External Correlations of the CMB and Cosmology

Cross Correlation Analysis Using Gibbs Sampling

Benjamin D. Wandelt
University of Illinois at Urbana-Champaign

The Gibbs Sampling Team So Far (alphabetical by institution)

- JPL/Caltech
  - Jeff Jewell
  - Ian O'Dwyer
  - Krzysztof Górski
- University of California at Davis
  - Lloyd Knox
  - Mike Chu
- University of Illinois at Urbana-Champaign
  - BDW
  - Greg Huey
  - David Larson
- University of Oslo
  - Hans-Kristian Eriksen

- Others have helped with aspects of this work not directly related to GS:
  - Tarun Souradeep and students
  - Frode Hansen

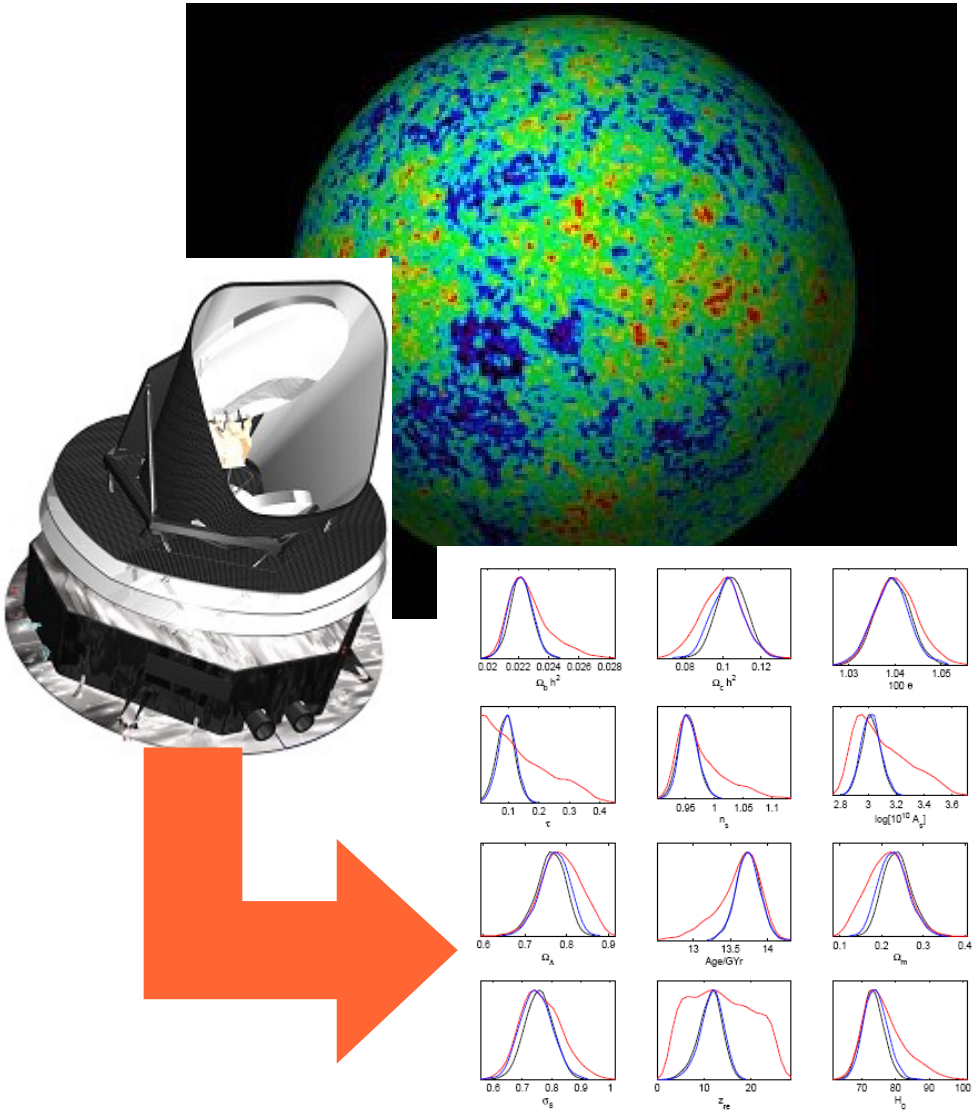
Statistical Challenges of Measuring External Correlations

- What to do with different masks/observed sky regions?
- Optimal use of non-uniformly observed sky area
- Optimal combination of different point spread functions in different data sets
- Correlated noise
- Systematic effects
- Estimating parameter posteriors from measured correlations
- Significance of a detection?

You do not want to have to worry about these technical issues.

Is there a way to overcome these challenges ?

Bayesian Cosmological Data Analysis



Cosmological data analysis takes

astronomical observations (D)

and turns them into

statistical statements about the parameters that define our Universe (θ)

- Conceptually straightforward:

$$P(\theta|D) \propto P(D|\theta)P(\theta)$$

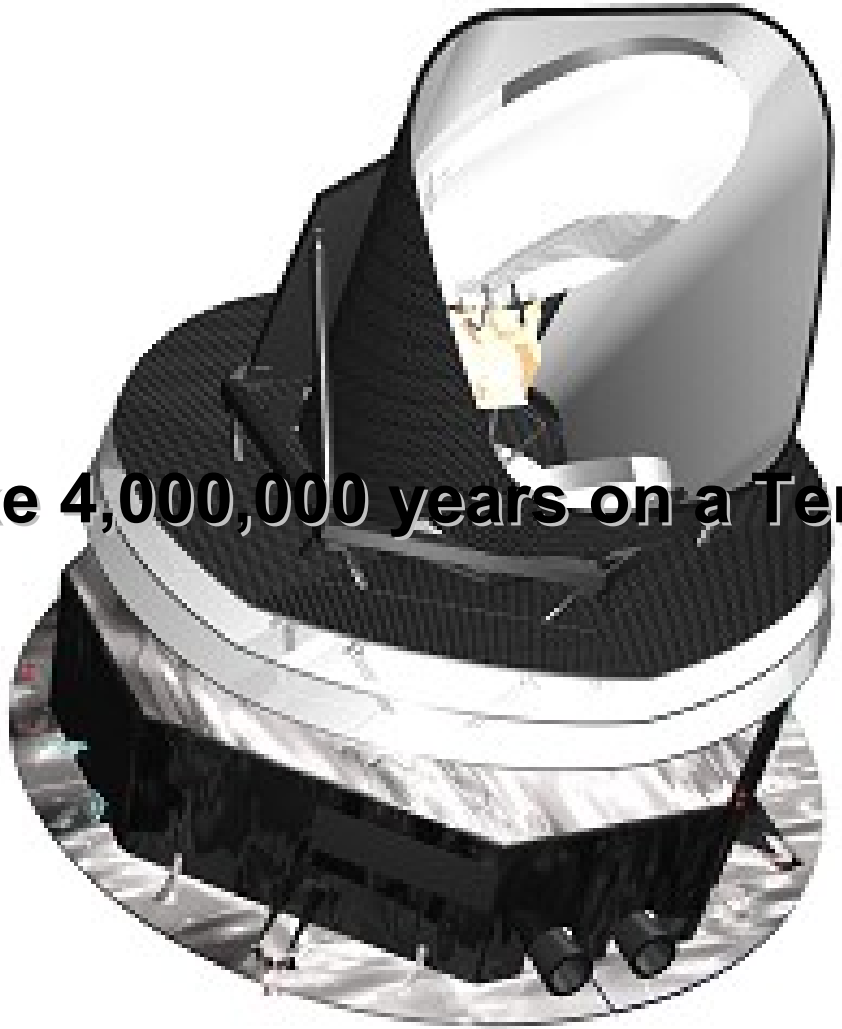
- After COBE –for more than a decade– the field has had to cope with approximations that avoid the computational difficulty of evaluating the terms in this equation.

THE COSMOSTATISTICS PROBLEM

- Black bar: size of data set
- Red area: work required to evaluate

$$P(\theta|D) \propto P(D|\theta)P(\theta)$$

THE COSMOSTATISTICS PROBLEM

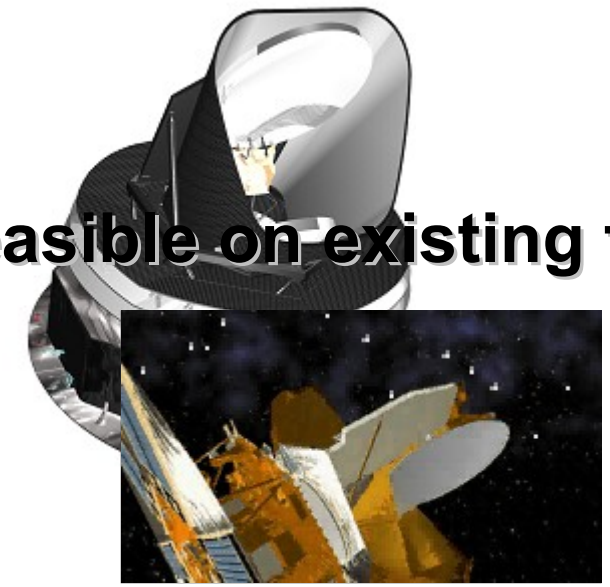


Planck would take 4,000,000 years on a TeraFlop facility

Computational Speed of Gibbs Sampling

- Black bar: size of data set
- Red area: work required to Gibbs sample

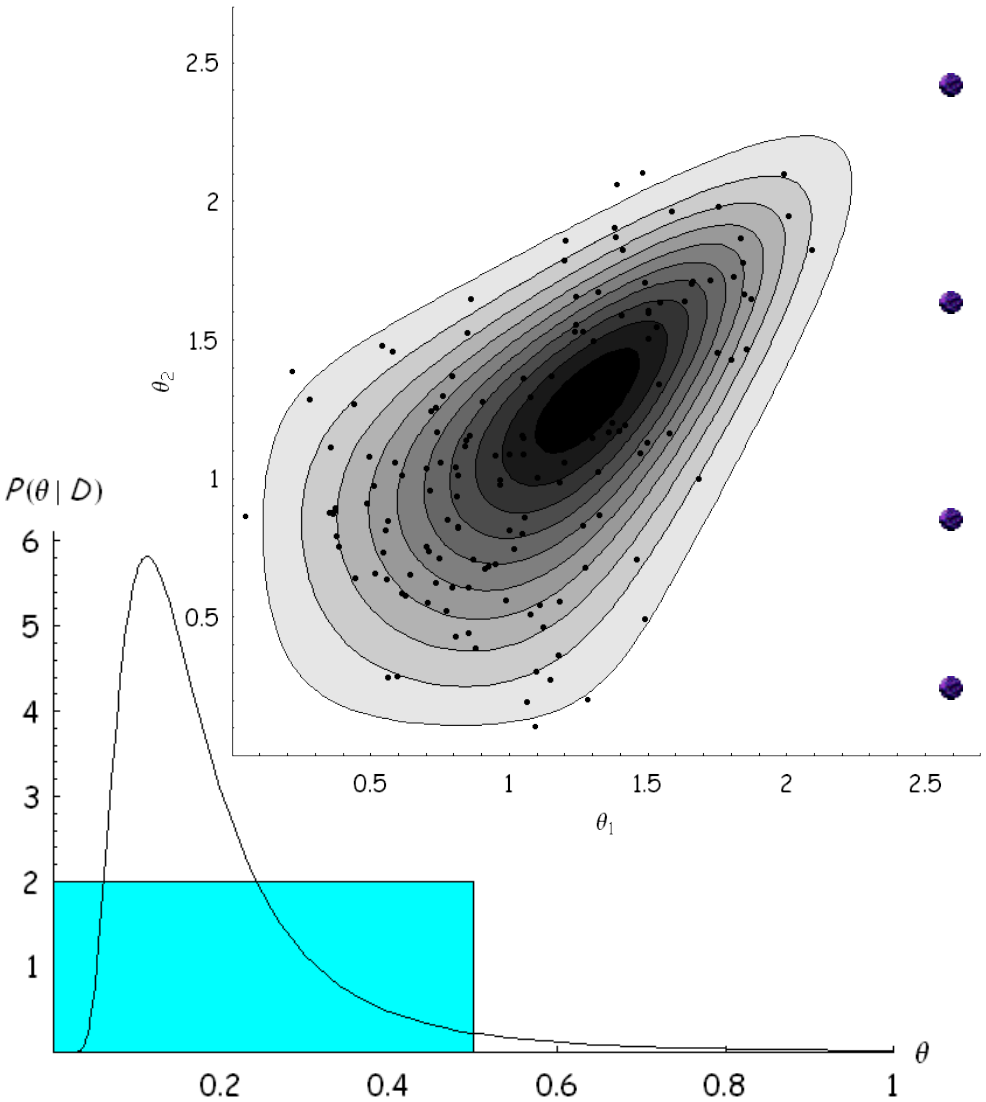
Computational Speed of Gibbs Sampling



Feasible on existing facilities

- The computational effort for each Gibbs sample is $O(N^{1.5})$ less than for the brute force techniques.
- For WMAP and Planck
  $N \sim 10^7 \rightarrow N^{1.5} \sim 10^{10.5}$
- This acceleration is of the same order as the approximate (Pseudo-C_ℓ) techniques.
- Possible optimizations and simplifications are very similar to the approximate techniques (e.g. destriping for 1/f noise;...)
- Challenges are also similar (memory, disk I/O,...)

Gibbs Sampling: How?



- Gibbs sampling is a Monte Carlo technique for generating samples from the likelihood/posterior.
- It recovers the results of the full Bayesian approach without brute force evaluation of the likelihood.
- The science products, θ, consist of between ~10 and ~100 numbers.
- Therefore ~100 – ~10,000 independent Gibbs samples characterize the detailed information, uncertainty, and interdependence in all the CMB science products.

(Jewell, Levin Anderson 2004, Wandelt, Larson, Lakshm. 2004; Eriksen et al. 2004)

Why Bayesian analysis?

Implemented

- Full statistical distributions for science products (auto-spectra, cross-spectra, cosmological parameters).
- Full correction and propagation of error for every modeled systematic effect.
- Optimal combination of different frequency bands/channels (achieving the best resolution, best combined mask, and smallest noise possible).
- Foregrounds handled with tunable combination of prior knowledge and data dependence
- Signal reconstruction and polarization E/B separation "for free"
  - use Gibbs sampling to clean out instrumental effects and apply analyses using the reconstructed signal.
- Can correct for arbitrary beam shapes/point spread function
- Can infer noise properties and take into account noise correlations.
  - Destriping or full treatment
- Exact parameter posterior without intermediate approximation
- Gibbs sampling is the only independent, production-quality technique that can offer an independent check on approximate results
  - All other techniques (MASTER, FASTER, Spice, and polarized versions) are based on the same underlying statistical approach (unbiased, lossy, quadratic estimator)
- <u>Any improvement to the instrument model immediately improves all parts of the cosmological infererence.</u>

Examples

Perfect Data

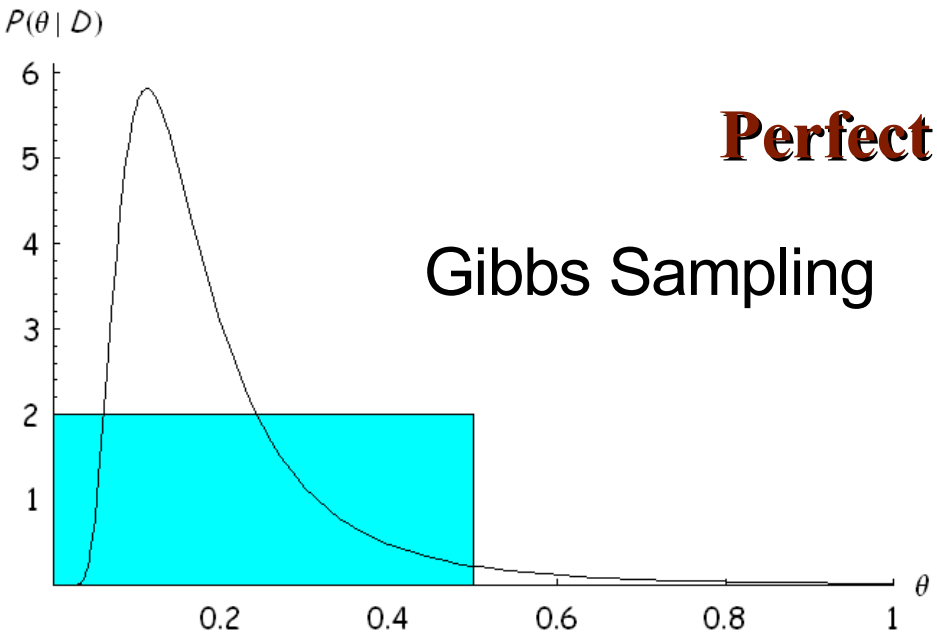
-

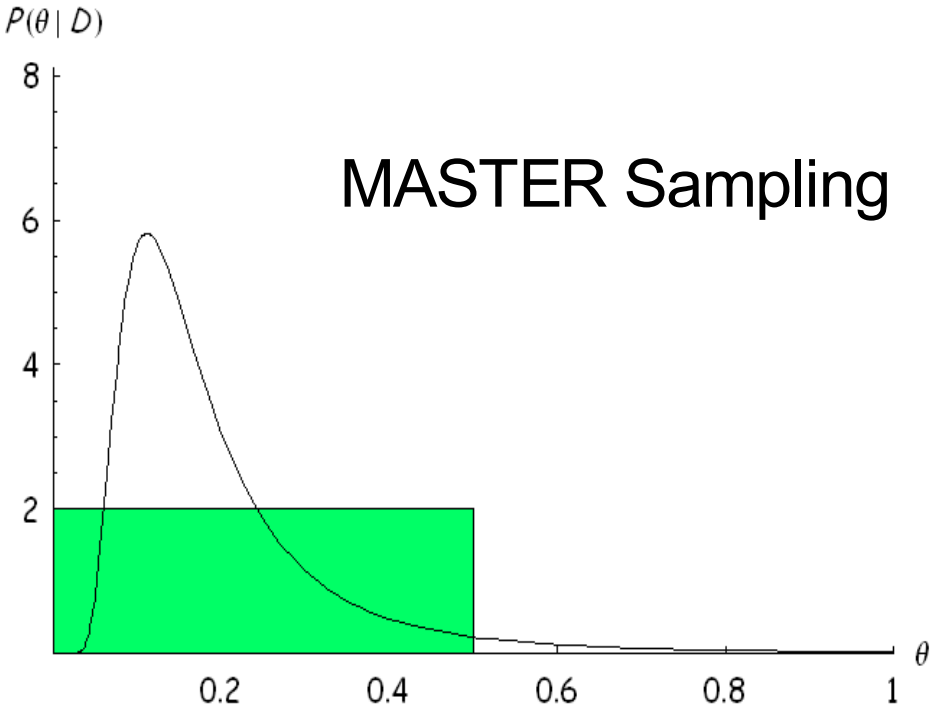
WMAP First Year Data

-

WMAP3

Perfect Data





- What do the error bars from a MASTER-like approx. technique mean?
- Here is an example using perfect, all sky data without any noise.
- We focus on $\theta = C_4$.
- In both cases, the solid line shows the actual likelihood for C_ℓ at ℓ=4.
- The samples are obtained using Gibbs sampling and by "regular" Monte Carlo, like MASTER.
- We see the effect of just using a single point estimate to simulate the uncertainty in the data.
- This is *neither frequentist nor Bayesian.*
- Finding the actual frequentist confidence ball is *extremely* costly.
- This will be even more important for real data.

WMAP1: Bayesian Analysis reduces "low power on large scales" from 99.5% to 90% effect

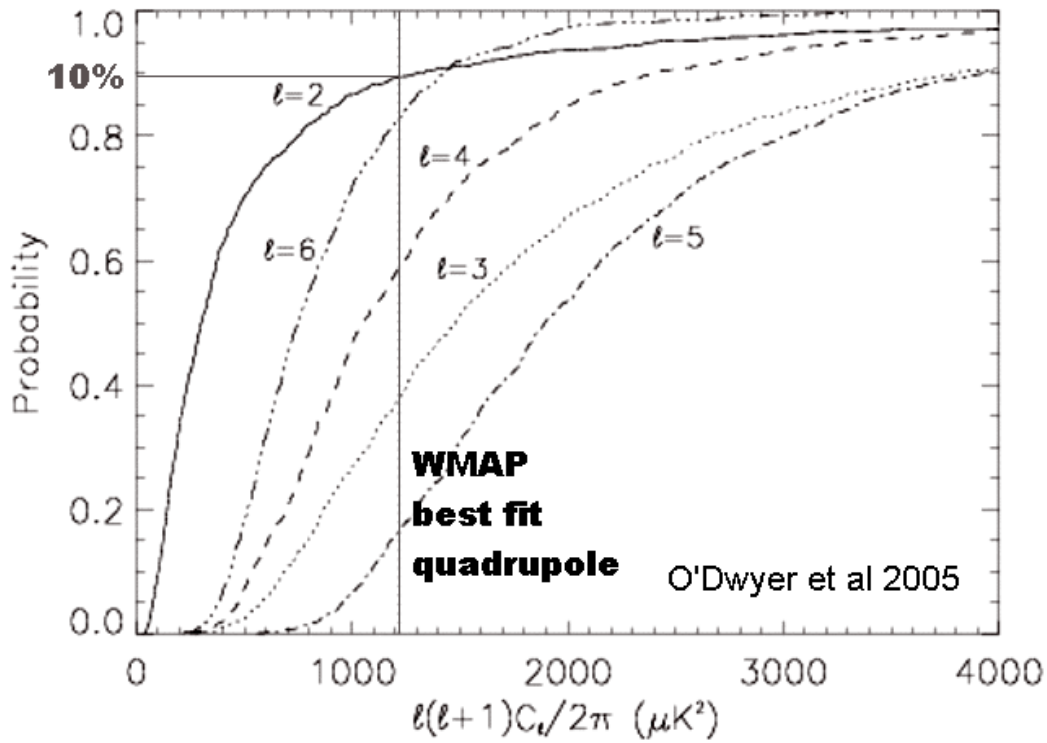
Our analysis demonstrated that the power spectrum likelihoods at low ℓ have strong tails to high C_ℓ.

This leads to a probability in excess of 10% that the true C_2 is even larger than the WMAP best fit C_2.

C_3 is unremarkable.

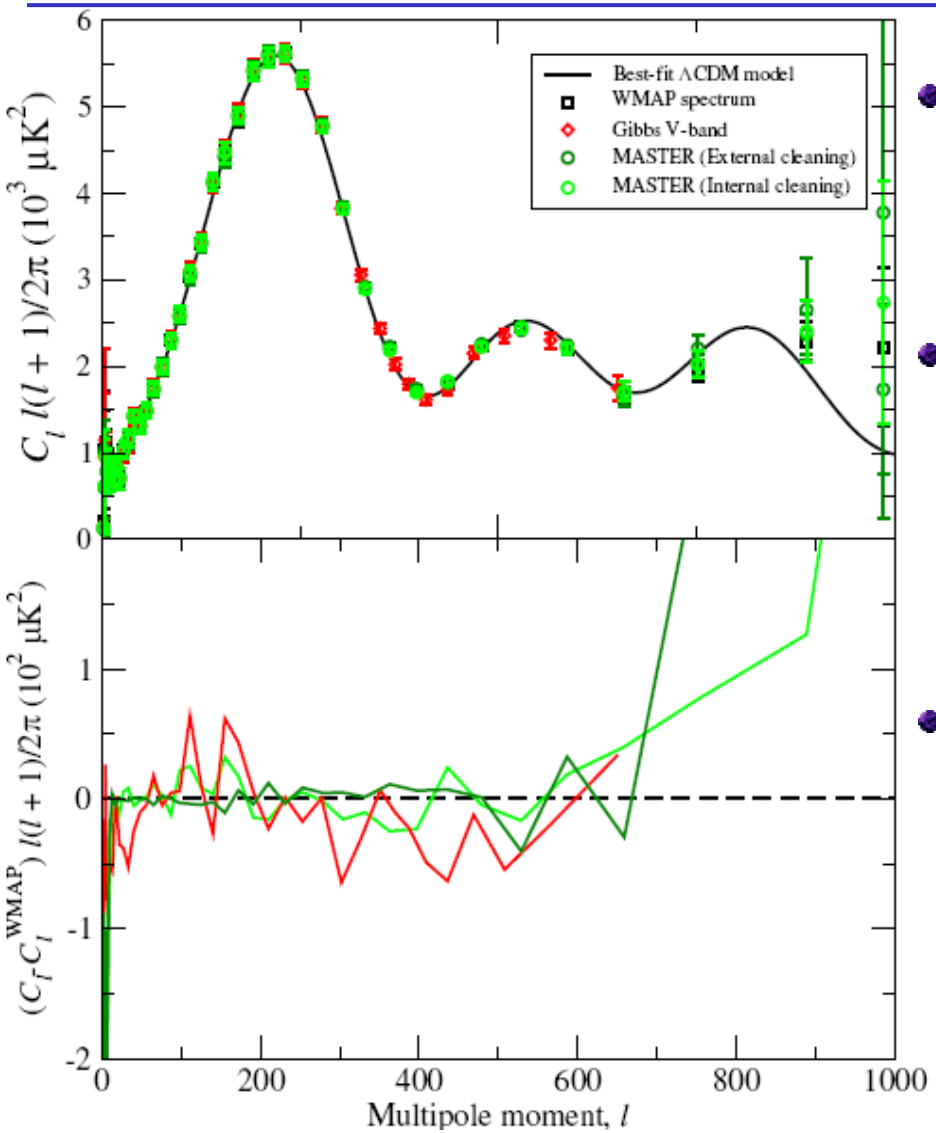
(<u>Note: this is due to statistics, not Foreground marginalization</u>)

$P(C_2 > x \mid \text{data})$



(O'Dwyer et al. 2005)

WMAP3: Making short work of 12,000,000 Pixels



- **Performance**
  - **We are routinely able to generate ~5,000 samples *per day per channel,* using 10 parallel samplers and 3 processors per channel.**
- **Independent techniques**
  - **Very useful check against blind separation method (Souradeep) and MASTER X-spectra (Hansen)**
  - **We think we are a useful check for the WMAP team!**
- **Watch out for our forthcoming results!**

Conclusions

- Everyone would perform a full likelihood analysis of CMB data if it was easy.
  - All early analyses were performed using Bayesian techniques. Pseudo-C_ℓ techniques were invented as work-arounds.
  - Pseudo-C_ℓ techniques are very convenient, but the error bars are "special." This is dangerous when S/N is ~1, which is always what we are most interested in.
- Gibbs sampling is a convenient Monte Carlo approach that makes full analysis possible again – even for Planck, including cross-correlations with other data.
- We are routinely analyzing WMAP-size data of temperature anisotropy on modest computational facilities.

- Extension to polarization is nearly complete (D. Larson)
- Extension to asymmetric beams and correlated noise is being tested (J. Jewell)
- Full coupling from maps to parameters will be coming soon.

- Most importantly, Bayesian analysis (via Gibbs sampling) enables *new science*.
  - Reconstructed signal maps are useful for measuring higher order and external correlations
  - What is the $P(\tau|D)$ after marginalizing over polarized foregrounds?
  - At what significance have the ISW effect/other X-correlations been detected?