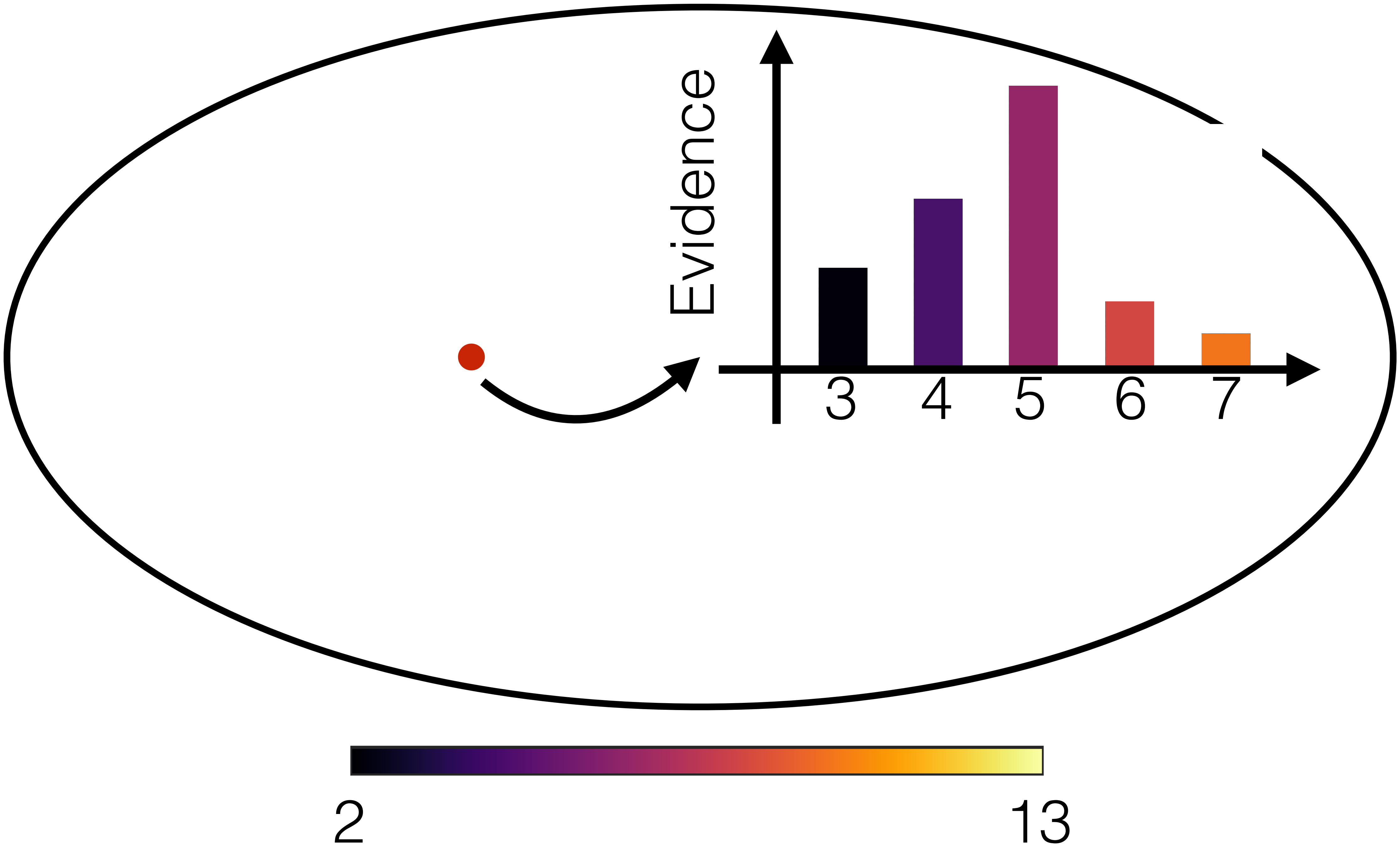
Image credit: Zoubin Ghahramani



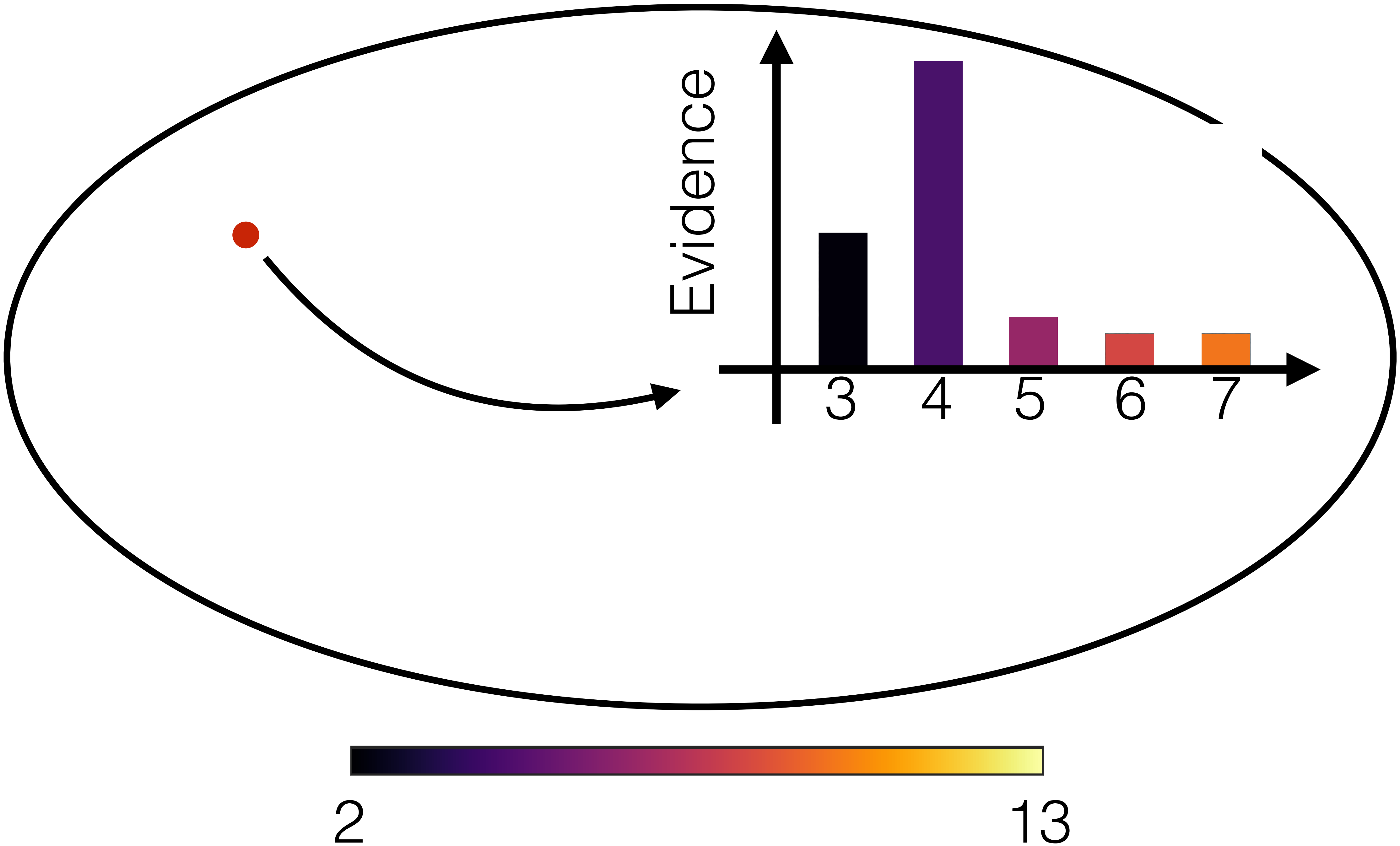
Optimal number of principal components



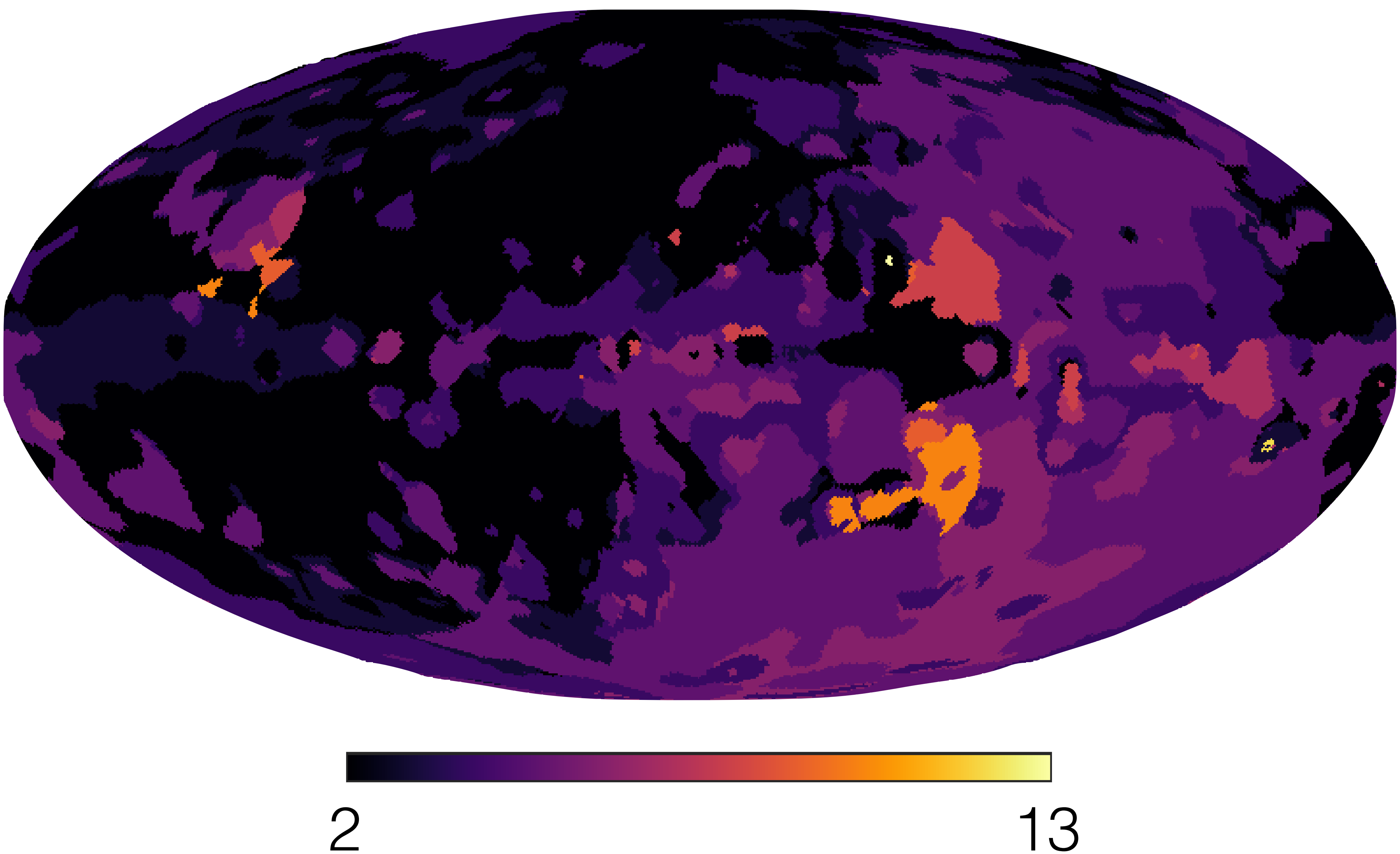
Optimal number of principal components



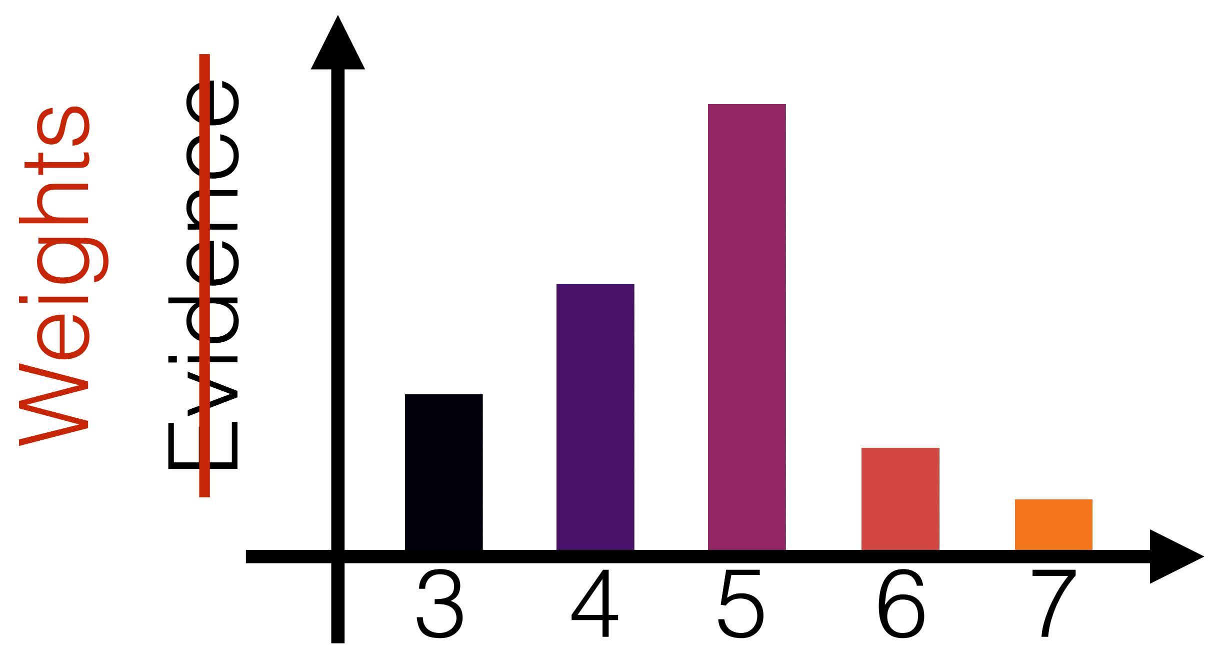
Optimal number of principal components



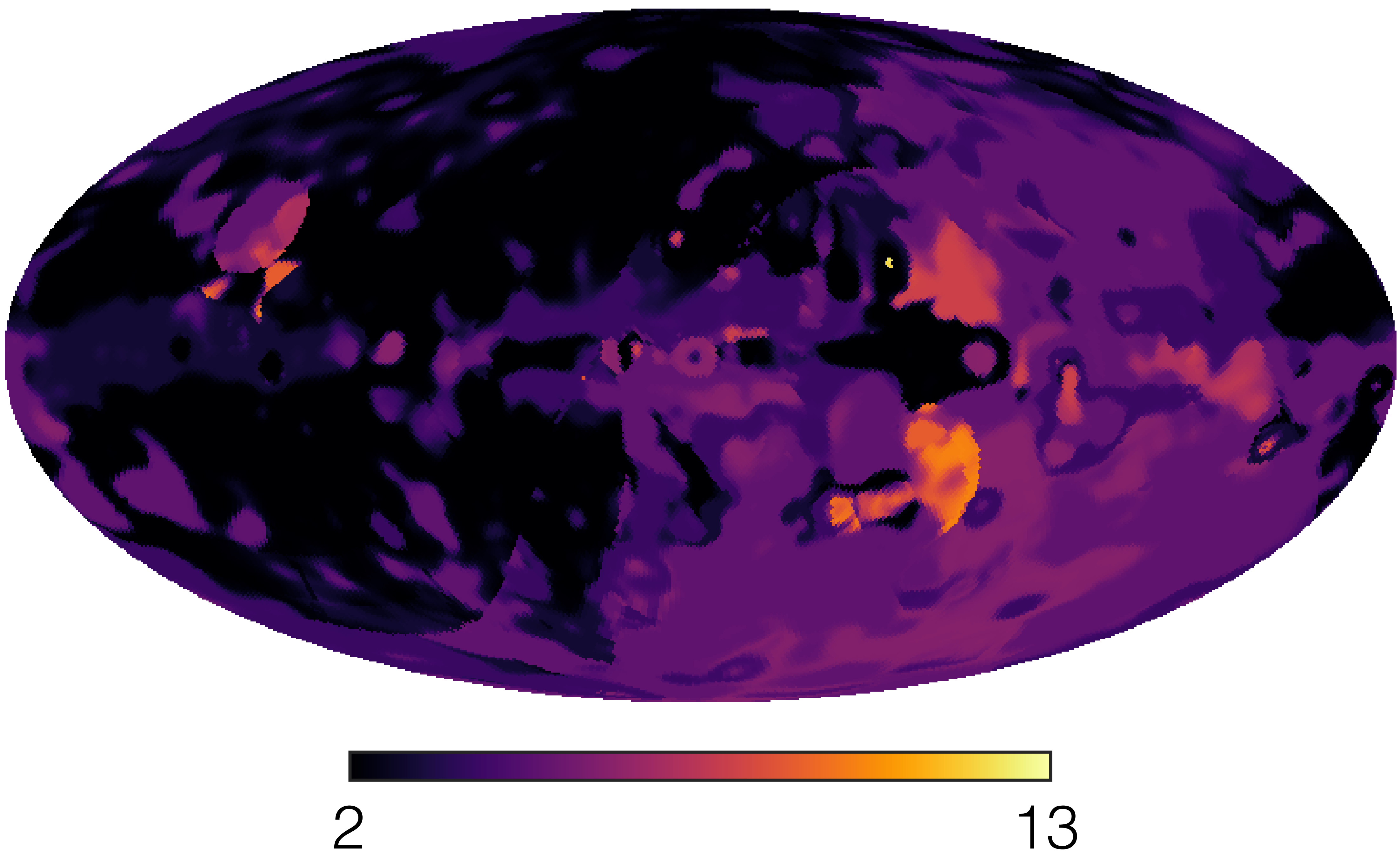
Optimal number of principal components



But why even commit to a model? Use evidence as a weight for constructing hybrid models that are noncommittal to the number of components



Effective number of principal components

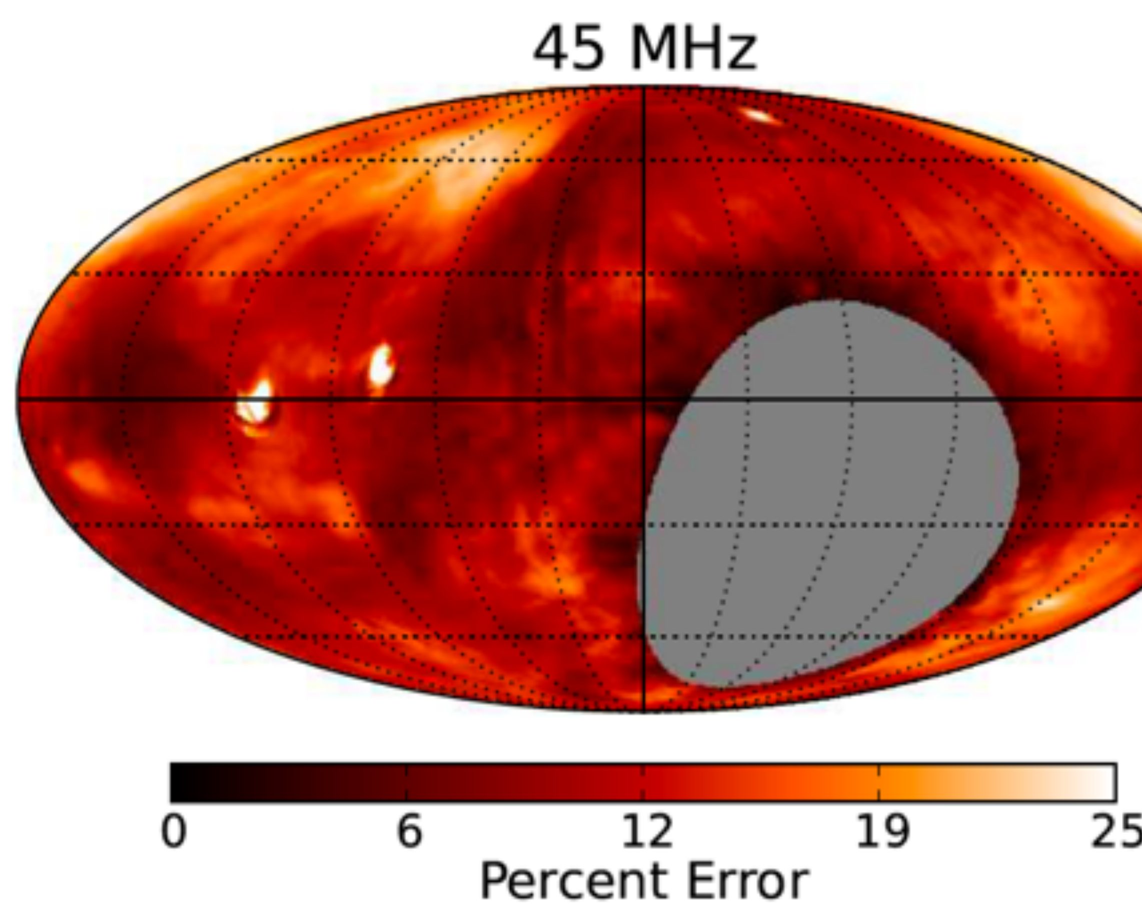


The old versions of the GSM had no error bars!

Solution: construct models for the errors in the input data, and Monte Carlo to get final errors in our predictions

Solution: construct models for the errors in the input data, and Monte Carlo to get final errors in our predictions

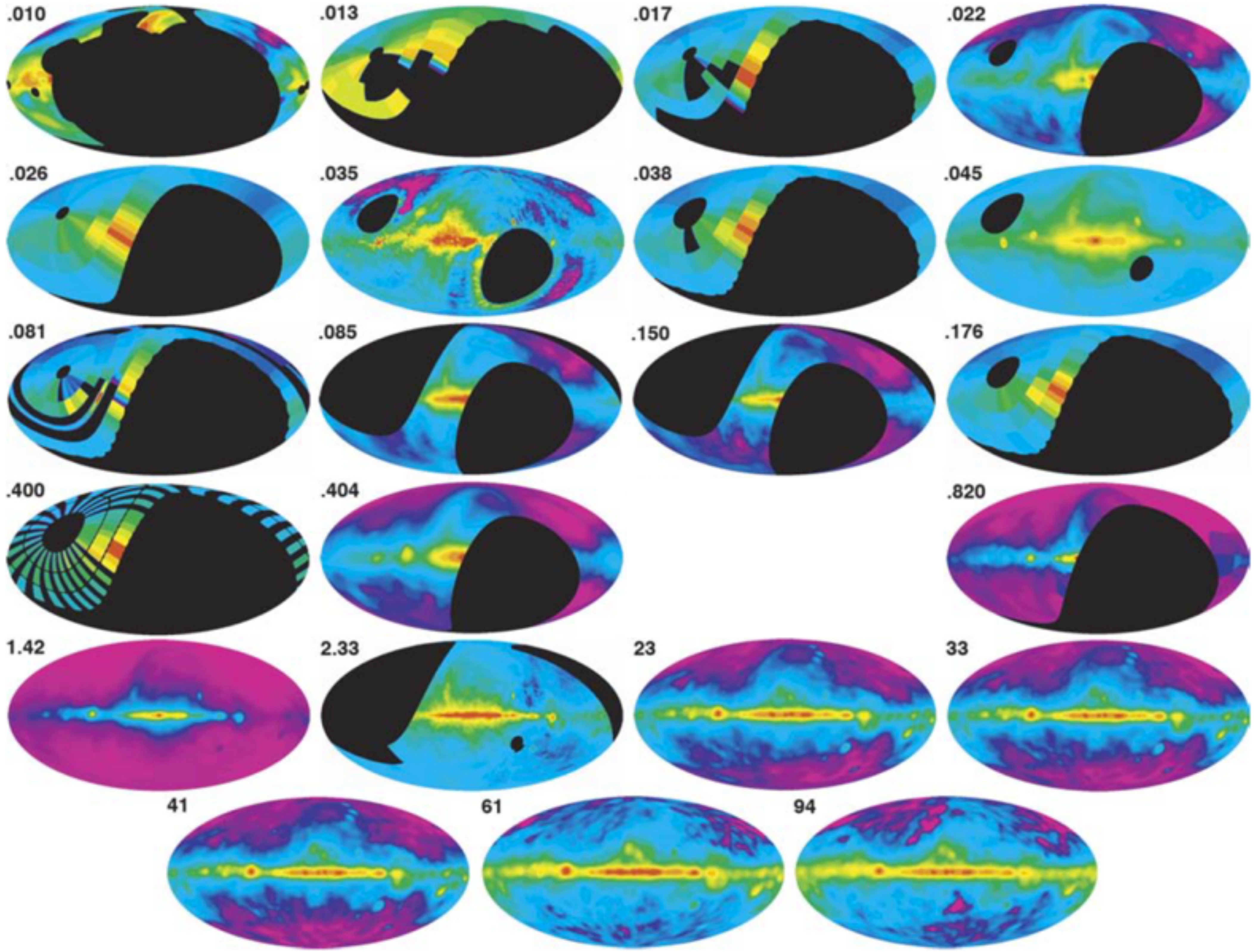
- Where available, use provided estimates of errors and covariances



LWA 74 MHz, Dowell et al. (2017)

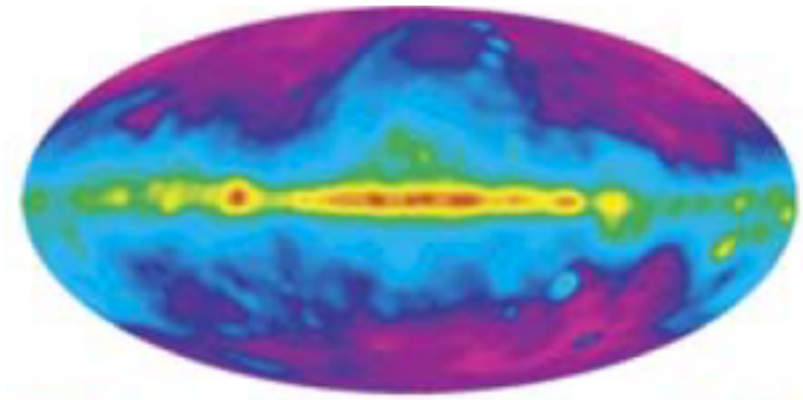
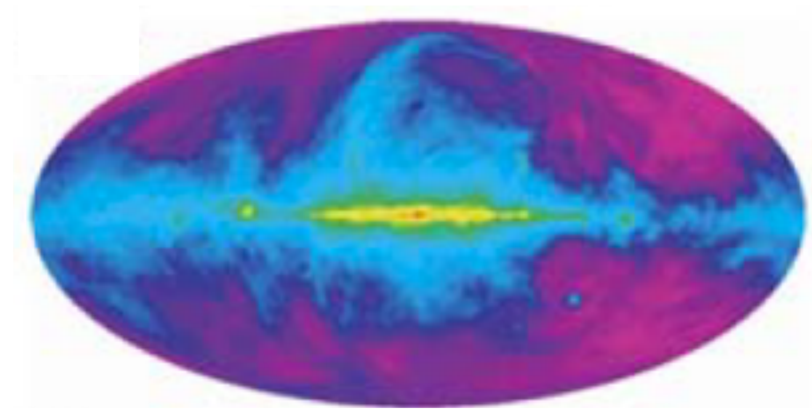
Solution: construct models for the errors in the input data, and Monte Carlo to get final errors in our predictions

- Where available, use provided estimates of errors and covariances
- Errors in the model itself modelled empirically

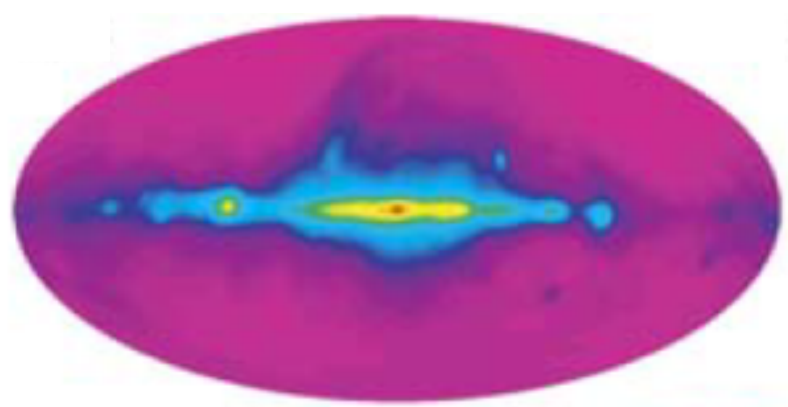


Run model again with an input map removed, making a prediction for the missing map

Prediction **Data**

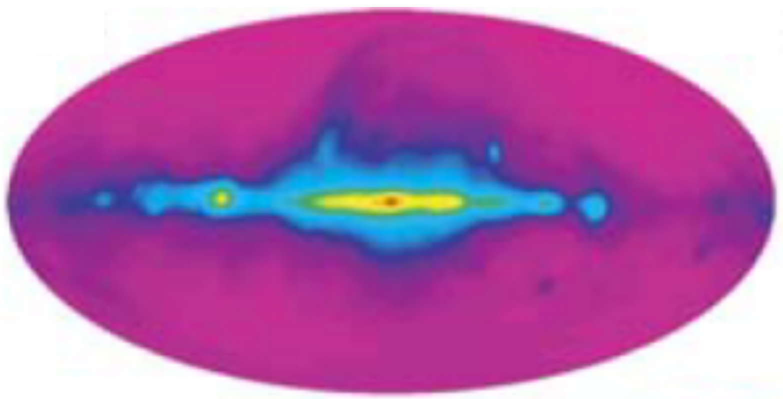
 − 

“Error”

= 

Subtract the new predicted map from the observed data

"Error"

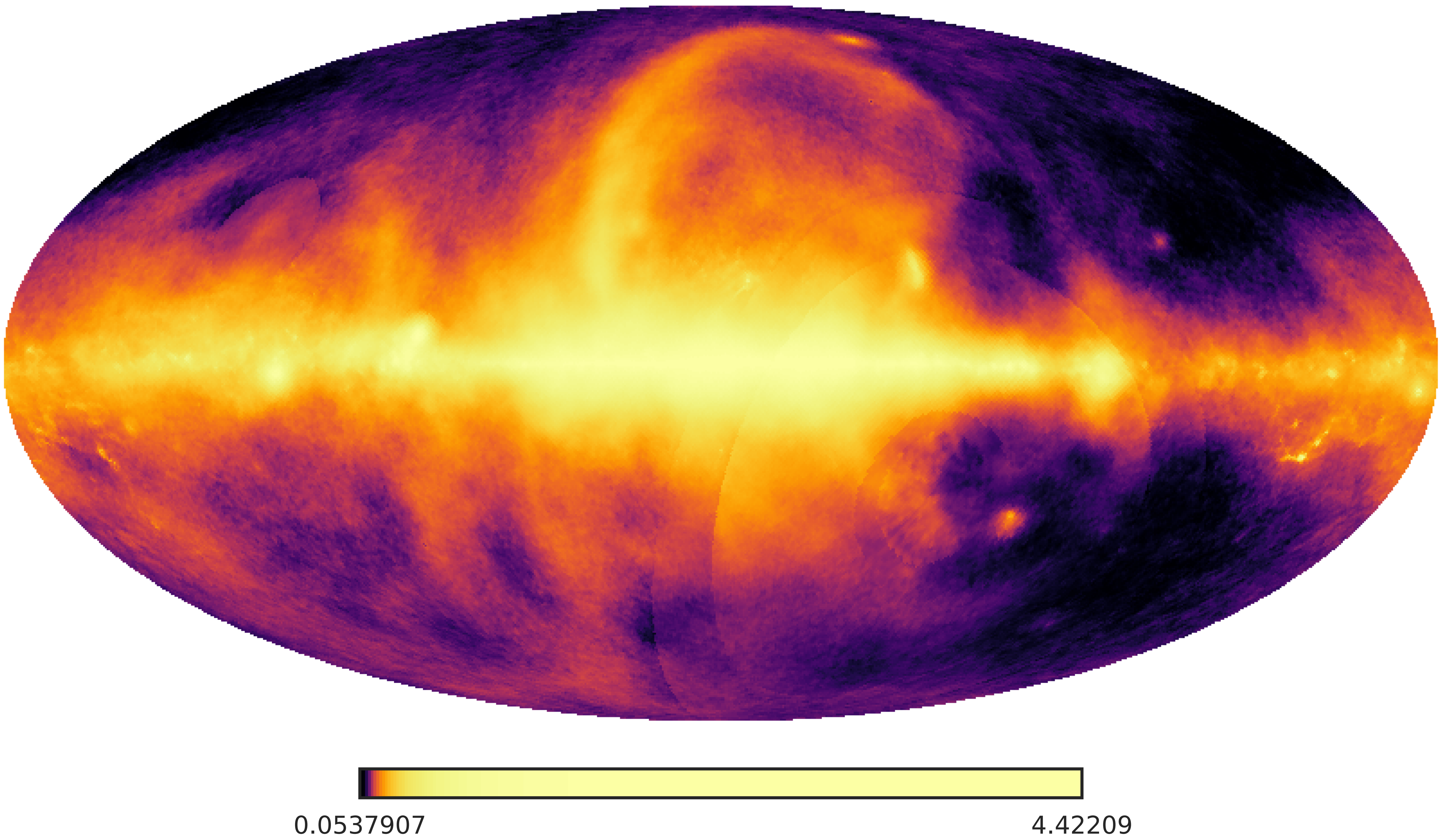


Ansatz for errors in image space: proportional to error map

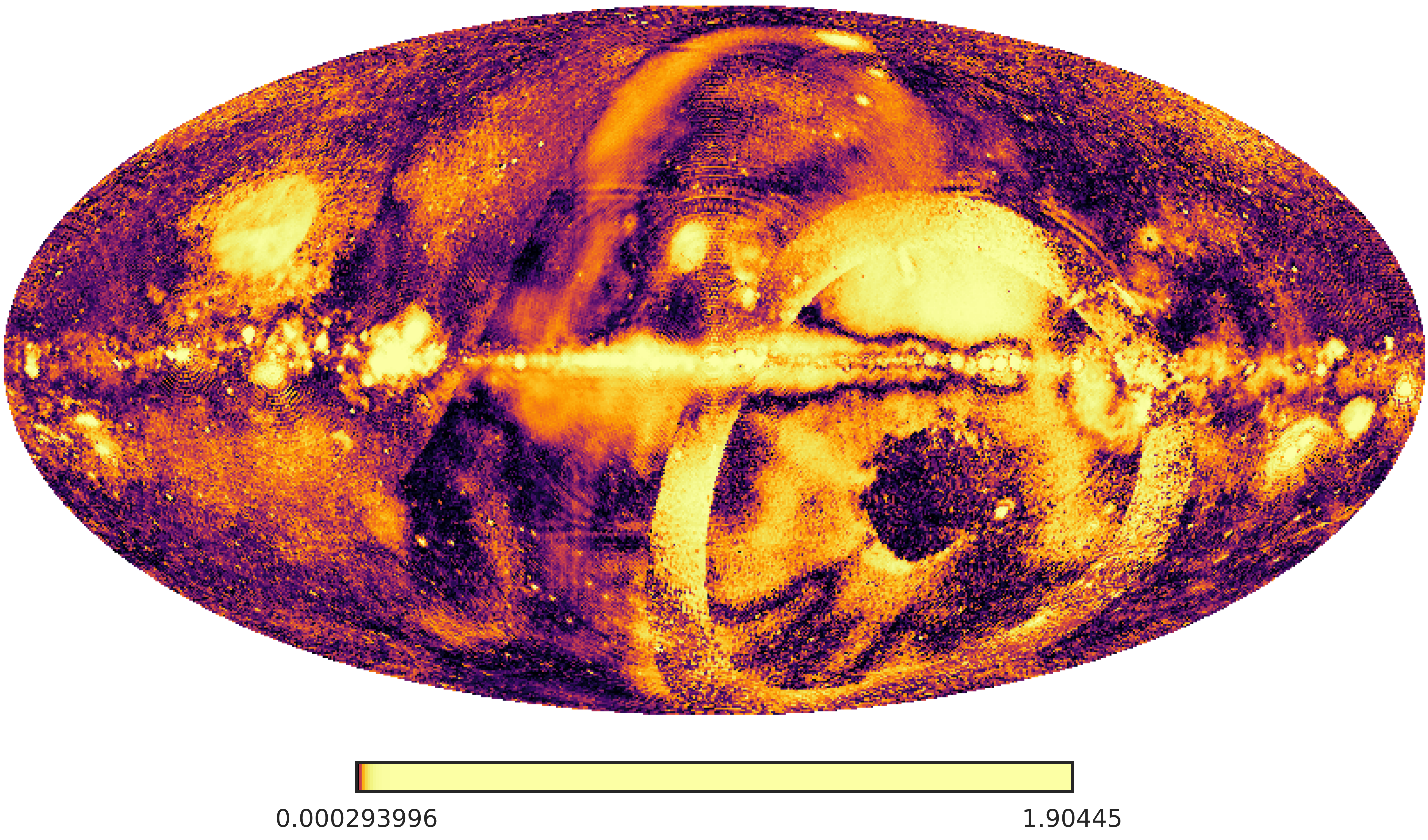
Ansatz for harmonic space: determined by C_l of whitened error map

Error model

An example 408 MHz prediction

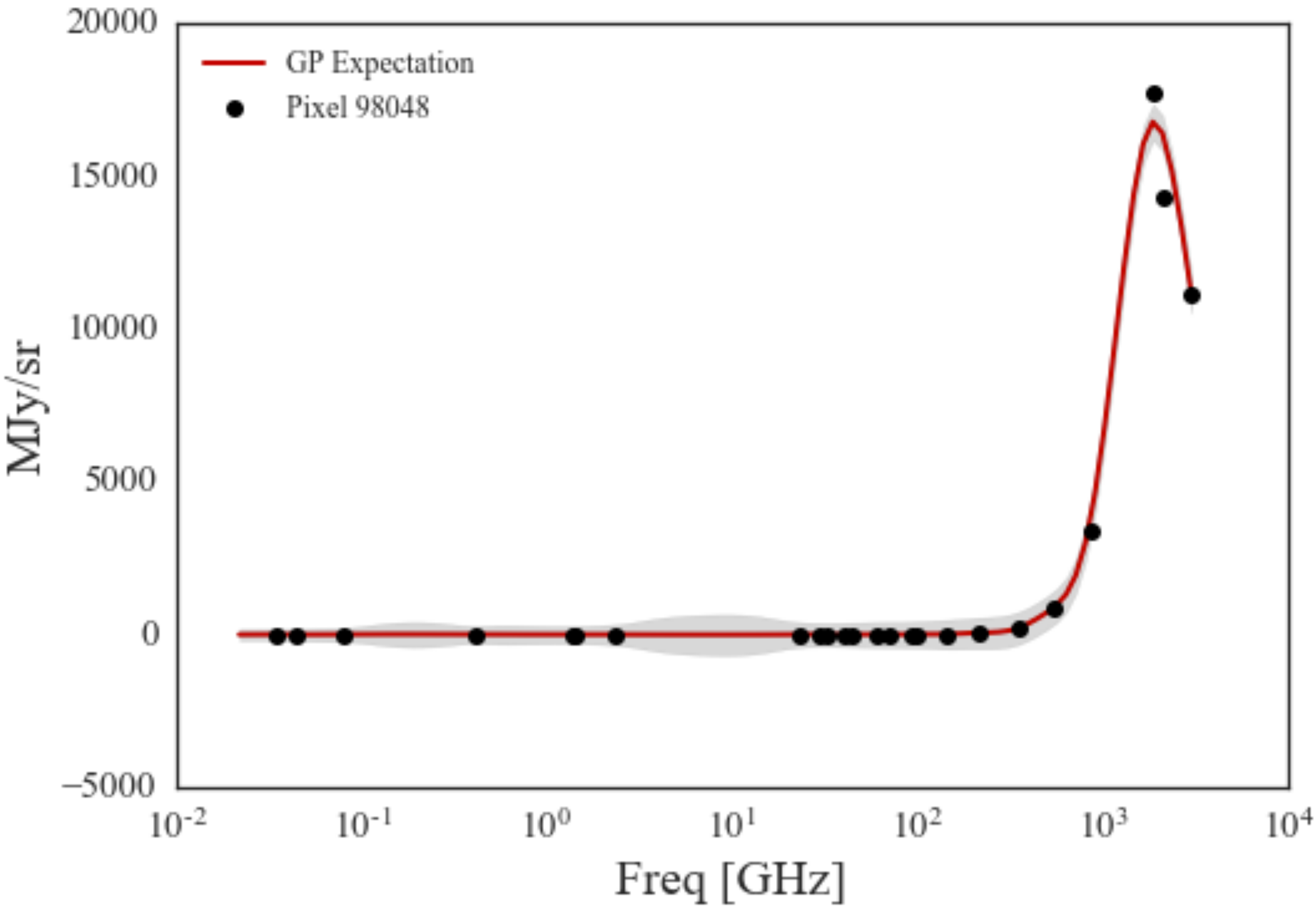


Errors on the 408 MHz prediction



Solution: construct models for the errors in the input data, and Monte Carlo to get final errors in our predictions

- Where available, use provided estimates of errors and covariances
- Errors in the model itself modelled empirically
- Interpolation errors accounted for using Gaussian Process regression.



GP Expectation
Pixel 98048
20000
15000
10000
5000
0
−5000
MJy/sr
10^{-2}
10^{-1}
10^{0}
10^{1}
10^{2}
10^{3}
10^{4}
Freq [GHz]

Lots more coming soon to a Github repo near you!

- Position-dependent number of components.
- Error bars in output maps.
- Framework for incorporating monopole measurements.
- Inclusion of new map data.

Lots more coming soon to a Github repo near you!

- Position-dependent number of components.

End goal: a publicly hosted, self-updating, best-guess model of the sky