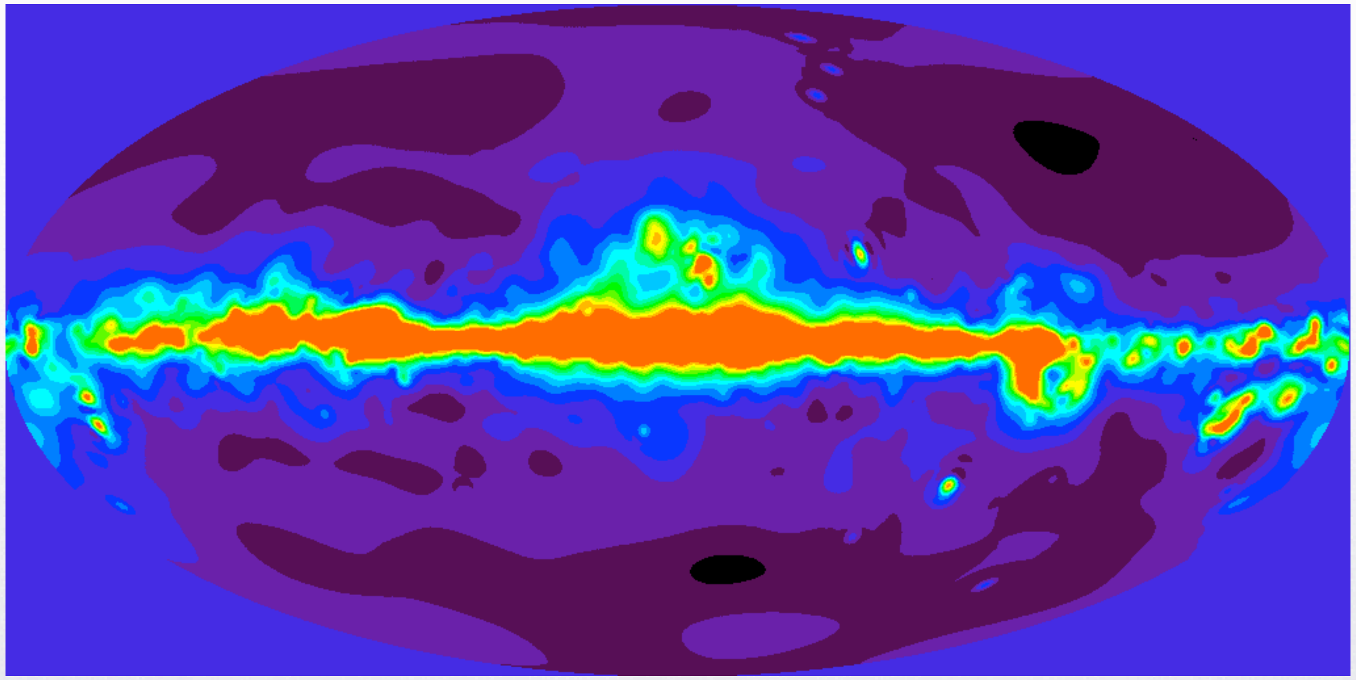
Nonparametric Estimation of CMB Foreground Emission

P. E. Freeman, C. R. Genovese, & L. Wasserman
Department of Statistics - Carnegie Mellon University



"The main goal of this paper is to remove foregrounds, not to understand or model them."

- Tegmark, de Oliveira-Costa, & Hamilton (2003)

Context: Simultaneous Inference

- Nonparametric foreground estimation is one piece of a simultaneous inference scheme:

1. Assume the CMB to be a realization of a Gaussian random field.

2. Nonparametrically estimate the unknown foreground function f given a CMB power spectrum $\{C_l\}$.

3. Use residual maps to inform movement from $\{C_l\} \rightarrow \{C_l'\}$.

4. . . . and determine the $\{C_l\}$ that maximizes the likelihood.

Nonparametric Statistics: Basics

- The goal is to make sharp inferences about an unknown function with minimal assumptions.
- Useful when a parametric model for the function is complex and a simpler nonparametric model may have similar inferential power.
- The name is a misnomer: nonparametric methods feature infinite-dimensional parameters (e.g., basis coefficients), that are further resolved with increasing data.

Nonparametric Foreground Estimation

- Our nonparametric regression problem:

$$T(\theta, \phi) \;=\; f(\theta, \phi) + \epsilon(\theta, \phi; C_l)$$

- Here, $f(\theta,\Phi)$ is the unknown, true foreground, while $\epsilon(\theta,\Phi;C_l)$ is the “noise,” consisting of both CMB signal and pixel noise.

- We expand f into the needlet basis (or frame):

$$f \;=\; \sum_{jk} \beta_{jk} \Psi_{jk}$$

Needlet Function

- The needlet function is:

$$\Psi_{jk}(\theta,\phi) \;=\; \sqrt{2\pi\lambda_{jk}}\sum_{lm} b\left(\frac{l}{B^j}\right) Y^*_{lm}(\theta,\phi)Y_{lm}(\xi_{jk})$$

where

(j,k): frequency and spatial pixel indices

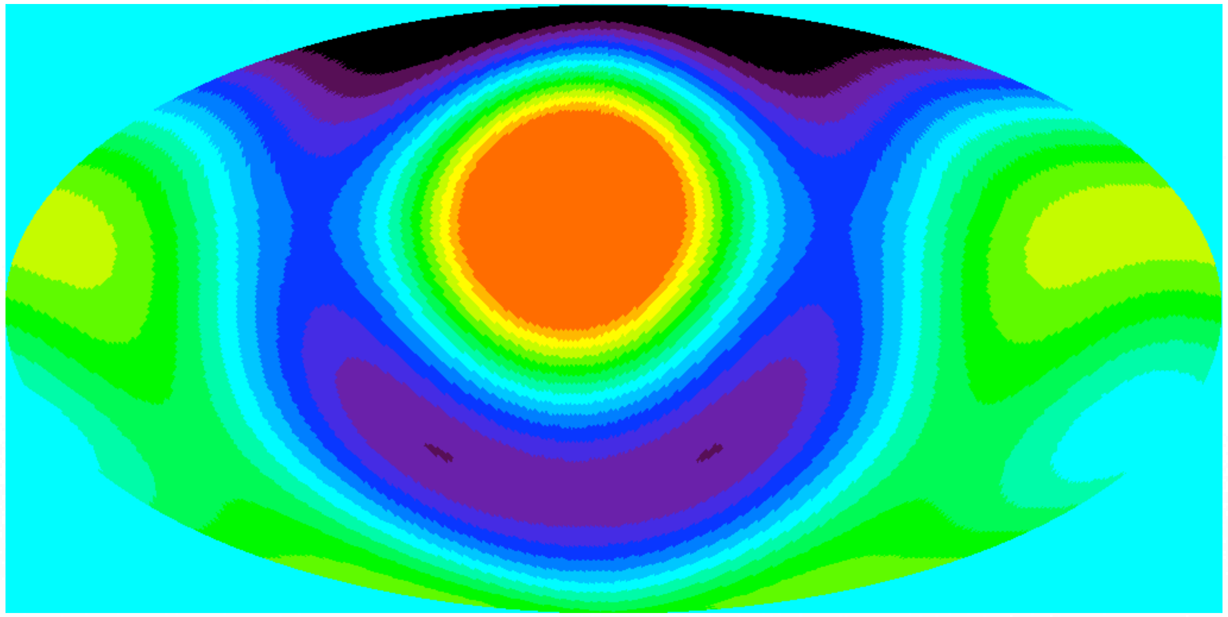
(l,m): multipole and phase indices

(ξ_{jk},λ_{jk}): cubature points and weights (GLESP; e.g., Doroshkevich et al. 2005)

B: multipole localization parameter ($l \in [B^{j-1}, B^{j+1}]$)

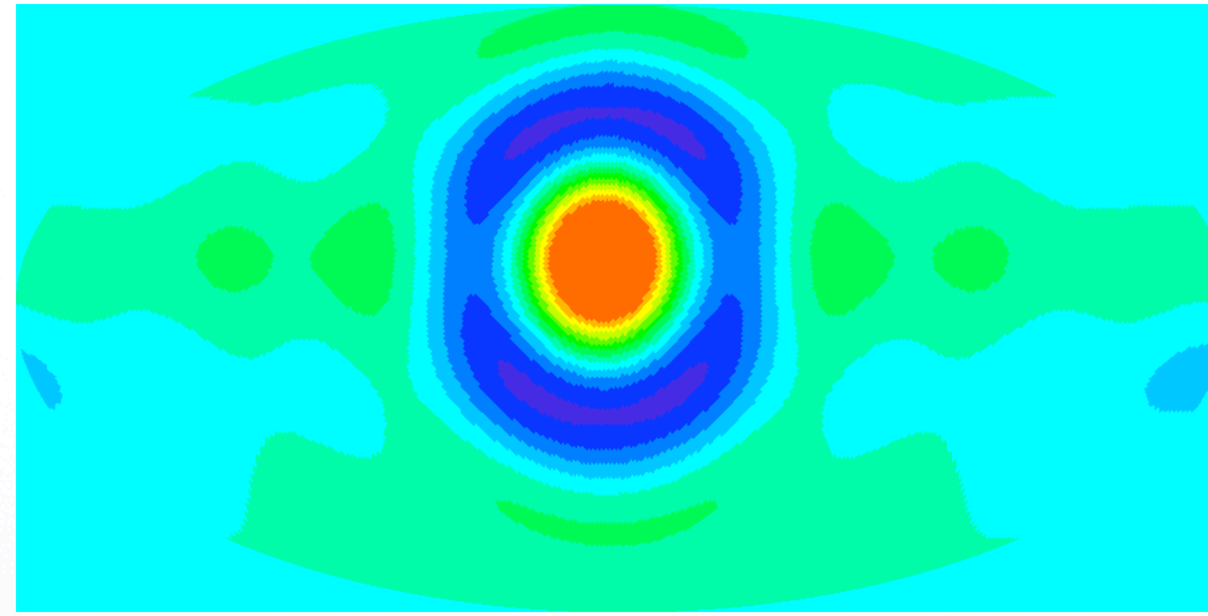
$b(\cdot)$: window function (e.g., Marinucci et al. 2007)

$Y_{lm}(\cdot)$: spherical harmonic function

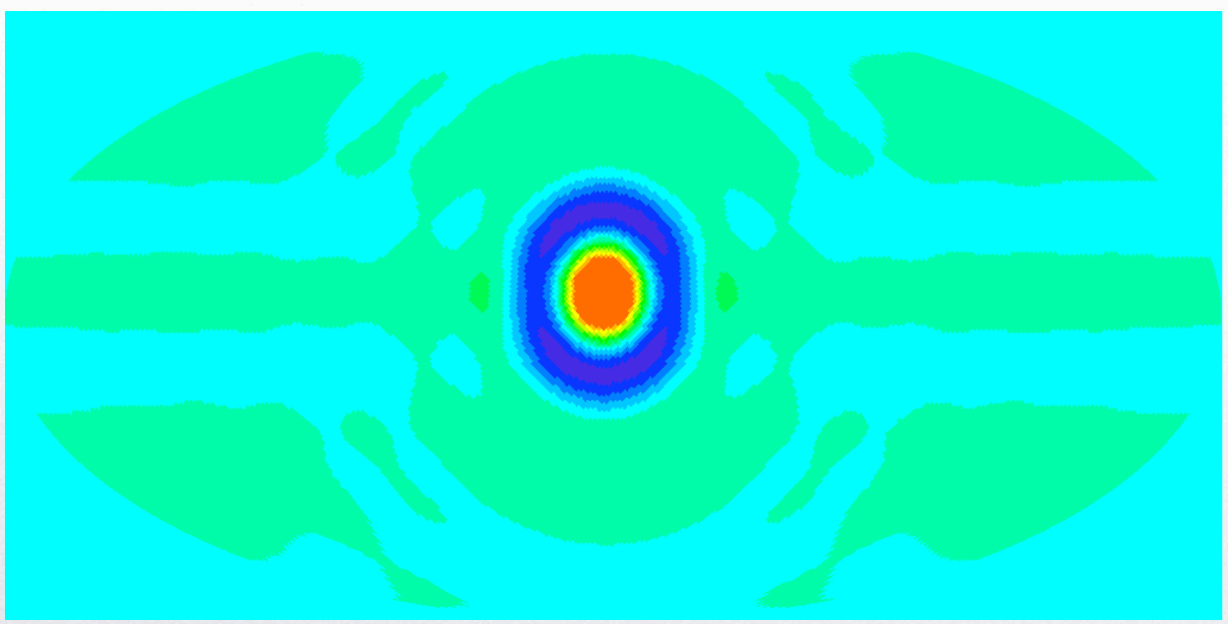


$j = 1 \qquad l \in [1, 4]$

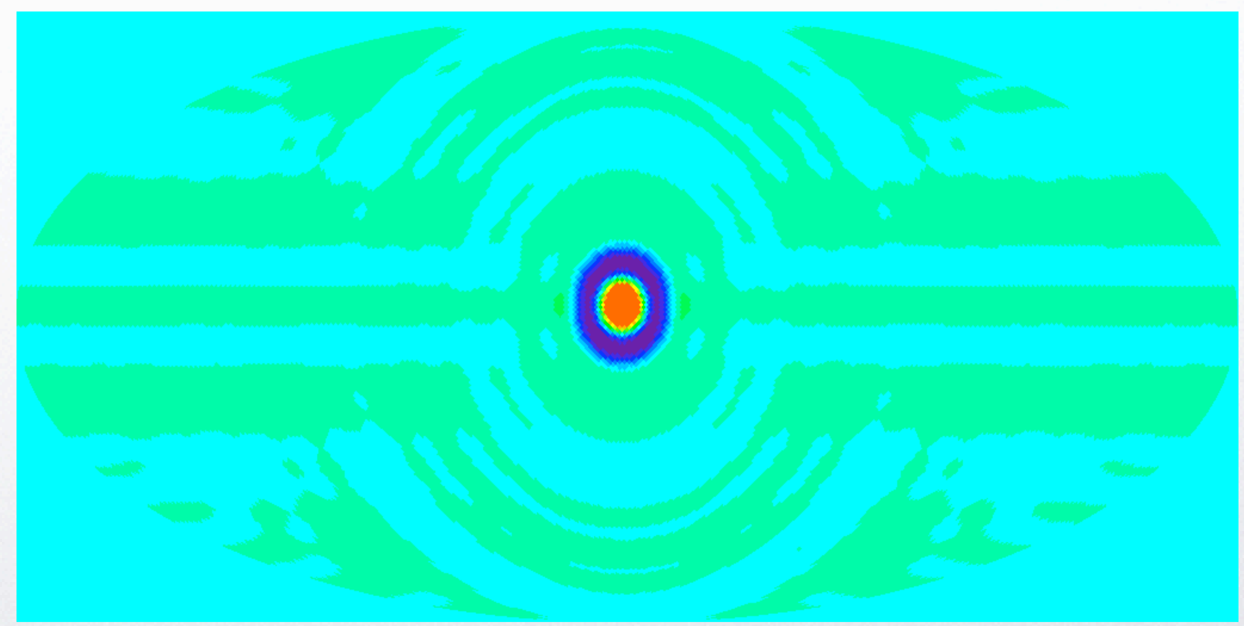
$B = 2$



$j = 2 \qquad l \in [2, 8]$



$j = 3 \qquad l \in [4, 16]$



$j = 4 \qquad l \in [8, 32]$

Algorithm

- Decompose full-sky map T into coefficients $\hat{a}_{lm}$.
- Compute needlet coefficients:

$$\hat{\beta}_{jk} = \sqrt{2\pi\lambda_{jk}} \sum_{lm} b\left(\frac{l}{B^j}\right) Y_{lm}(\xi_{jk}) \hat{a}_{lm}$$

- Estimate shrinkage coefficients $\hat{\mu}_{jk}(\hat{\beta})$
- Transform back to spherical harmonic space:

$$\hat{a}_{lm}^{FG} = \sum_{jk} \sqrt{2\pi\lambda_{jk}} b\left(\frac{l}{B^j}\right) Y_{lm}^{*}(\xi_{jk}) \hat{\mu}_{jk}(\hat{\beta}) \hat{\beta}_{jk}$$

Tested Shrinkage Procedure

- We test shrinkage via hard thresholding:

$$\hat{\mu}_{jk}(\hat{\beta}) \;=\; \begin{cases} 1 & \text{if } |\hat{\beta}_{jk}/\sigma_{\beta_{jk}}| \geq t_j \\ 0 & \text{if } |\hat{\beta}_{jk}/\sigma_{\beta_{jk}}| < t_j \end{cases}$$

where σ is a function of the assumed $\{C_l\}$.

- A simplistic shrinkage procedure, but: if foreground is nearly sparse in the needlets representation, we may achieve identifiability.

Tested Shrinkage Procedure

- For a given $\{C_l\}$, we would estimate the optimal thresholds t_j by minimizing the mean-squared error (MSE), estimated by computing the bias B and variance V of estimates $\hat{f}$:

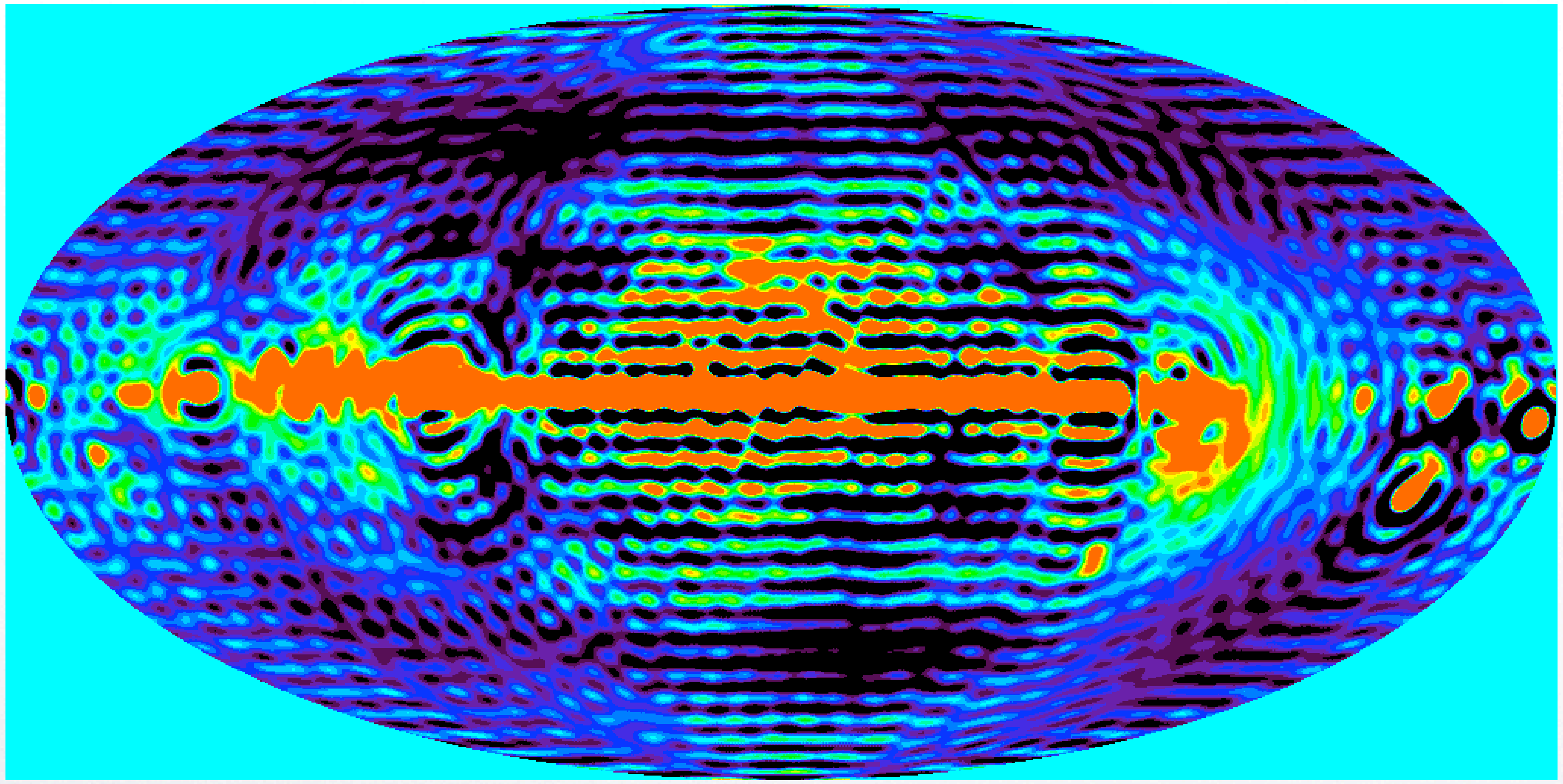
$$\mathrm{MSE}(f, \hat{f}, t_j) \;=\; B^2(f, \hat{f}, t_j) + V(f, \hat{f}, t_j)$$

- This requires computationally intensive simulations.
- Instead we simulate the distribution of the maximum of $\hat{\beta}_{jk}/\sigma_{\beta_{jk}}$; then $t_j = t_j(\alpha)$.

Results: WMAP 5YR Data

- We estimate diffuse foregrounds in all five WMAP bands, given the best-fit ΛCDM C_l.
  - We co-add data in each band over DA and year with inverse-variance noise weighting.
  - We set $B = 2$ and examine frequencies $j = [1,\ldots,6]$. (We use $j = 0$ needlets to remove the dipole components of the foregrounds.) Thus we have full needlet function coverage over the multipole range $l = 2 \rightarrow 64$.
  - We ignore pixel noise in the selection of thresholds.

Caveat: Strong Radio Point Sources



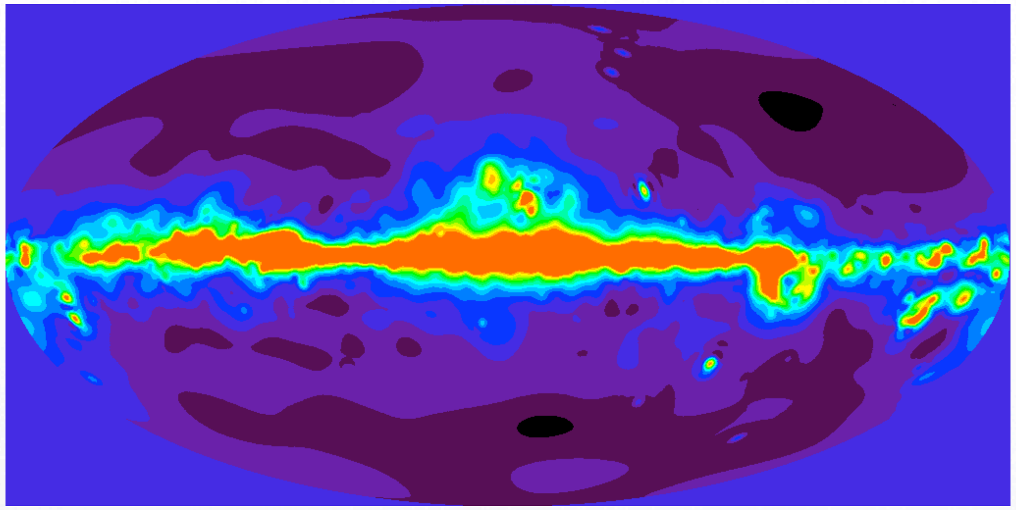
Q Band

Issue:
our current
statistical procedure
is sensitive to strong
radio point sources.

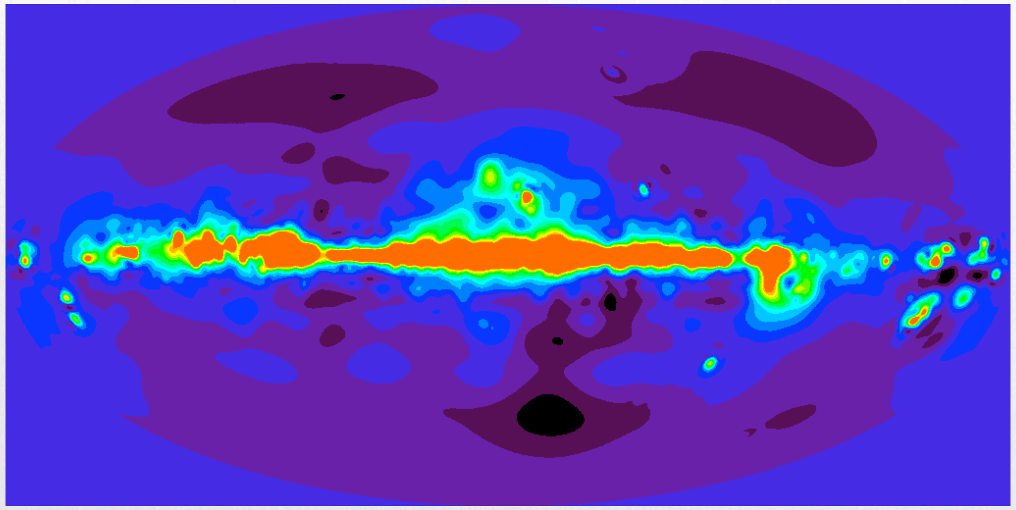
(cf. Abrial et al., arXiv:0804.1295)

Foreground Estimates

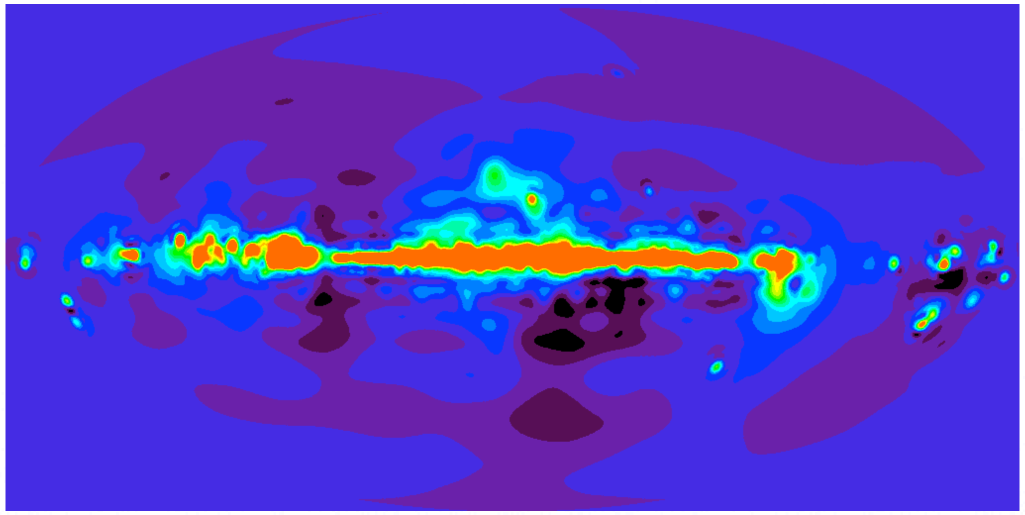
(Standardized T: $-1\sigma \rightarrow 3\sigma$)



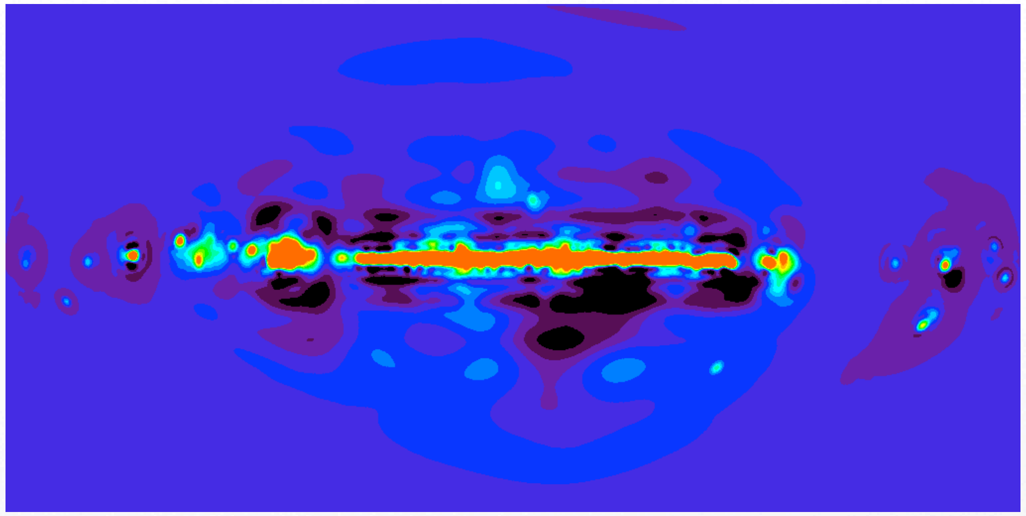
K



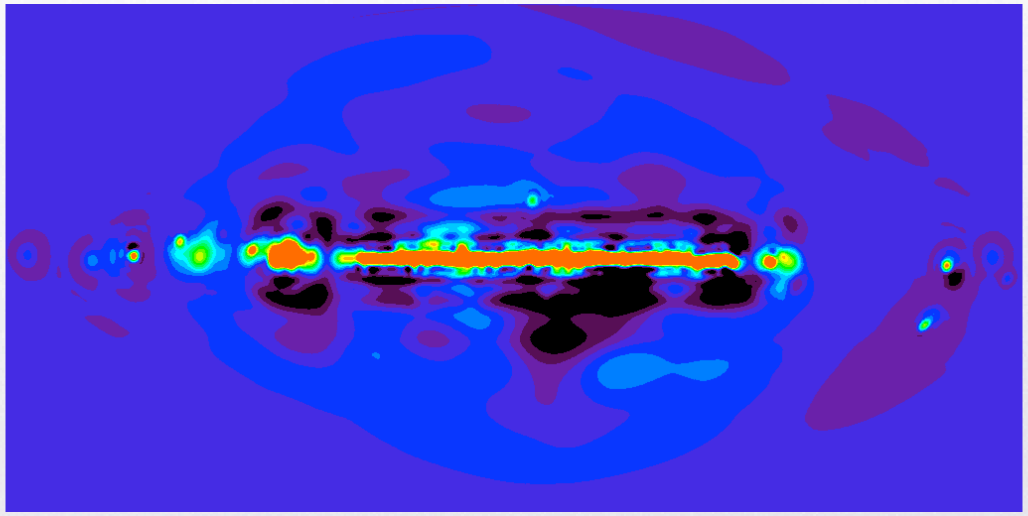
Ka



Q



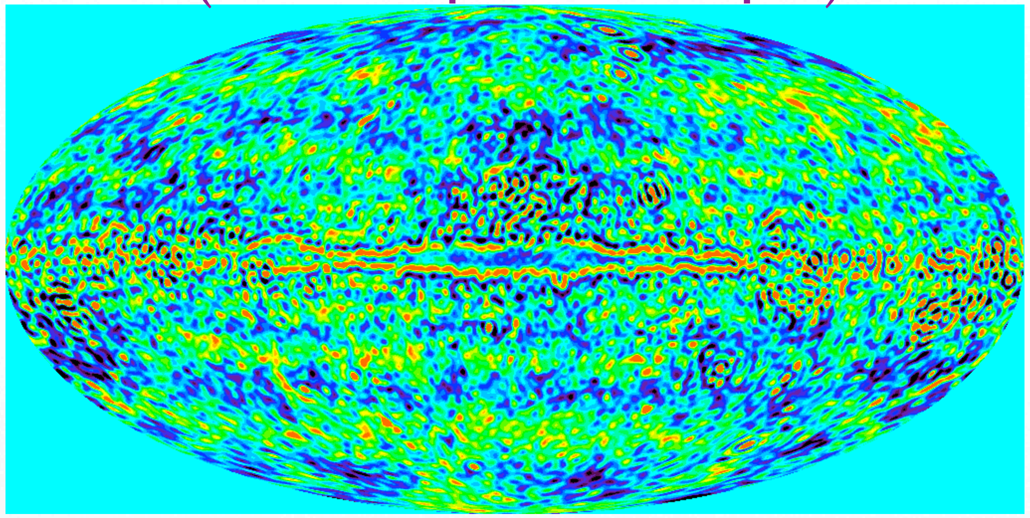
V



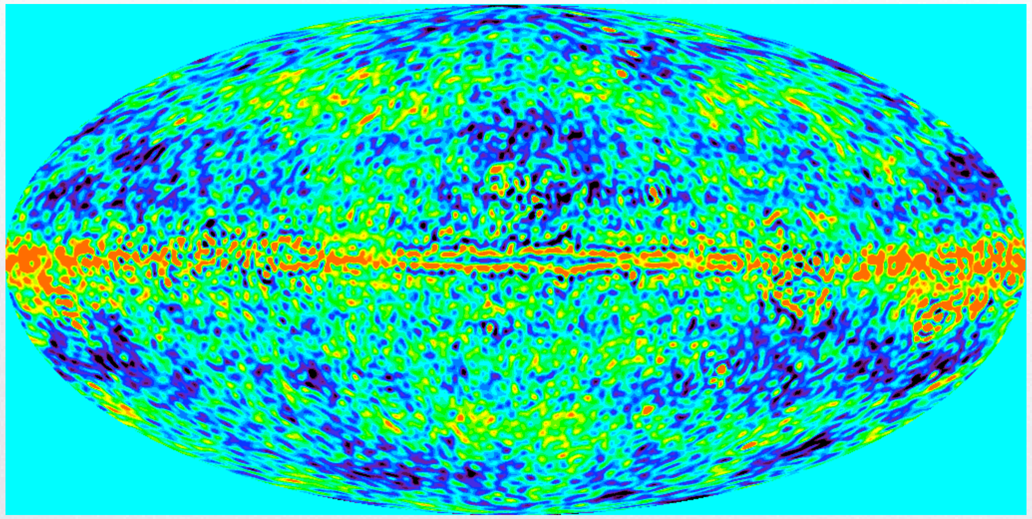
W

CMB Estimates

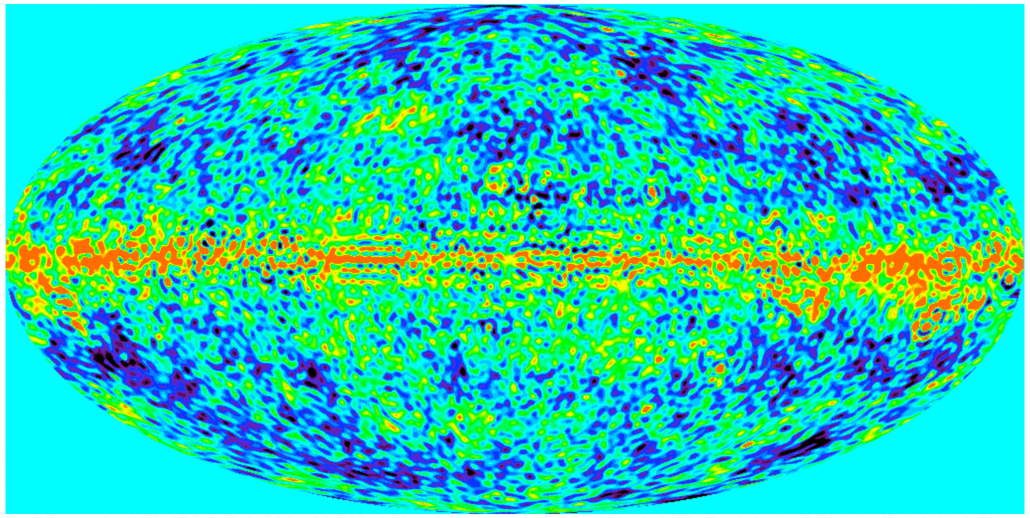
(T: −200 μK → 200 μK)



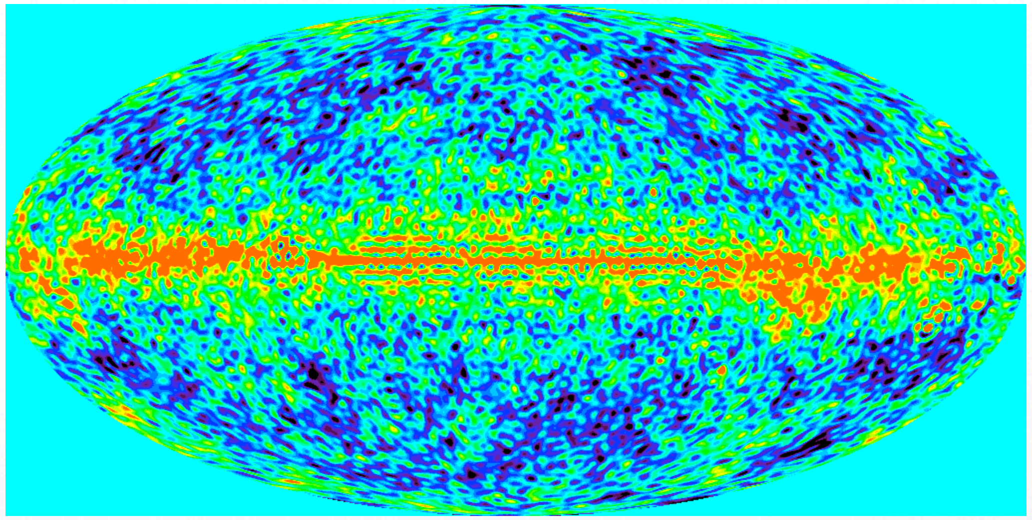
K



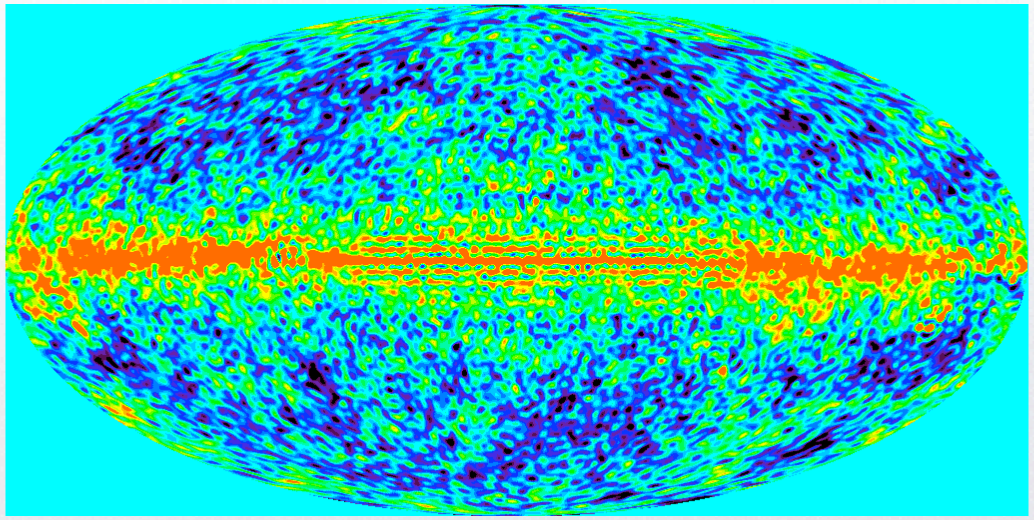
Ka



Q

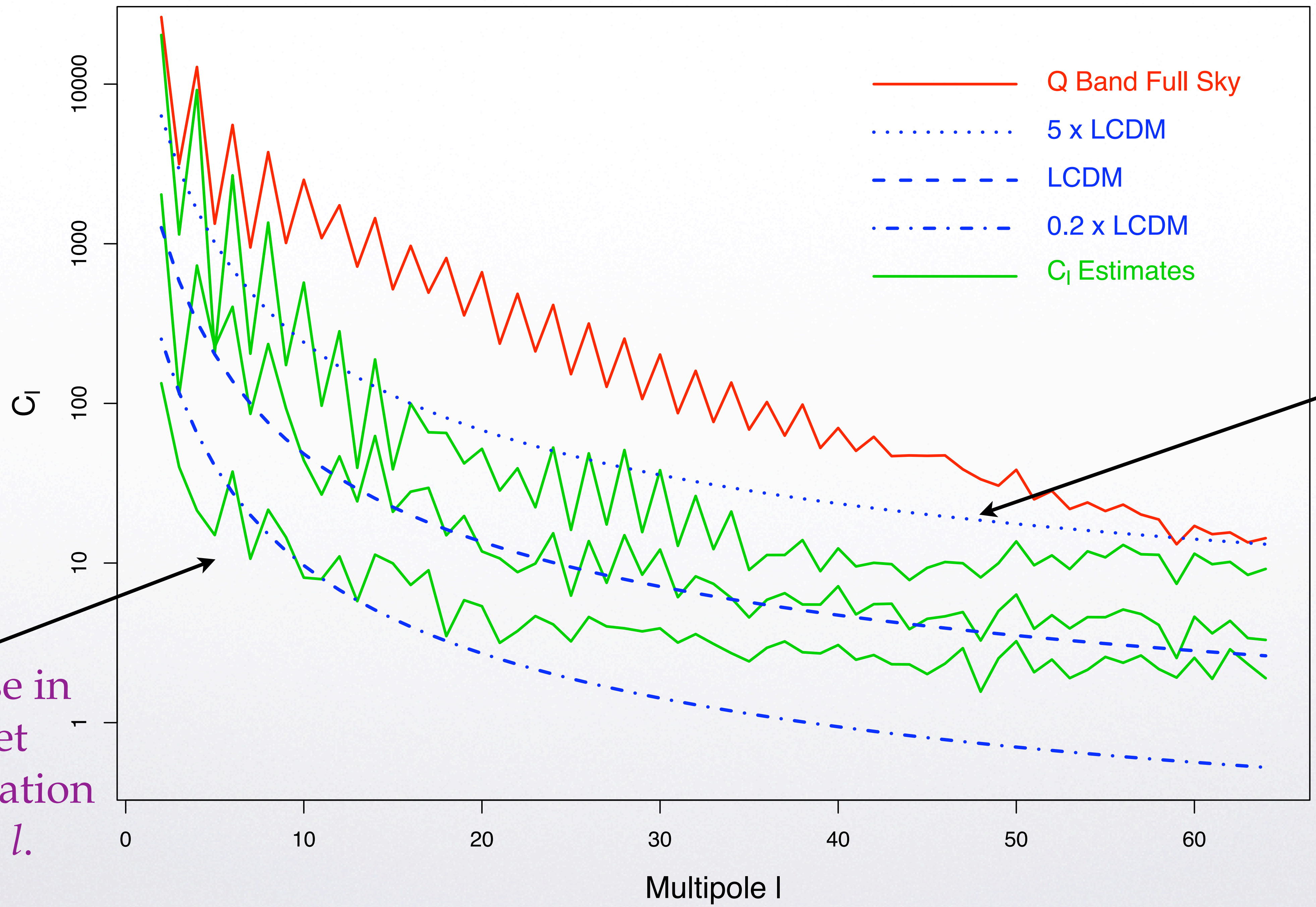


V



W

Sparsity Drives Identifiability



FG sparse in needlet representation at high *l*.

FG dense in needlet representation at low *l*.

Future Work

- Our prototype nonparametric shrinkage procedure shows great promise in being able to separate components in an identifiable manner.
- We will develop a more powerful statistical procedure in which we will build in constraints that, e.g., deal with strong radio sources, while remaining computationally efficient.