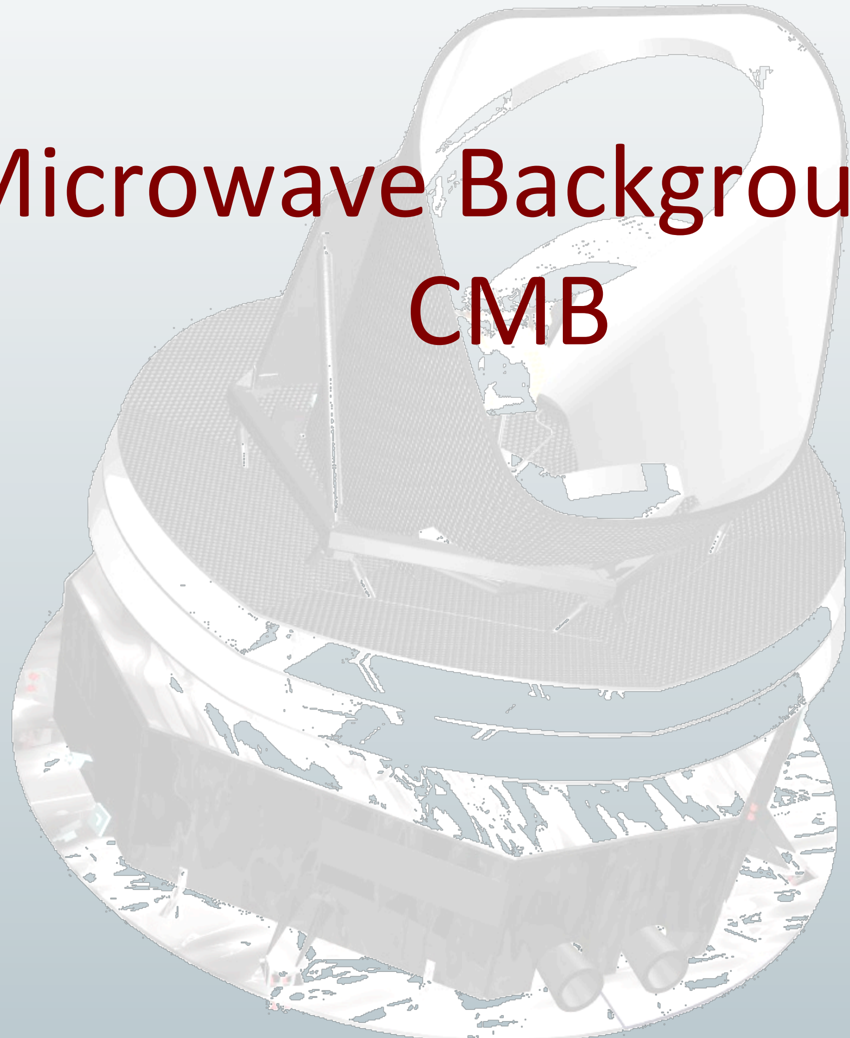
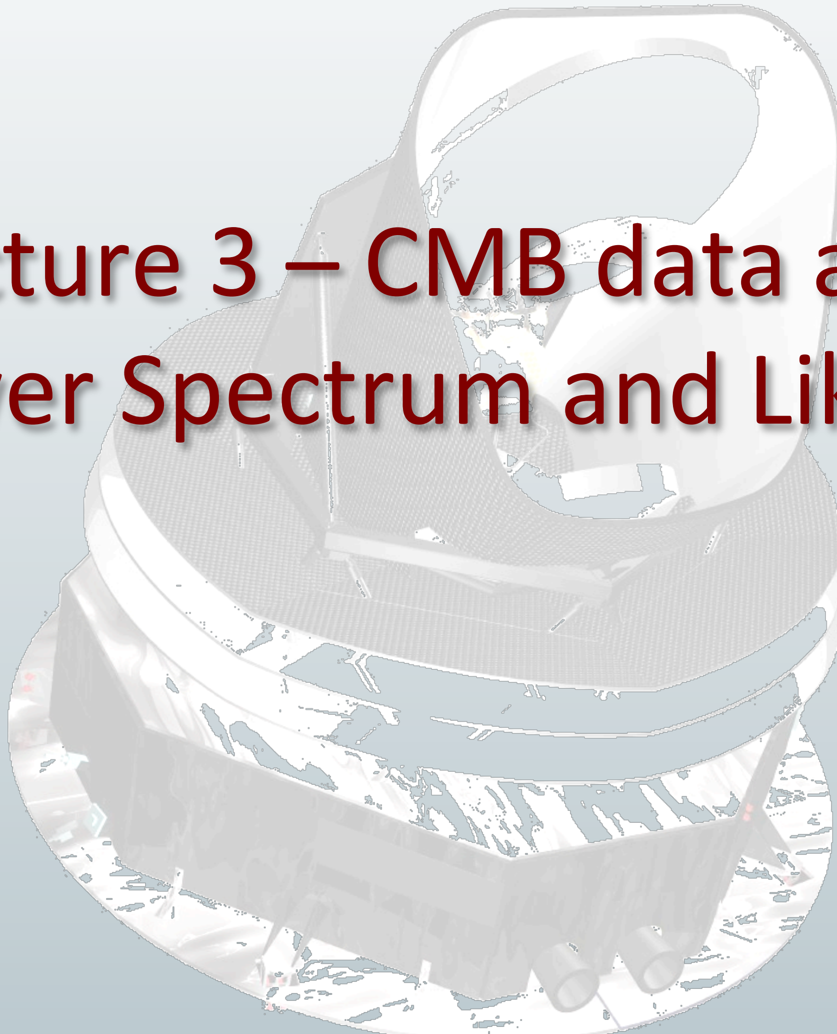
Cosmic Microwave Background Radiation
CMB



Graça Rocha
JPL/Caltech
graca@caltech.edu

Vth INPE Advance Course on Astrophysics, INPE, September 2013



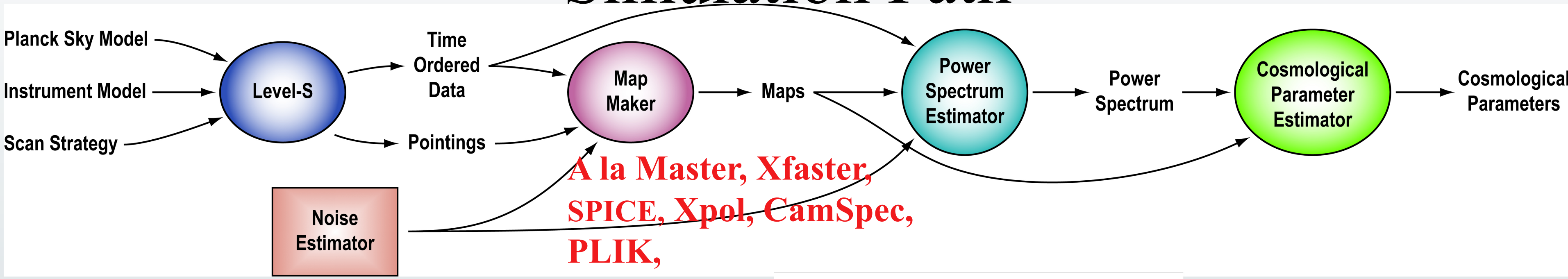
Lecture 3 – CMB data analysis Power Spectrum and Likelihood

Graça Rocha
JPL/Caltech
graca@caltech.edu

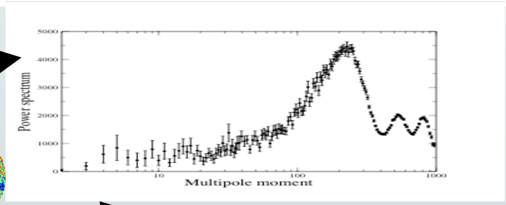
Vth INPE Advance Course on Astrophysics, INPE, September 2013

CMB Analysis

Simulation Path

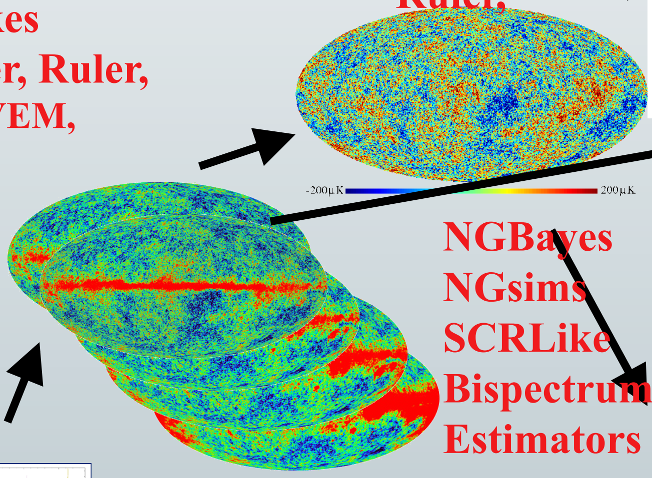


A la Master, Xfaster, SPICE, Xpol, CamSpec, PLIK, Commander Ruler,

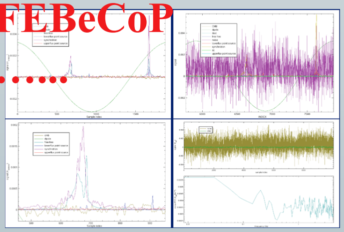


Xfaster, PiXFaster ..Gasussian correlated …

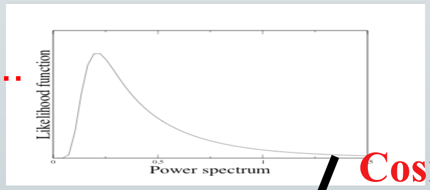
PowellSnakes Commander, Ruler, SMICA, SEVEM, NILC …..



MADAM, Springide, Polkapix, … MCBR FEBeCoP ……



CamSpec, PLIK, Xfaster, Xpol, etc..



NGBayes NGsims SCRLike Bispectrum , trispectrum Estimators …..

Xfaster …..

CosmoMC MCBR ……

Non-Gaussianity Physics

Ω_b n_s τ h Ω_0 σ_8

Cosmological parameters

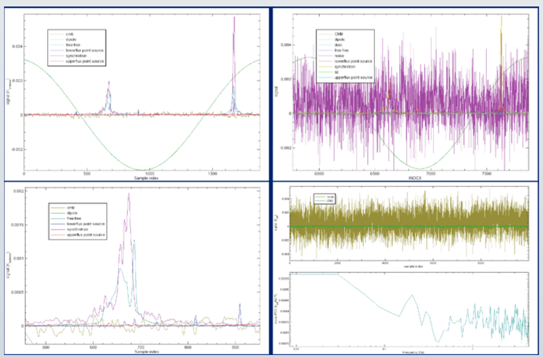
Bayesian approach

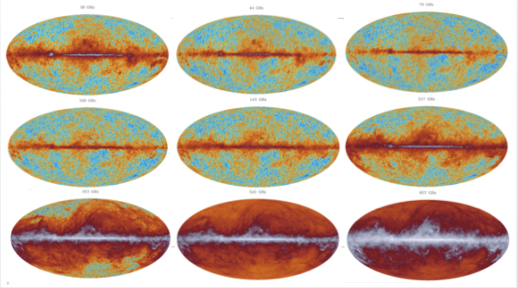
Data Analysis Path

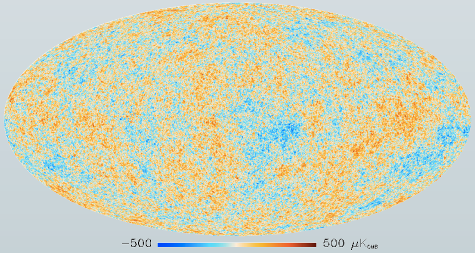
Angular Power Spectrum - Motivation

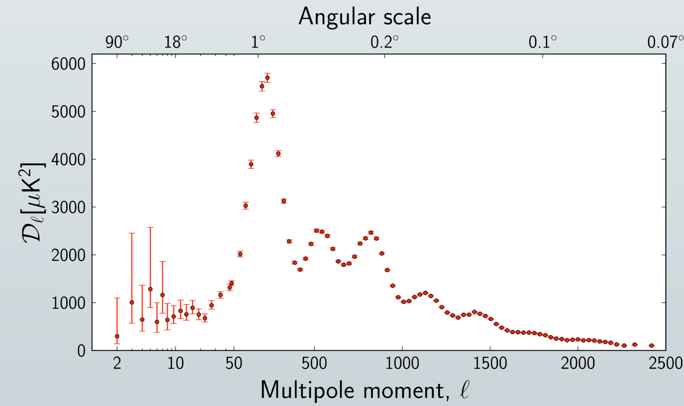
With the advent of large, high-quality data sets direct extraction of science from the pixelized maps is computationally expensive and in fact unfeasible. Accurate estimation of the angular power spectrum enables the extraction of science with minimal loss of information

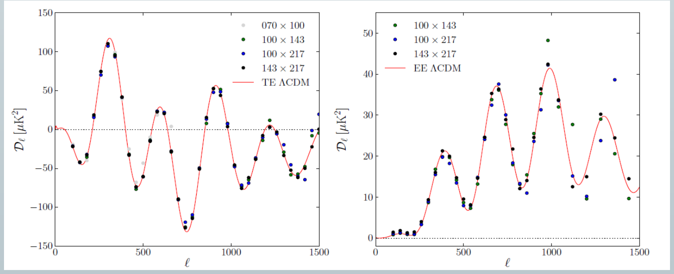
$N \approx 10^{11}$ $\quad$ $N \approx 10^{7}$ $\quad$ $N \approx 10^{3}$











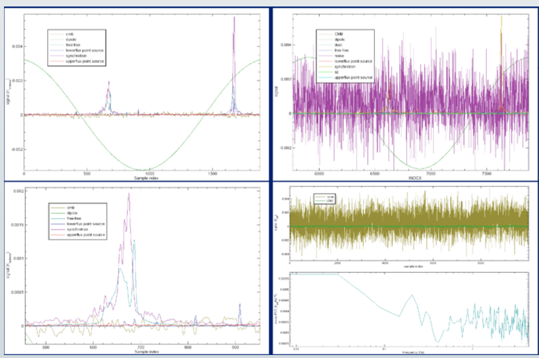
To estimate cosmological parameters one needs to compare Theory with Observations – via the Likelihood function

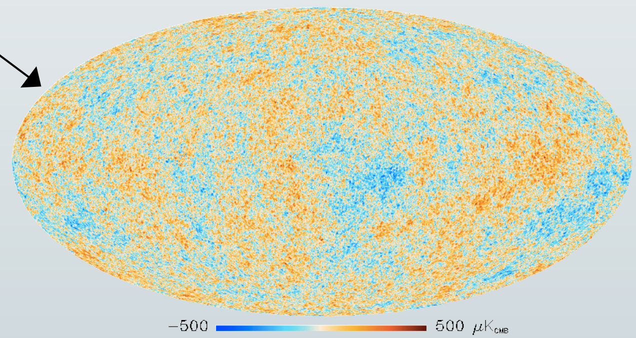
Bayesian statistics

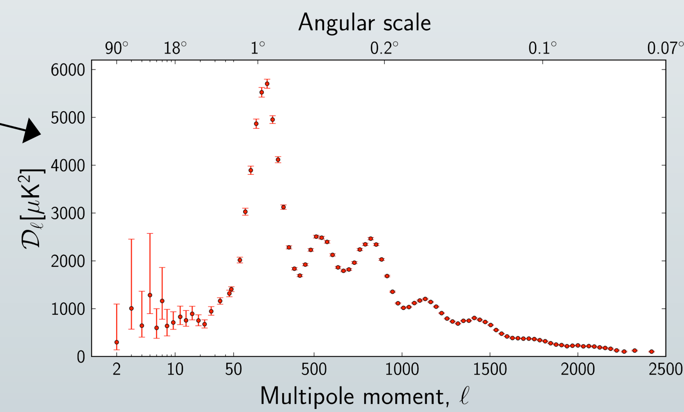
Angular Power Spectrum - Motivation

With the advent of large, high-quality data sets direct extraction of science from the pixelized maps is computationally expensive and in fact unfeasible. Accurate estimation of the angular power spectrum enables the extraction of science with minimal loss of information

$N \approx 10^{11}$ $\quad$ $N \approx 10^{7}$ $\quad$ $N \approx 10^{3}$

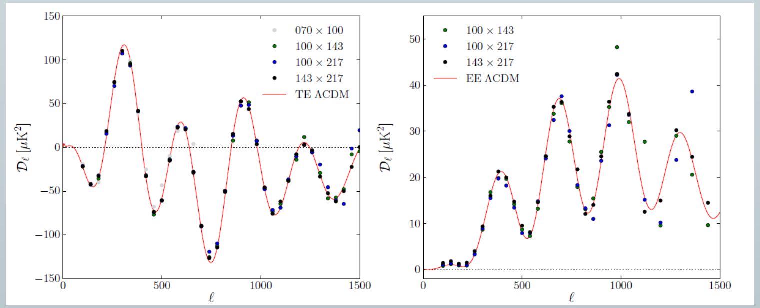






To estimate cosmological parameters one needs to compare Theory with Observations – via the Likelihood function

Bayesian statistics



The angular power spectrum of the CMB anisotropy

- Current measurements are quite consistent with the primary fluctuations in the CMB being a Gaussian random field.
- The primary fluctuations in the CMB are thus well described by their angular power spectrum (see lecture 5)

$$\frac{\Delta T}{T} \equiv \sum_{\ell m} a_{\ell m} Y_{\ell m}(\theta, \phi)$$

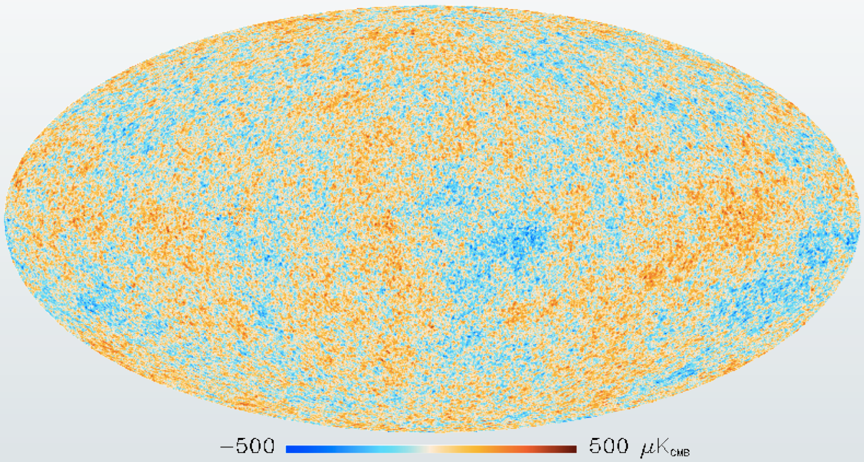
Curved sky analogue of FT

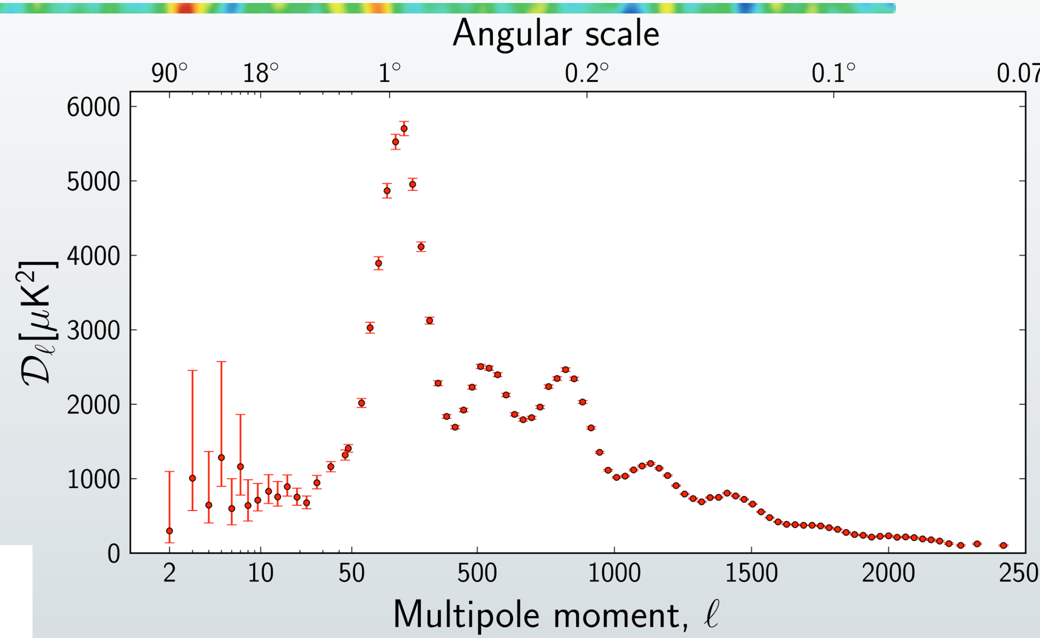
$$\langle a_{\ell m} a_{\ell' m'} \rangle = C_\ell \, \underbrace{\delta_{\ell \ell'} \delta_{m m'}}$$

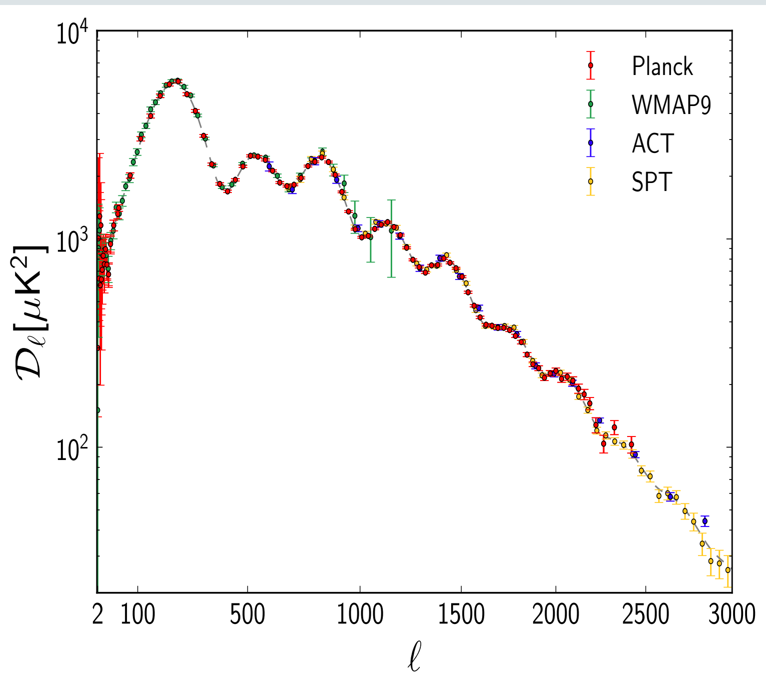
Statistical isotropy

Rotational invariance

T - current state-of-the-art of CMB temperature measurements



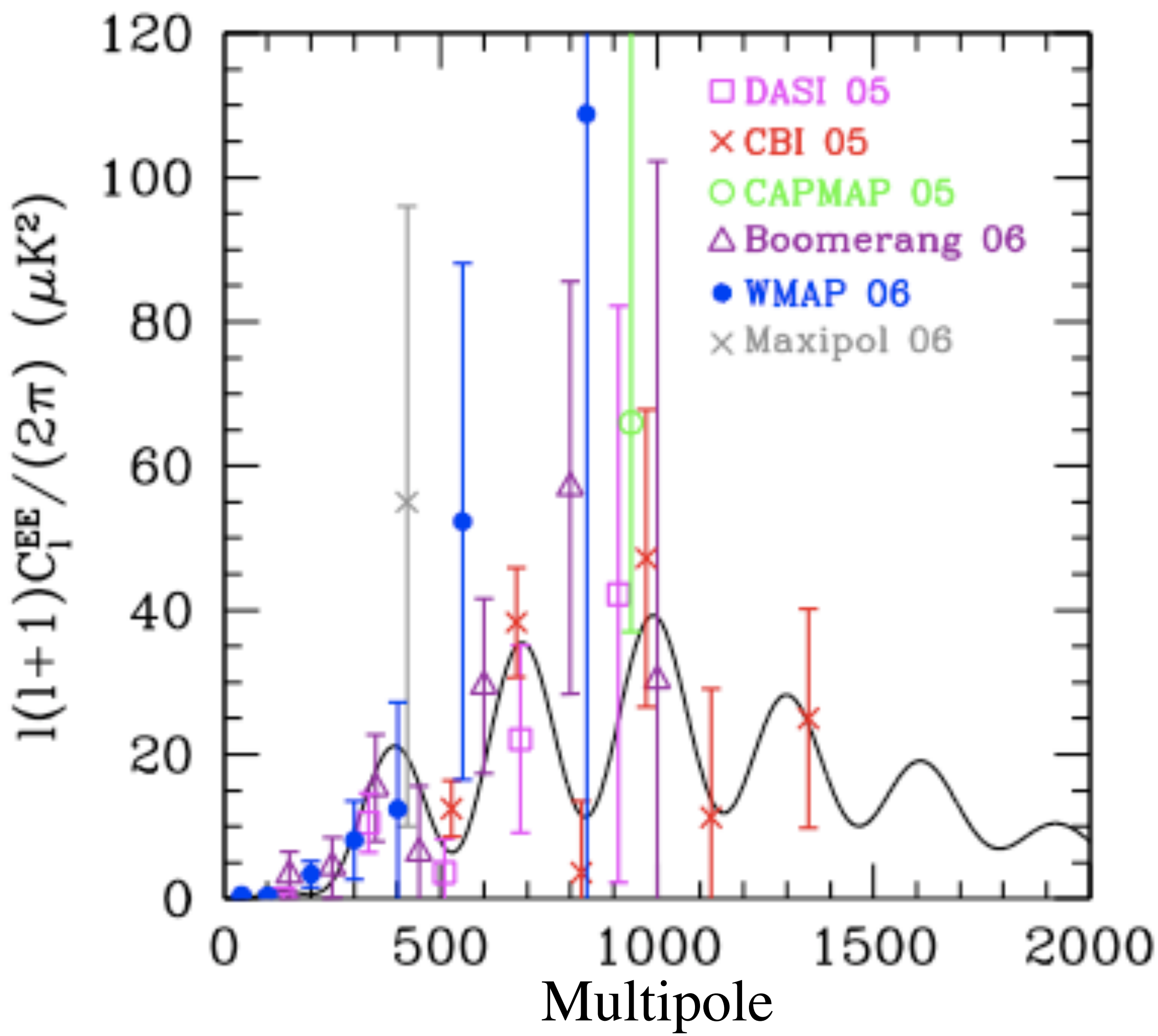




- *Planck 2013 XV, Likelihood*
- *Planck 2013 XVI Cosmological parameters*

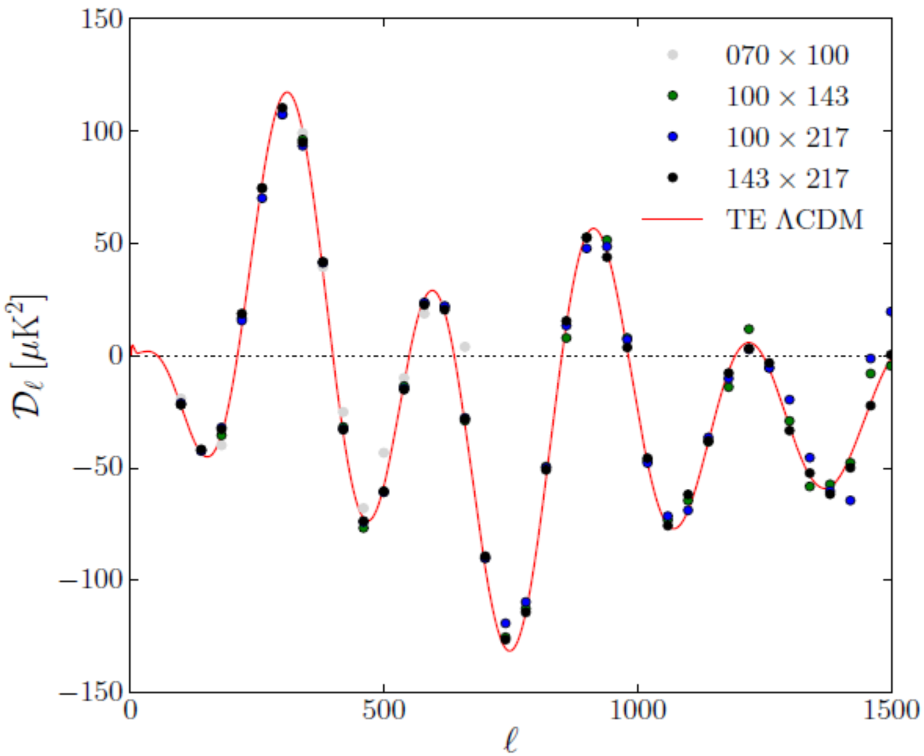
P - the world compilation of CMB polarization measurements

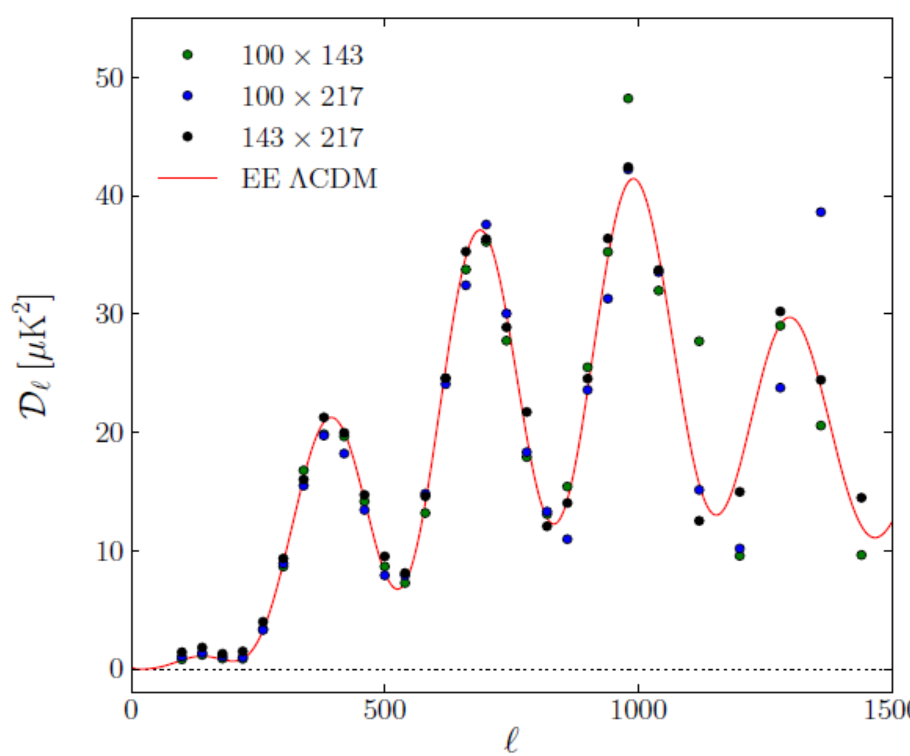
Courtesy Lewis Hyatt



Polarization with Planck (preliminary)

TE and EE Power Spectra (preliminary!) - red line is not a fit to the polarized spectra – it is the TT best fit model





Excellent quality of the data
Foregounds and systematics are not dominant

Types of Power Spectrum Estimators

MLE **Maximum Likelihood Estimator**	**MADspec** **Cambridge ML**	**Computes the Power Spectrum C_l that maximizes the Likelihood:** $$L(d \mid p) = \frac{1}{2\pi^{N/2} \mid C \mid^{1/2}} \exp\left(-\frac{1}{2} dC^{-1} d^t\right)$$
	Gibbs sampling **Commander** **MAGIC**	**MCMC posterior estimation code that samples from the full CMB posterior by a Gibbs sampling scheme.** **It iteratively samples from the conditional densities:** **P(sky signal I power spectrum, data)** **P(power spectrum I sky signal, data)**
Pseudo - C_l **Monte Carlo**	**Spice** **ROMAster** **Xpol** **CrossSpec** **CamSpec** **PLIK**	**Estimates C_l : or $C(\theta)$** $$C_l = \frac{1}{2l+1}\sum_{m=-l}^{l} \mid a_{lm} \mid^2 \; ; \quad C(\theta) = \langle T(q_1)T(q_2) \rangle = \sum_l \frac{l+1/2}{l(l+1)} C_l P_l(\cos\theta)$$ **Uses either fast spherical transforms or fast evaluation of the 2-point Correlation function (Spice)**
Hybrid	**Xfaster**	**Computes the C_l that maximizes the Likelihood:** $$C^{obs}_{lm,l'm'} = a^{obs}_{lm} a^{obs*}_{l'm'}$$ $$C_l = \frac{1}{2}\sum_{l'} F^{-1}_{ll'} Tr\left[C^{-1}\frac{\partial S}{\partial C_{l'}} C^{-1}\left(C^{obs} - N\right)\right]$$
	GL- Hybrid	**Combines a Quadratic MLE at low-l and Pseudo - C_l at high-l** **With a smooth transition**

Power Spectrum Estimators - some details…

MADspec	Parallel implementation of the Bond, Jaffe & Knox maximum-likelihood algorithm for the estimation of CMB temperature and polarization angular power spectra. It consists of two stages: ➢Calculating the pixel-pixel noise correlation matrix ➢ Estimating the angular power spectra MADspec noise correlator uses the time-ordered data, pointing, and time-time noise correlations to generate the pixel-pixel noise correlation matrix and map. The ML solution is found iteratively using a Newton-Raphson algorithm It requires computing the inverse of the Signal+Noise Covariance matrix : $C(p) = S(p) + N$
Gibbs sampling	Draw samples from the posterior distribution $P(C_l \mid d)$ - difficult -> though It can be proved that if one can sample from the conditional distributions: $P(s \mid C_l, d)$ $P(C_l \mid s, d)$ Then one can sample form the joint density in an iterative fashion: $P(C_l, s \mid d)$ and marginalise over s: Begin from a starting guess C_l^0 ; and iterate the following equations: $s^{i+1} \leftarrow P(s \mid C_l^i, d)$ $C_l^{i+1} \leftarrow P(C_l \mid s^{i+1})$ After some burn-in the $\left(C_l^i, s^i\right)$ converge to being samples of the joint distribution -> Gibbs sampler technique Only requires the inverse of the Noise covariance matrix for the 1st step and the inverse of the signal Covariance for the 2nd step of this iterative procedure.

Power Spectrum Estimators - more details…

Pseudo - C_l	**Estimate:** $$C_l = \frac{1}{2l+1}\sum_{m=-l}^{l} \lvert a_{lm} \rvert^2$$ **But,** ➢ Signal on the sky is a superposition of galactic and extra-galactic fg + experimental noise --> bias above estimator wrt underlying cosmological signal ➢ Scanning strategy - usually non-trivial ➢ Filtering of the timestream removes power on both the large and small scales --> bias the Power Spectrum wrt the underlying signal Change above estimator to include these effects: $$\langle \tilde{C}_l \rangle = \sum_{l'} K_{ll'} F_{l'} B_{l'}^2 \langle C_{l'} \rangle + \langle \tilde{N}_l \rangle$$ Master like methods - rely on estimating the biasing terms by MC simulations of identical observations on a known signal: ➢ Auto-spectra codes - require MC's of the noise to compute the Noise bias; Cross-spectra does not ➢ require MC's of the signal alone to compute the Transfer function (info on the pre-processing of Tods) ➢ Covariance matrix of the estimated PS, uncertainty on the estimator - requires MC's of signal+noise ❖ CamSpec, Xpol produces analytical estimates of the correlation matrix - it does not require MC's
Hybrid (Xfaster)	the (l,m) MLS for the PS is: $$C_l = \frac{1}{2}\sum_{l'} F_{ll'}^{-1} Tr\left[C^{-1}\frac{\partial S}{\partial C_{l'}} C^{-1}\left(C^{obs} - N\right)\right]$$ where $C^{obs}_{lm,l'm'} = a^{obs}_{lm} a^{obs*}_{l'm'}$ ➢ Recast in isotropic, diagonal approximation, ➢ Parameterize cut-sky PS wrt full sky through a set of deviations q_l from a template shape $C_l^{(S)}$. ➢ Requires MC's of signal alone (for transfer function) and noise alone (for noise bias) ➢ The uncertainty in the estimator given by Fisher matrix - a byproduct of the method - no need for MC's of (s+n)! ➢ Multiple Maps analysis, uses full Covariance of the a_{lm}'s (uses both Auto and Cross-spectra)

How do they compare ?

<table>
<tr><td>MADspec</td><td>Exact</td><td>Computationally expensive –> can estimate PS on low-resolution full sky maps

(can compute high-resolution small patches of the sky)</td><td rowspan="2">suitable for PSE @ low-l</td></tr>
<tr><td>Gibbs sampling</td><td>Exact</td><td>Issues with MCMC convergence
for high-l due to low S/N</td></tr>
<tr><td>Pseudo - C_l</td><td>Approx</td><td>Have to assume an approximation to the Likelihood –> Possibly ok for high-l, There is no good approximation at low-l
Auto-Spectra- requires MC's of noise; signal, signal+noise –> computationally expensive</td><td rowspan="2">Suitable for PSE @ high-l</td></tr>
<tr><td>Xfaster</td><td>Approx</td><td>Same as above but:
Auto-Spectra - requires MC's for noise and signal separately - no need for MC's for signal+noise</td></tr>
</table>

How do they compare ?

MADspec	Exact	Computationally expensive –> can estimate PS on low-resolution full sky maps (can compute high-resolution small patches of the sky)	suitable for PSE @ low-l
Gibbs sampling	Exact	Issues with MCMC convergence for high-l due to low S/N Blackwell-Rao Likelihood	
Pseudo - C_l Hivon, Gorski et al.	Approx	Have to assume an approximation to the Likelihood –> high-l Gaussian Correlated Likelihood Auto-Spectra- use the half difference ring maps to estimate noise bias, need signal+noise though (use pseudo-analytic approach,or Cross-Spectra – noise bias less of an issue, need to get signal+noise	Suitable for PSE @ high-l
Xfaster Rocha, Contaldi et al.	Approx	Same as above and Xfaster Likelihood Auto-Spectra – use the half difference ring maps to estimate noise bias, signal come naturally from Fisher approx. Cross-Spectra as above	

A la MASTER (Hivon, Gorski et al, 2001)

Pseudo - C_l $\longrightarrow$ $C_l = \frac{1}{2l+1}\sum_{m=-l}^{l} |a_{lm}|^2$

Run one of HEALpix module: anafast

Relate full-sky to observed C_l:

$$\langle \tilde{C}_l \rangle = \sum_{l'} K_{ll'} F_{l'} B_{l'}^2 \langle C_{l'} \rangle + \langle \tilde{N}_l \rangle$$

Coupling Matrix due to the cut sky

Filter function to account for pre-processing of data – toi and mapmaking

Beam transfer function

Noise

MASTER

Pseudo - C_l

$$C_l = \frac{1}{2l+1}\sum_{m=-l}^{l} |a_{lm}|^2$$

Run one of HEALpix module: anafast

Relate full-sky to observed C_l :

$$\langle \tilde{C}_l \rangle = \sum_{l'} K_{ll'} F_{l'} B_{l'}^2 \langle C_{l'} \rangle + \langle \tilde{N}_l \rangle$$

Coupling Matrix due to the cut sky

Signal MC simulations

Beam transfer function

Noise Monte Carlo simulations

Another example: XFaster
(Rocha et al, 2011a)

Xfaster is a quadratic maximum likelihood estimator allows to go straight from the map to parameters

$$L(d \mid p) = \frac{1}{2\pi^{N/2} \mid C \mid^{1/2}} \exp\left(-\frac{1}{2} d C^{-1} d^t\right) \qquad C(p) = S(p) + N$$

$$C_l = \frac{1}{2} \sum_{l'} F_{ll'}^{-1} Tr\left[C^{-1} \frac{\partial S}{\partial C_{l'}} C^{-1} \left(C^{obs} - N\right)\right] \qquad \text{MLE}$$

$$F_{ll'} = \frac{1}{2} Tr\left[\frac{\partial S}{\partial C_l} C^{-1} \frac{\partial S}{\partial C_{l'}} C^{-1}\right] \qquad C^{obs}_{lm,l'm'} = a^{obs}_{lm} a^{obs*}_{l'm'}$$

Fisher information matrix

Employ an iterative scheme to reach the maximum of the likelihood:

Start with a guess -> Estimate $F_{ll'}$ -> evaluate C_l

Only feasible for a few thousand – Xfaster solves it by recasting the estimation in the isotropic+diagonal Master like approach -> N is now diagonal and S is summed up in bins to reduce the coupling of modes due to cut-sky

XFaster

$$\tilde{C}_{lm,l'm'} = \delta_{ll'}\delta_{mm'}\left(\tilde{C}_l + \left\langle \tilde{N}_l \right\rangle\right)$$

$$\tilde{C}_{l'} = \sum_{l'} K_{ll'} F_{l'} B_{l'}^2 C_{l'}^S q_{l'}$$

$$\tilde{C}_{l'} = \sum_b q_b \tilde{C}_{bl}^S = \sum_b q_b \sum_{l'} K_{ll'} F_{l'} B_{l'}^2 C_{l'}^S \chi_b(l')$$

Signal MC simulations

$$q_b = \frac{1}{2}\sum_{b'} F_{bb'}^{-1} \sum_l (2l+1) g \frac{\tilde{C}_{b'l}^S}{\left(\tilde{C}_l + \left\langle \tilde{N}_l \right\rangle\right)^2}\left(\tilde{C}_l^{obs} - \left\langle \tilde{N}_l \right\rangle\right)$$

Noise MC simulations

$$F_{ll'} = \frac{1}{2} Tr\left[\frac{\partial S}{\partial C_l} C^{-1} \frac{\partial S}{\partial C_{l'}} C^{-1}\right] = \frac{1}{2}\sum_l (2l+1) g \frac{\tilde{C}_{bl}^S \tilde{C}_{lb'}^S}{\left(\tilde{C}_l + \left\langle \tilde{N}_l \right\rangle\right)^2}$$

- Look for biases in the estimator with Signal+Noise MC simulations
- Assume a given shape Power Spectrum (eg WMAP, flat), Iterate q_b and $F_{bb'}$
- Get the uncertainty in the estimator for free with the inverse of $F_{bb'}$

Xfaster - Polarization

(Rocha et al, 2011a)

Extending the above formalism to polarization, the `XFaster` estimator takes a matricial form, implemented trivially since the matrix C is now block diagonal:
$\tilde{\mathbf{C}} \rightarrow diag(\tilde{\mathbf{D}}_{\ell_{min}}, \tilde{\mathbf{D}}_{\ell_{min}+1}, \ldots, \tilde{\mathbf{D}}_{\ell_{max}})$, where each multipole's covariance is a 3×3 matrix:

$$\tilde{\mathbf{D}}_\ell = \begin{pmatrix} \tilde{C}_\ell^{TT} & \tilde{C}_\ell^{TE} & \tilde{C}_\ell^{TB} \\ \tilde{C}_\ell^{TE} & \tilde{C}_\ell^{EE} & \tilde{C}_\ell^{EB} \\ \tilde{C}_\ell^{TB} & \tilde{C}_\ell^{EB} & \tilde{C}_\ell^{BB} \end{pmatrix}, \tag{24}$$

Similarly its inverse is a block diagonal of the inverses of $\tilde{\mathbf{D}}_\ell$ matrices and therefore simple to compute. The noise covariance matrix is also of this form, the $\tilde{N}_\ell^{XY}$ in each block diagonal is obtained by noise only Monte Carlo simulations. The band power deviations, q_b, take now the following form:

$$q_b = \frac{1}{2} \sum_{b'} \mathcal{F}_{bb'}^{-1} \sum_\ell (2\ell+1) g \,\mathrm{Tr}\left[\tilde{\mathbf{D}}_\ell^{-1} \frac{\partial \tilde{\mathbf{S}}}{\partial q_{b'}} \tilde{\mathbf{D}}_\ell^{-1} \left(\tilde{\mathbf{D}}_\ell^{obs} - \tilde{\mathbf{N}}_\ell\right)\right], \tag{25}$$

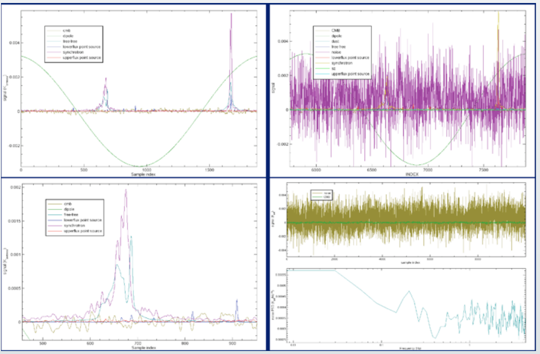
and the Fisher matrix is now given by:

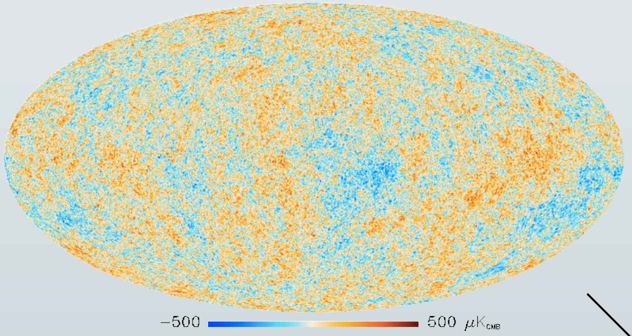
$$\mathcal{F}_{bb'} = \frac{1}{2} \sum_\ell (2\ell+1) g \,\mathrm{Tr}\left[\tilde{\mathbf{D}}_\ell^{-1} \frac{\partial \tilde{\mathbf{S}}_\ell}{\partial q_b} \tilde{\mathbf{D}}_\ell^{-1} \frac{\partial \tilde{\mathbf{S}}_\ell}{\partial q_{b'}}\right]. \tag{26}$$

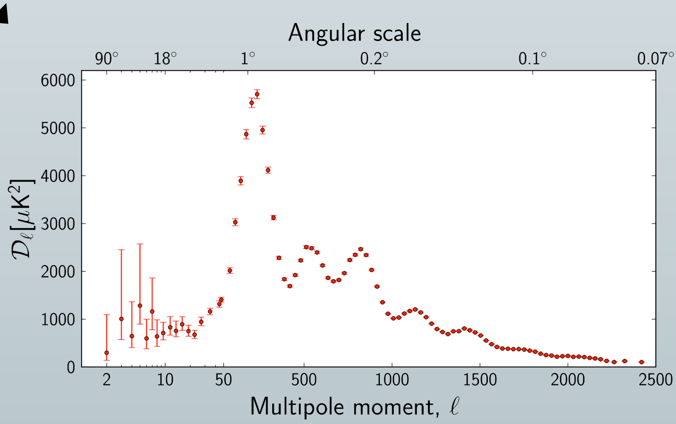
where the band index, b, spans bands in all polarization types. The derivatives of the signal matrices with respect to the deviations q_b are given by:

Angular Power Spectrum

$N \approx 10^{11}$ $N \approx 10^{7}$ $N \approx 10^{3}$



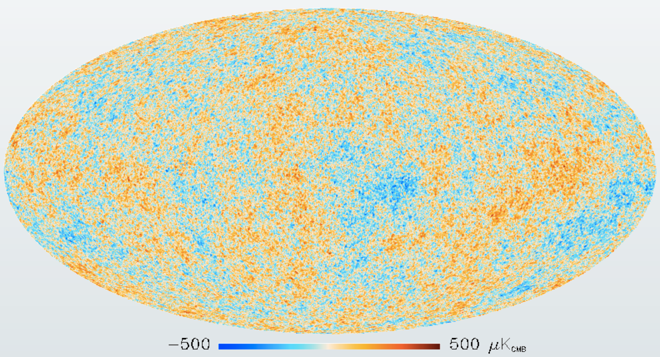




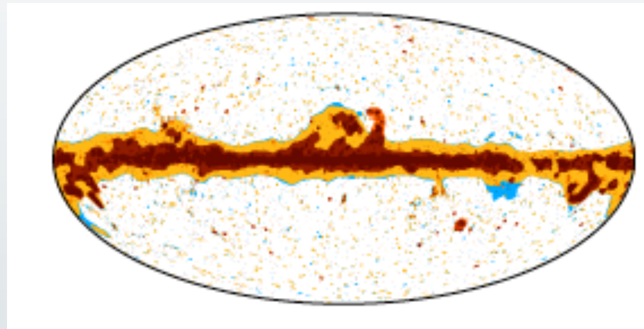
How Planck Unveils the most Ancient Light

Angular Power Spectrum

Maps

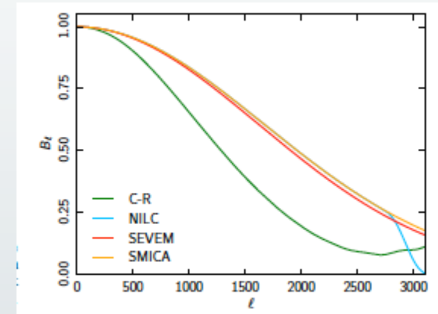


Masks

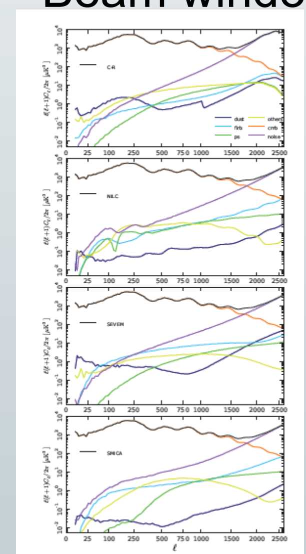


Masks $K_{ll'}$

Beams

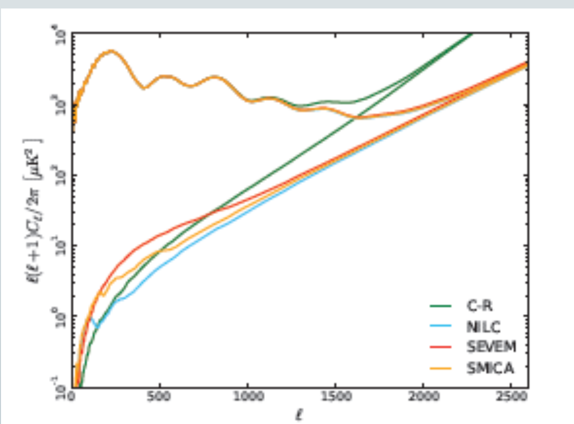


Beam windows W_l



Parameters

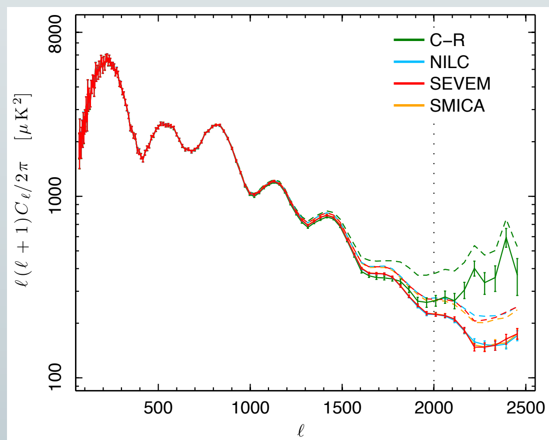
Spectra of residual fg from sims



Pseudo-C_l

Estimated with anafast – Healpix

Signal +Noise and Noise bias

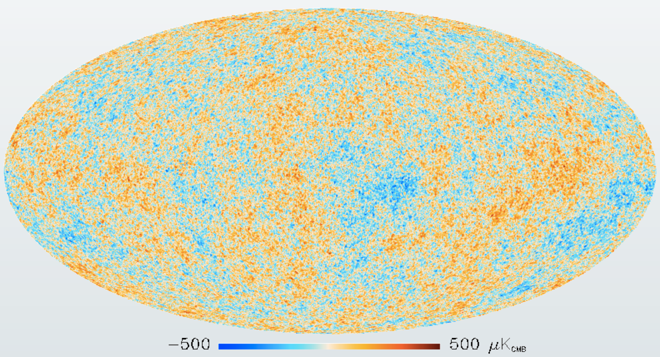


(Full sky) C_l estimated with Xfaster before/after residual foreground s ubtraction (best fit)

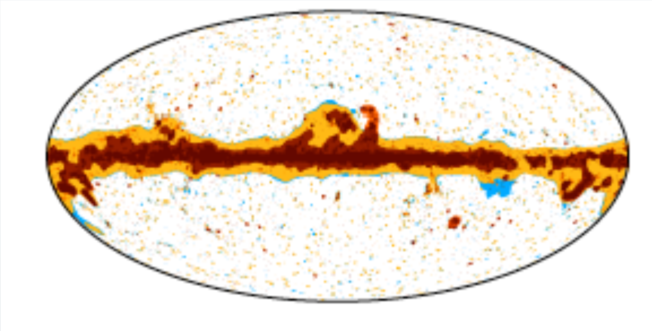
- *Planck 2013 XII, Component Separation*
- *Planck 2013 XV, Likelihood*

Angular Power Spectrum

Maps

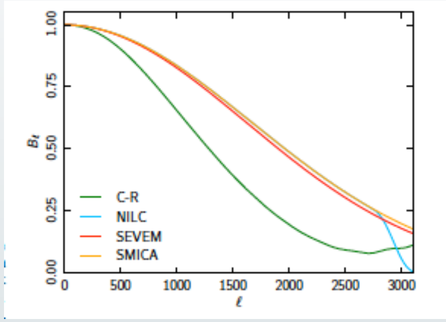


Masks

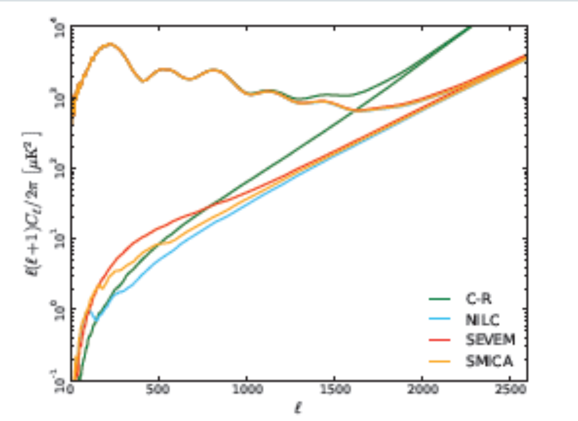


Masks K_{ll}'

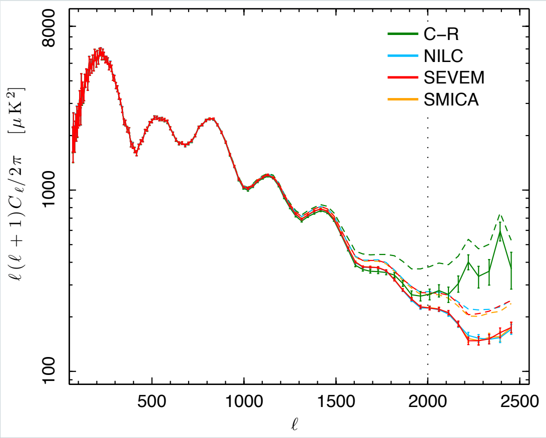
Beams



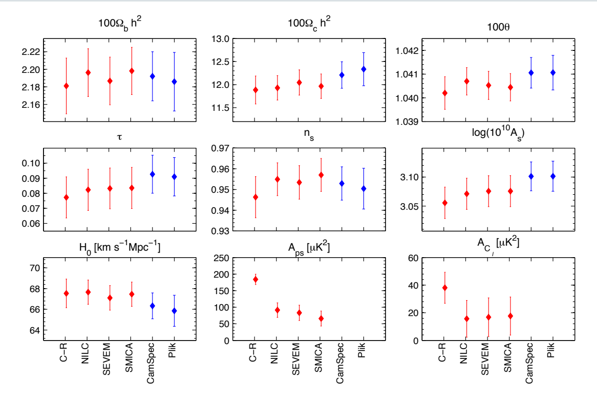
Beam windows W_l



Pseudo-C_l
Estimated with anafast – Healpix
Signal +Noise and Noise bias



(Full sky) C_l estimated with Xfaster before/after residual foreground s ubtraction (best fit)



Parameters
Cosmological and residual foreground: point sources ,CIB

Filter functions, biases…

Xfaster is a quadratic maximum likelihood estimator allows to go straight from the map to parameters

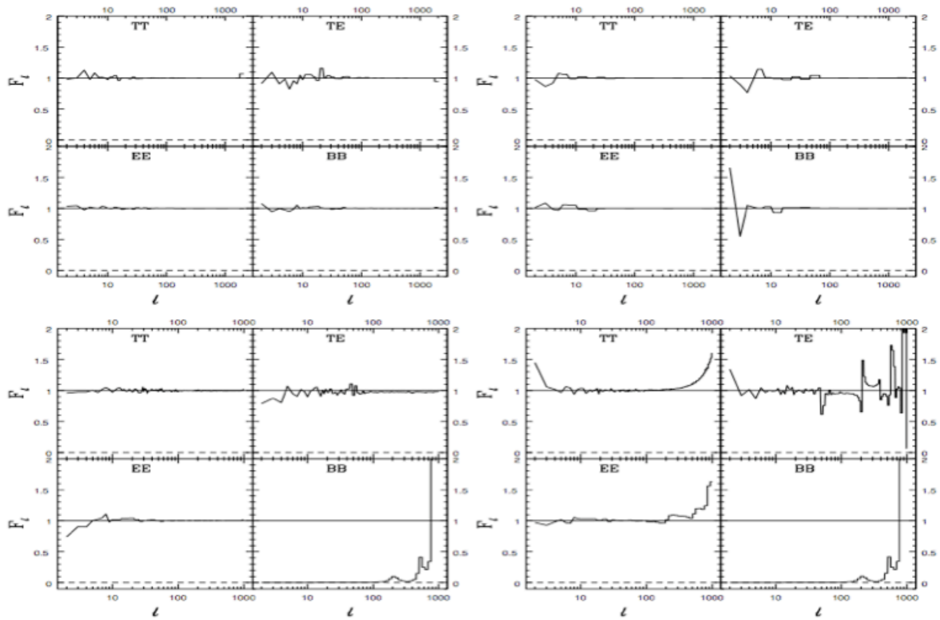


Figure 7. Transfer Function: Top row - for Phase 1a (left hand side) and Phase 1b (right hand side); Bottom row - for Phase 2a symmetric beam (left hand side) and Phase 2b asymmetric beam (right hand side).

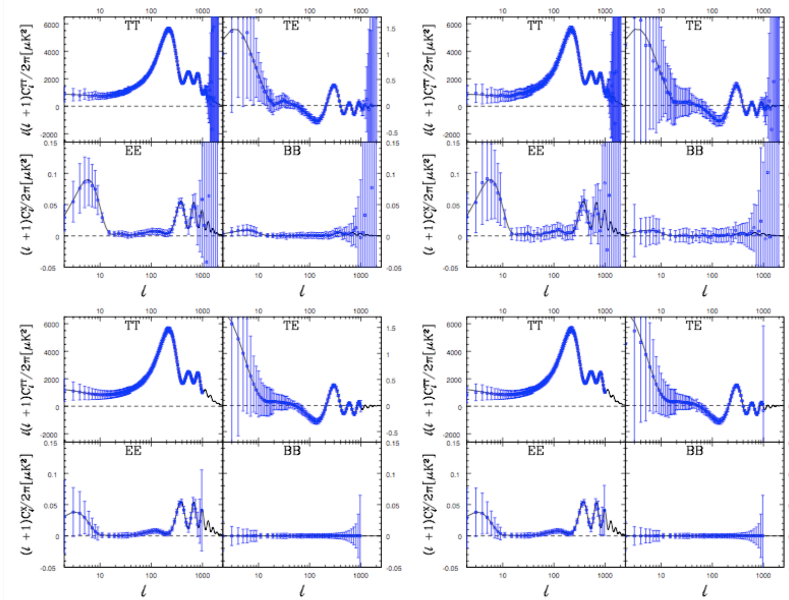


Figure 8. Power spectrum obtained with `XFaster` in average mode, meaning that we replaced the observed map by the average of the signal+noise simulated maps (blue). Top row—Phase 1a (left hand side) and Phase 1b (right hand side) for map generated with a quadruplet of detectors. Bottom row—Phase 2a symmetric beam (left hand side) and Phase 2b asymmetric beam (right hand side) for map generated with all twelve detectors, overplotted with the C_ℓ fiducial model used as input in our Phase 1 signal simulations, first year WMAP best fit model (black) for Phase 1 and first year WMAP+CBI+ACBAR best fit model (black) for Phase 2. It serves the purpose of checking for possible biases of the power spectrum estimator—in principle the power spectrum estimated in average mode should follow closely the input signal C_ℓ model used to generate the signal simulations (black).

Estimating the filter functions $F_{ll'}$

$F_{ll'}$ Reflects the pre-processing in TOI and Map Domain <- Use Signal MC simulations

Look for biases in your estimator use the Signal+Noise Monte Carlo simulations

Filter functions – $F_{ll'}$

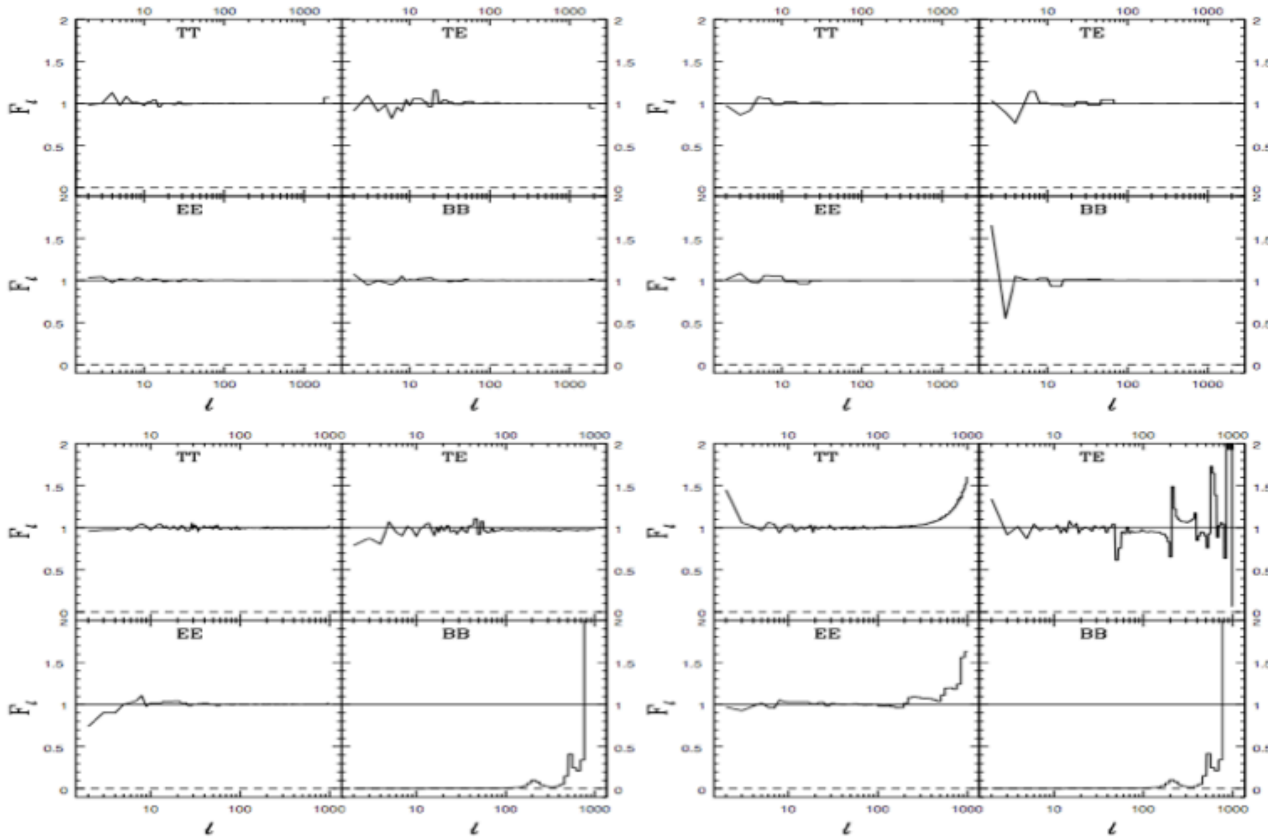


Figure 7. Transfer Function: Top row - for Phase 1a (left hand side) and Phase 1b (right hand side); Bottom row - for Phase 2a symmetric beam (left hand side) and Phase 2b asymmetric beam (right hand side).

$F_{ll'}$ reflects the pre-processing in TOI and Map Domain <- Use Signal MC simulations

Biases

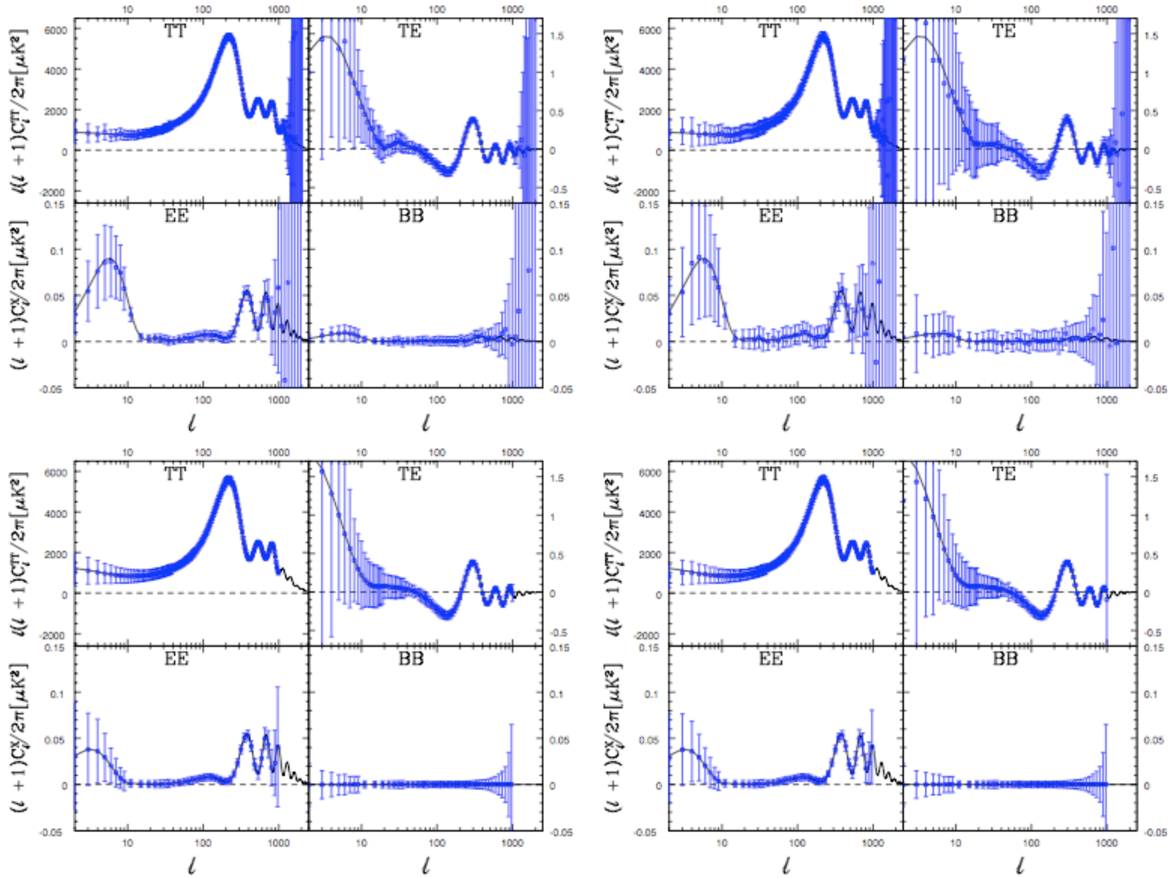
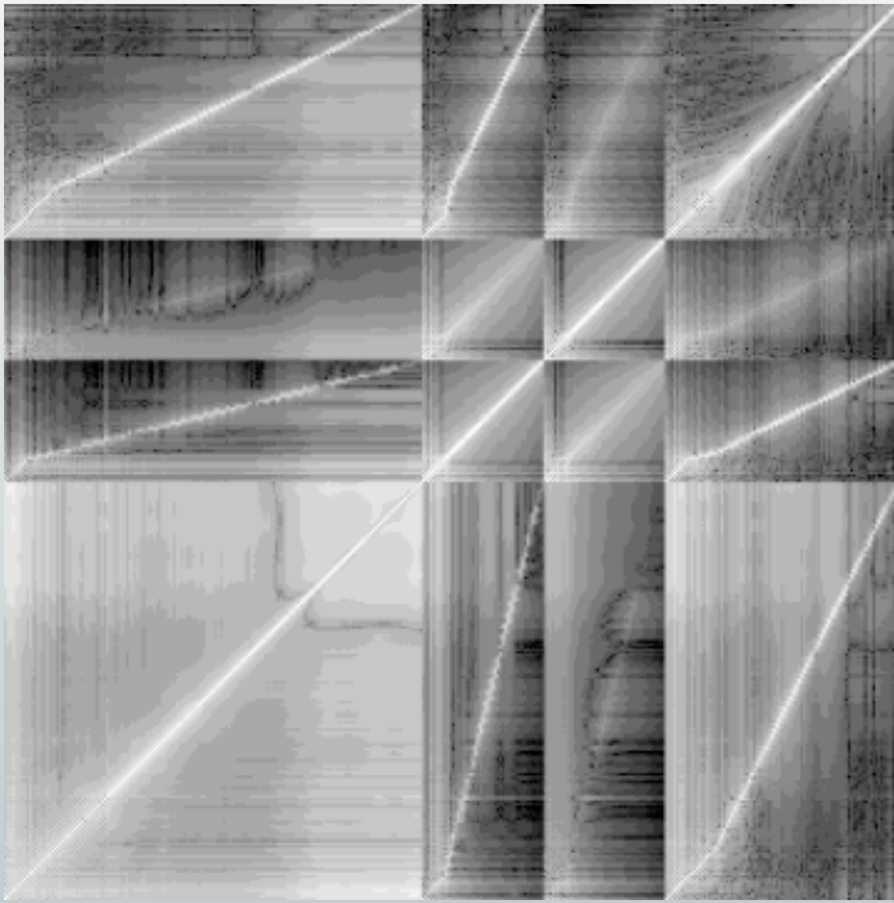


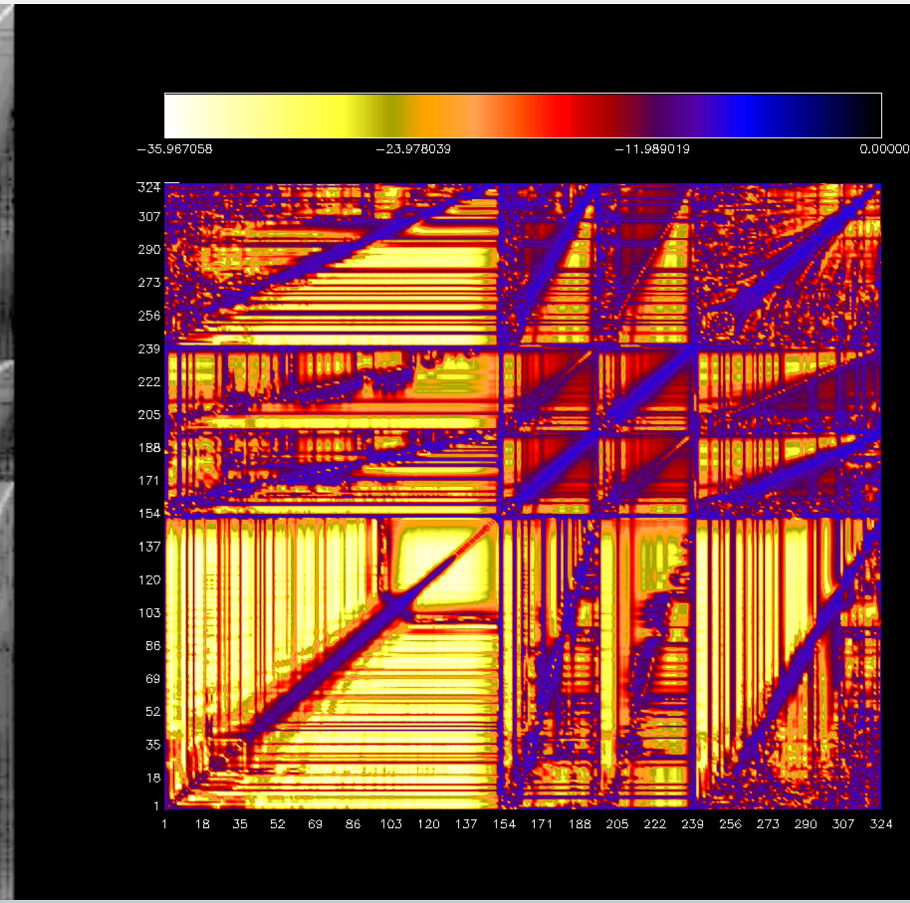
Figure 8. Power spectrum obtained with `XFaster` in average mode, meaning that we replaced the observed map by the average of the signal+noise simulated maps (blue). Top row—Phase 1a (left hand side) and Phase 1b (right hand side) for map generated with a quadruplet of detectors. Bottom row—Phase 2a symmetric beam (left hand side) and Phase 2b asymmetric beam (right hand side) for map generated with all twelve detectors, overplotted with the C_ℓ fiducial model used as input in our Phase 1 signal simulations, first year WMAP best fit model (black) for Phase 1 and first year WMAP+CBI+ACBAR best fit model (black) for Phase 2. It serves the purpose of checking for possible biases of the power spectrum estimator—in principle the power spectrum estimated in average mode should follow closely the input signal C_ℓ model used to generate the signal simulations (black).

Look for biases in your estimator use the Signal+Noise Monte Carlo simulations

Covariance matrices - $C_{ll'}$

Logarithm of the absolute value of the 'normalized' (ie set to 1 at the maximum value) inverse of the Fisher matrix (covariance matrix)





TT, EE, BB and TE modes are displayed sequentially from bottom left-hand side corner to the upper right-hand side corner along the diagonal

Kernels $K_{ll'}$

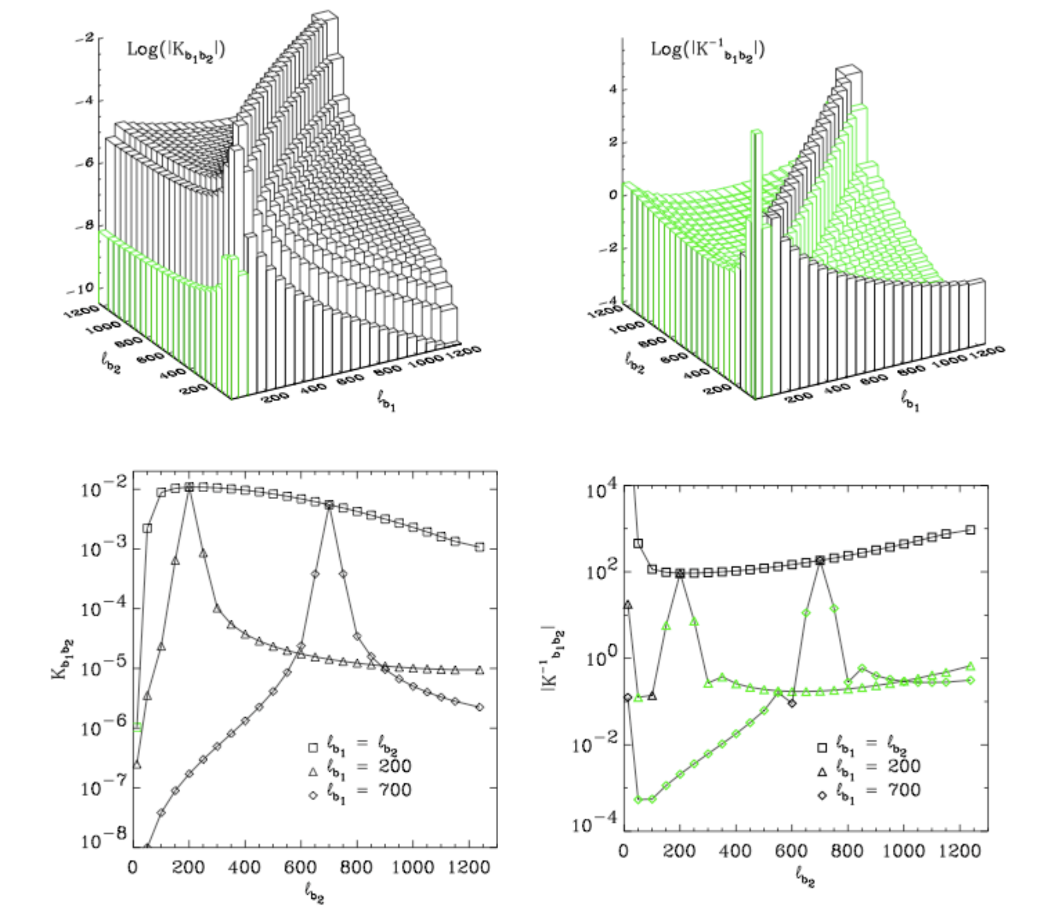
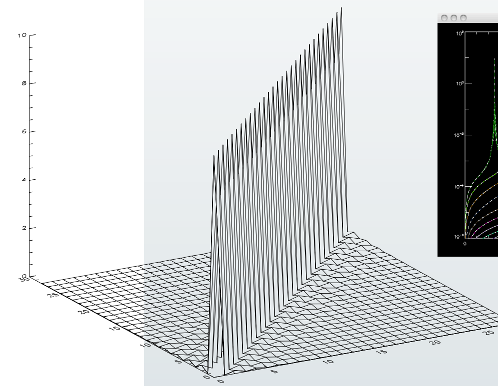
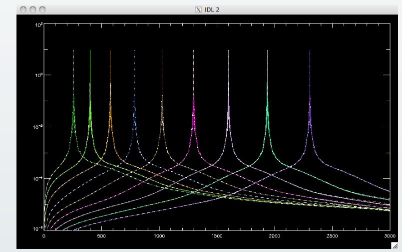
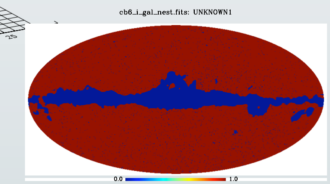
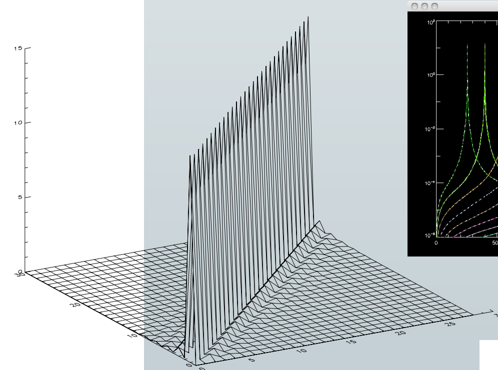


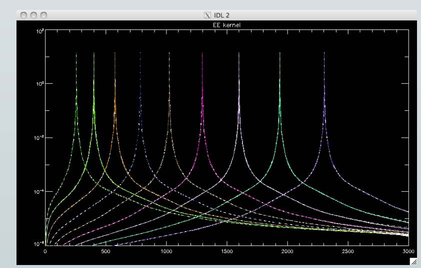
Fig. 3.— Binned power spectrum coupling kernel $K_{bb'}$ and its inverse (absolute values shown, with green color indicating the negative elements) for an elliptically shaped top hat sky window covering 1.8% of the sky. Binwidth is $\Delta\ell = 50$, except for the last bin, for which $\Delta\ell = 150$. The diagonal elements and the $\ell_{b_1} = 200$, and 700 rows of both matrices are shown in the bottom panels.

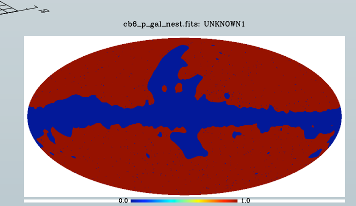






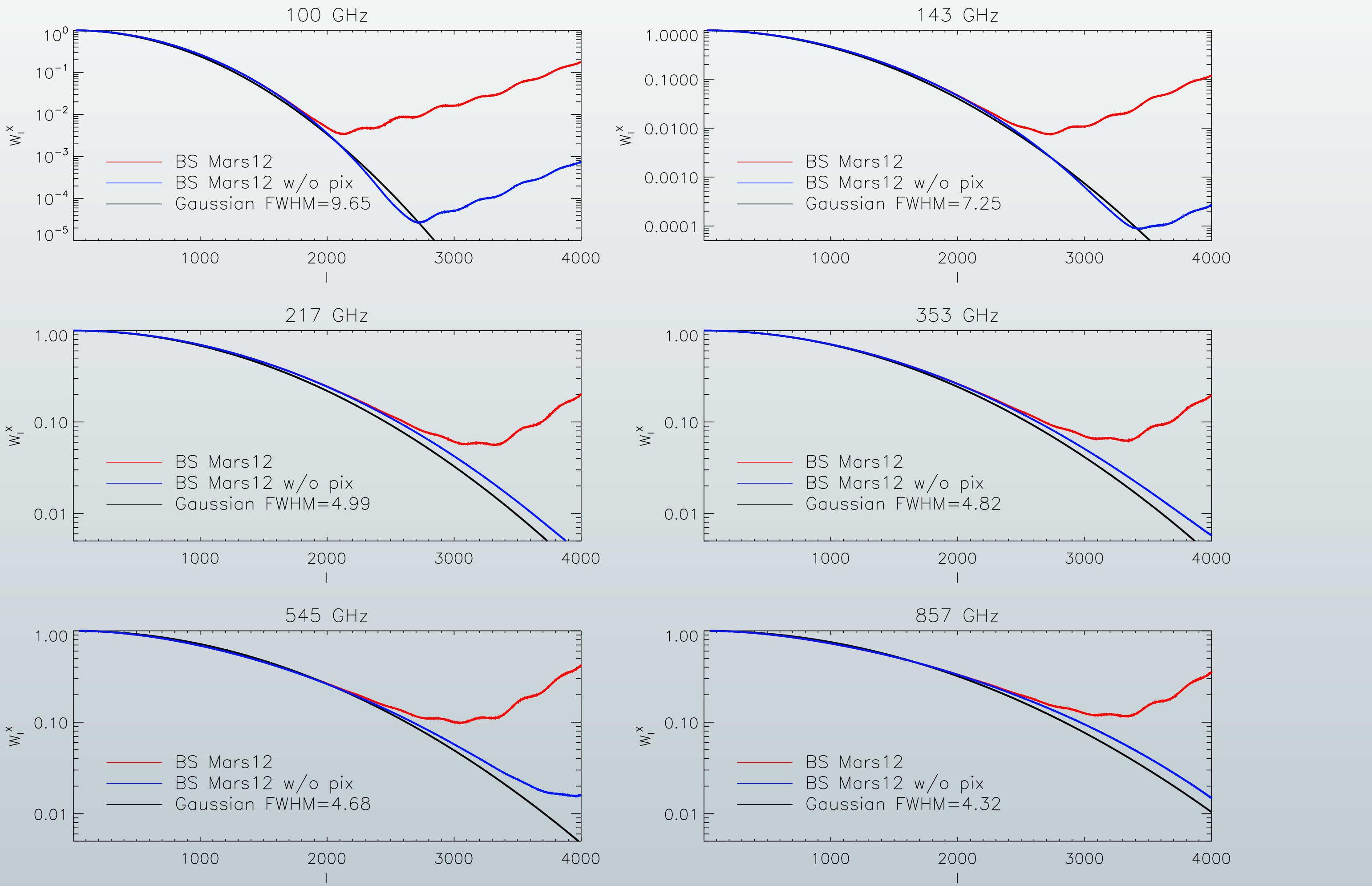






Beam transfer or window functions $W_{ll'}$

Example for Planck HFI channels (estimated with FEBeCoP)



Planck paper: Planck HFI Time Response and Beams

PowerSpectrum (Planck simulations 70GHz)

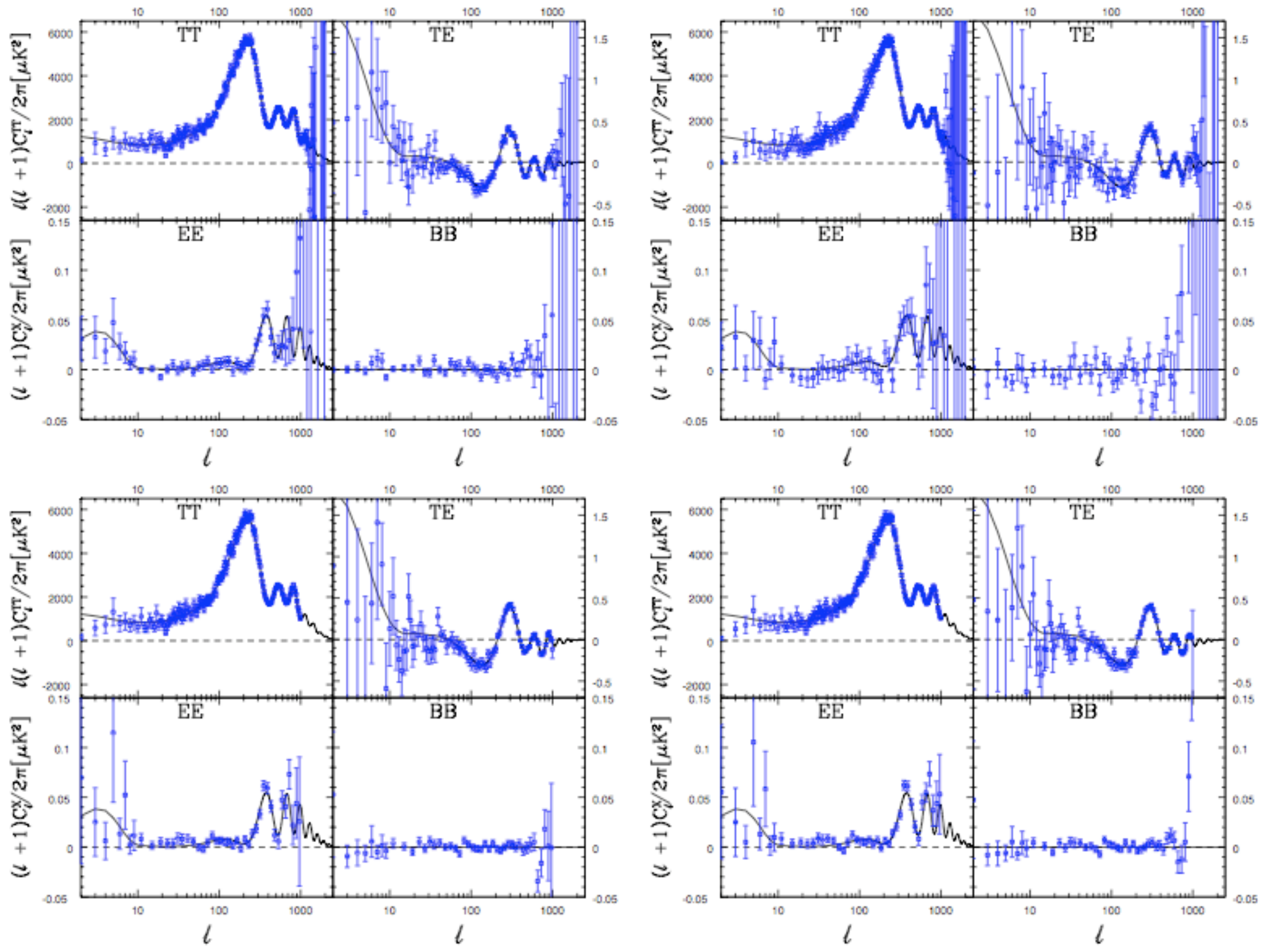
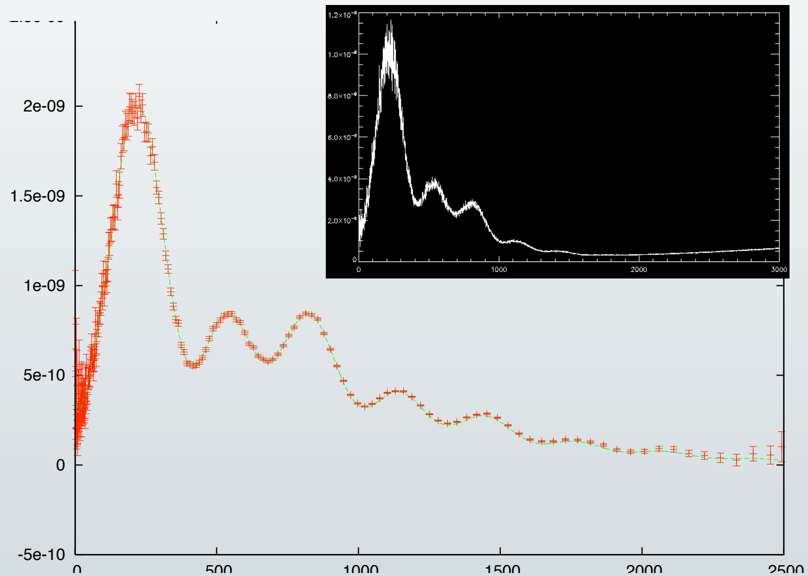
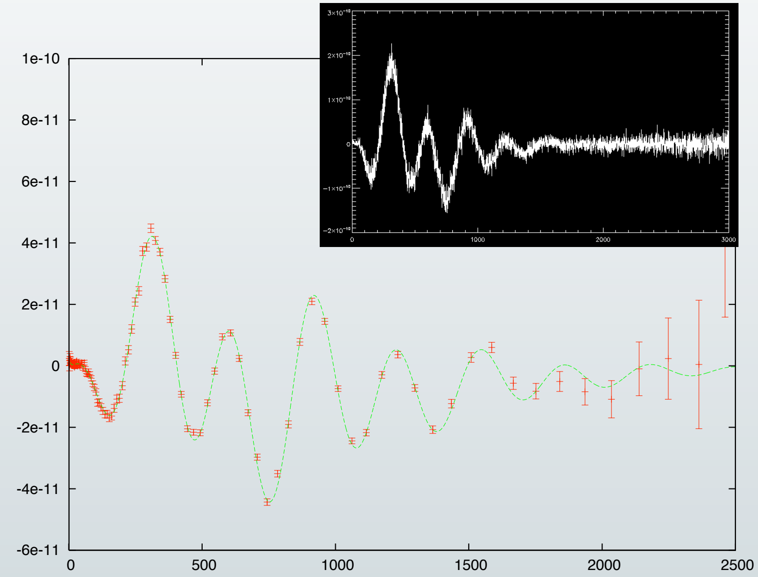
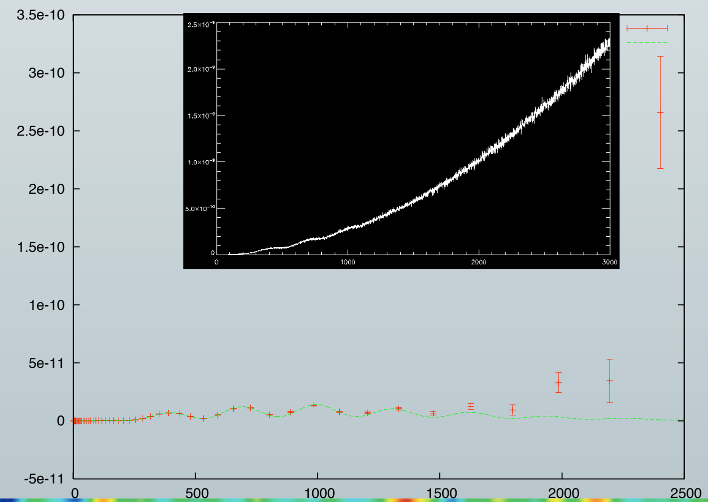


Figure 5. Power spectrum estimated with `XFaster` and 1σ error bars. Top row—Phase 1a (left hand side) and Phase 1b (right hand side) for map generated with a quadruplet of detectors. Bottom row—Phase 2a symmetric beam (left hand side) and Phase 2b asymmetric beam (right hand side) for map generated with all twelve detectors,. The plot displays the estimated power spectrum (blue) of the observed map, overplotted with the C_ℓ fiducial model used as input in Phase 2 signal simulations, first year WMAP+CBI+ACBAR best fit model (black).

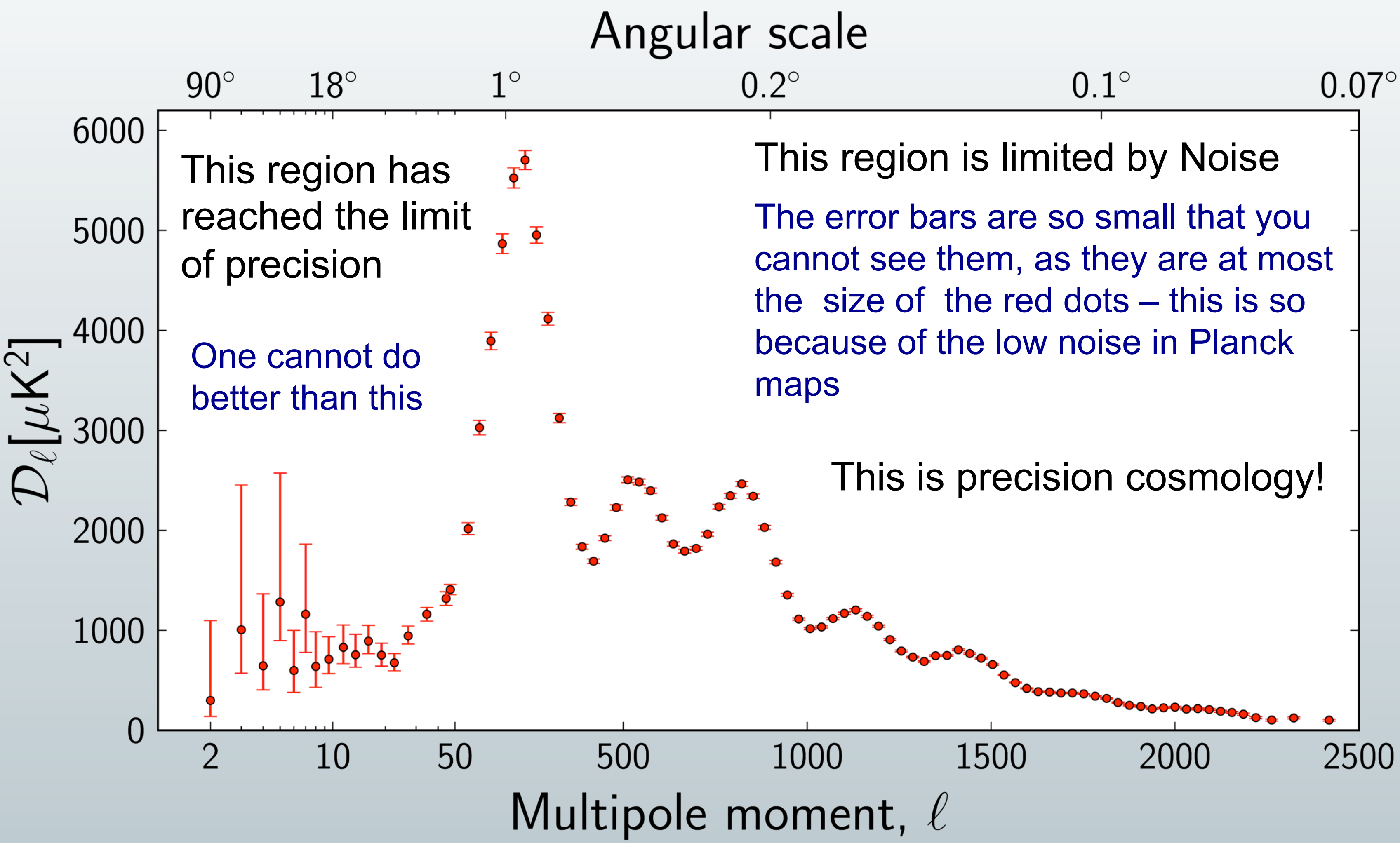
PowerSpectrum (Planck simulations 143GHz)



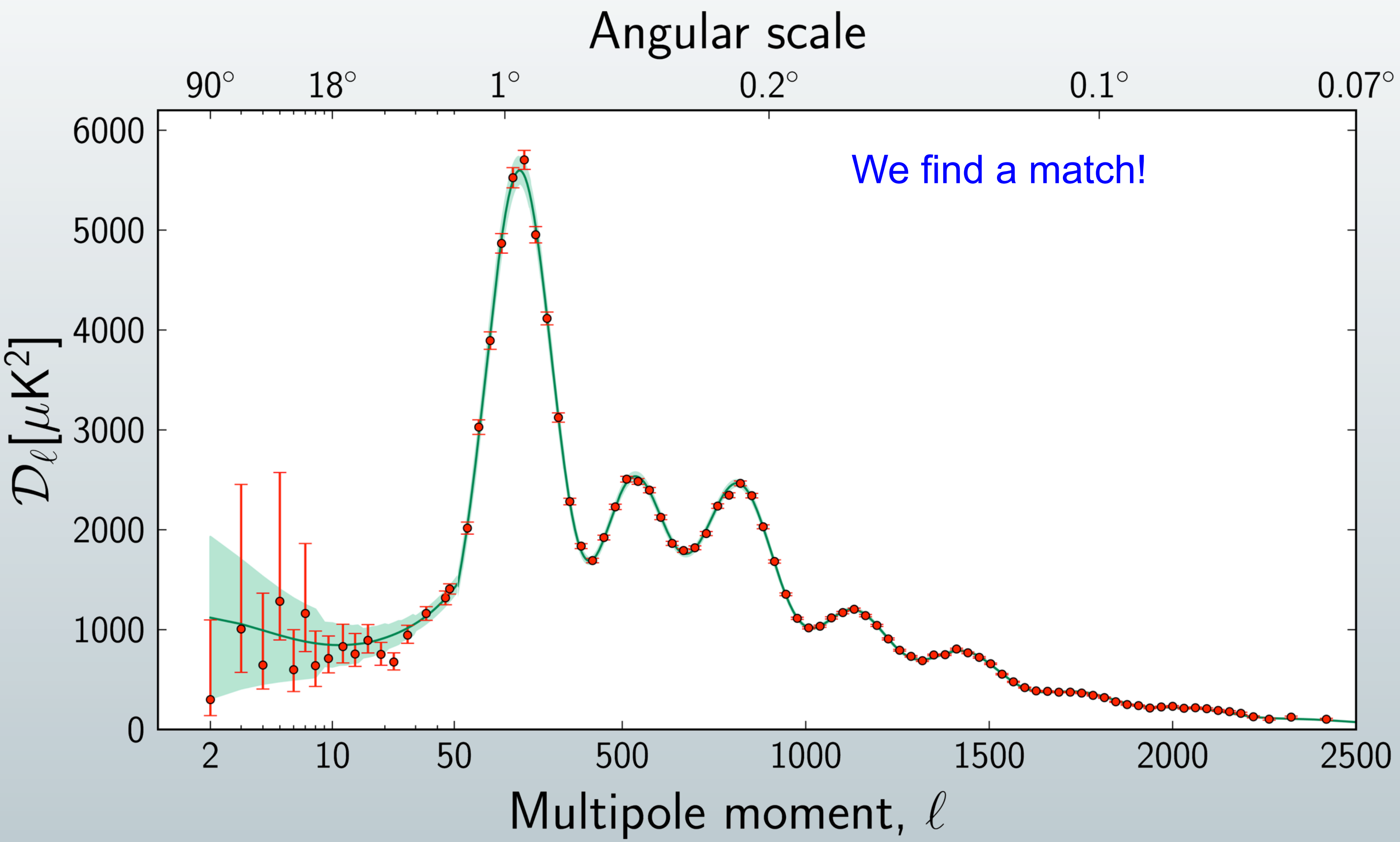




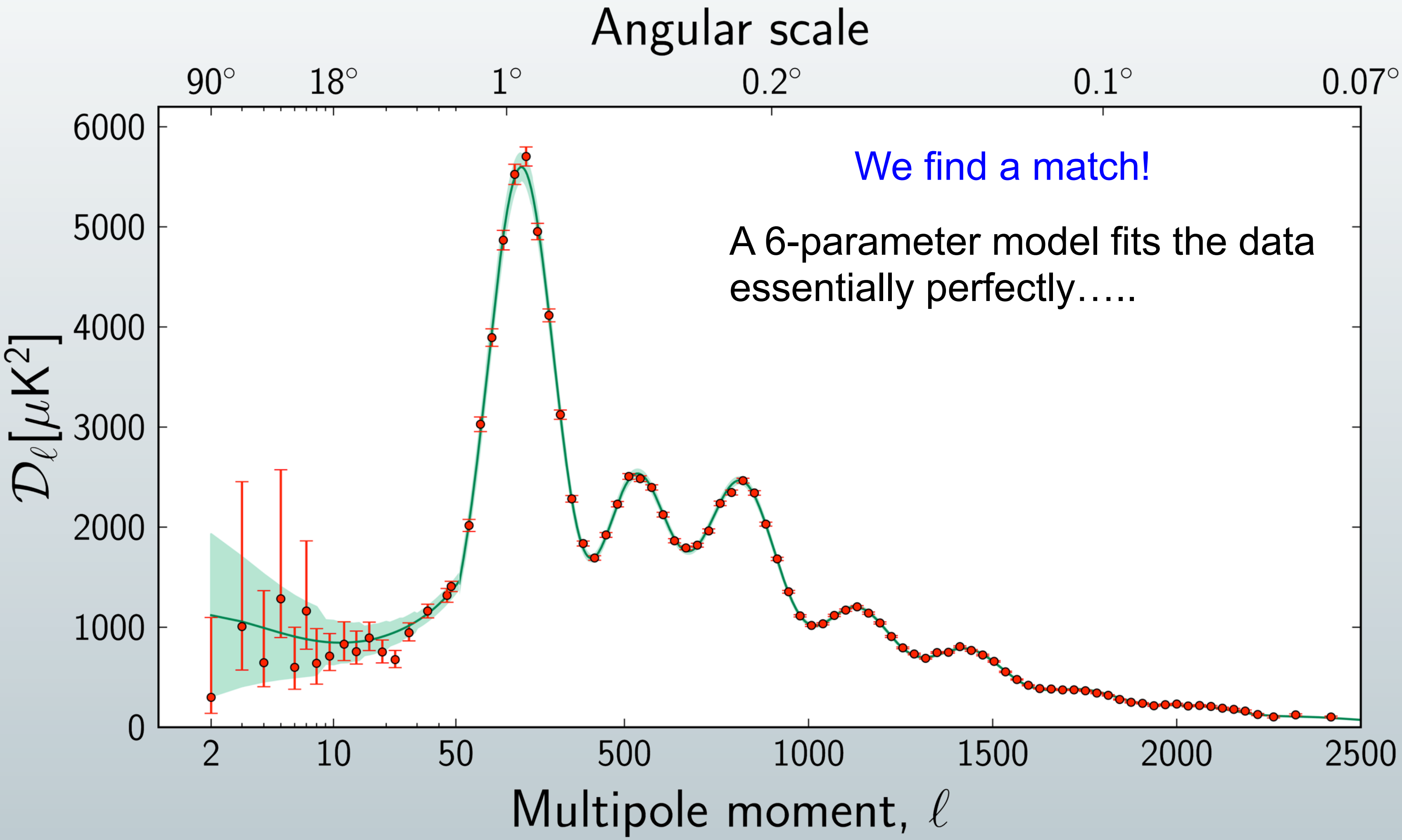
CMB angular power spectrum
what we measure from Planck



CMB angular power spectrum from Planck measurement vs models



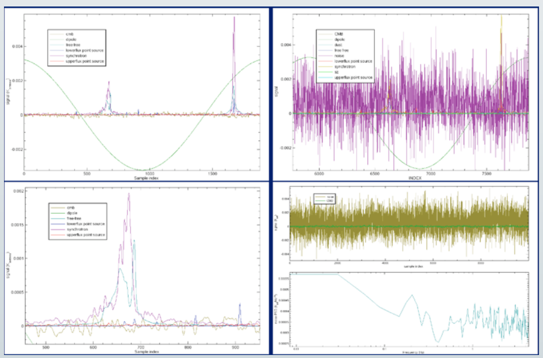
CMB angular power spectrum from Planck measurement vs models



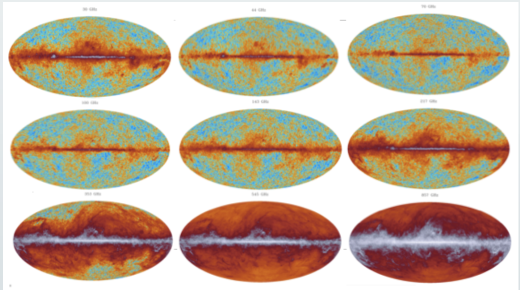
Likelihoods

With the advent of large, high-quality data sets direct extraction of science from the pixelized maps is computationally expensive and in fact unfeasible. Accurate estimation of the angular power spectrum enables the extraction of science with minimal loss of information

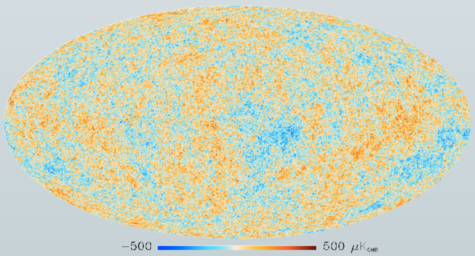
$N \approx 10^{11}$ $N \approx 10^{7}$ $N \approx 10^{3}$



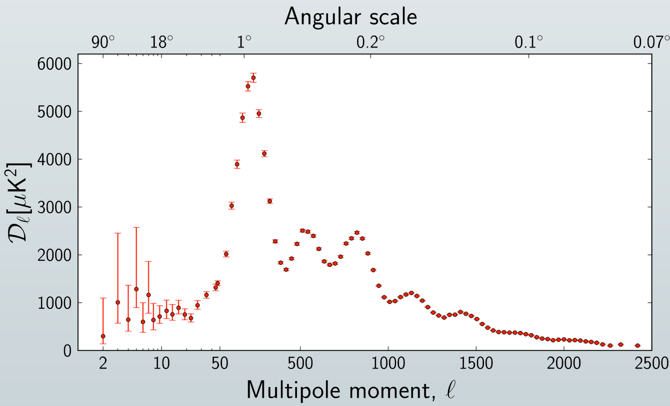
→

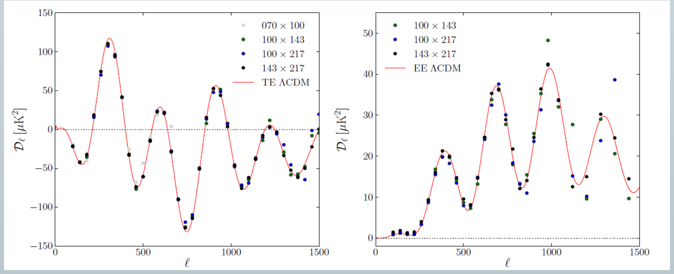


↓



→





To estimate cosmological parameters one needs to compare Theory with Observations – via the Likelihood function

Bayesian statistics

Likelihoods

(Rocha et al, 2011b)

2.1 Exact likelihood in harmonic space

If the CMB is Gaussian, its statistical properties are represented fully by the underlying power spectrum C_ℓ. The multipole harmonic coefficients, $a^X_{\ell m}$ (where X is T, E, or B), on different scales are independent of one another, and we can write

$$\left\langle (a^X_{\ell m})^* a^{X'}_{\ell' m'} \right\rangle = \delta_{\ell\ell'}\delta_{mm'} C^{XX'}_\ell . \tag{4}$$

The $a^X_{\ell m}$ are complex. Under the assumption of Gaussianity, their real and imaginary parts are independent and Gaussian distributed, hence their phases are random. Since the CMB is real, they must satisfy $a^{X*}_{\ell m} = (-1)^m a^X_{\ell -m}$. This means that $a^X_{\ell 0}$ is real.

For $m = 0$ we have

$$P(\boldsymbol{a}_{\ell 0}|\boldsymbol{C}_\ell)\mathrm{d}\boldsymbol{a}_{\ell 0} = \frac{1}{(2\pi)^{3/2}|\boldsymbol{C}|^{1/2}} \exp\left\{ -\frac{1}{2} \boldsymbol{a}^T_{\ell 0} \boldsymbol{C}^{-1}_\ell \boldsymbol{a}_{\ell 0} \right\} . \tag{5}$$

For $m \neq 0$, we have for the real part of the $a^X_{\ell m}$

$$P(\Re\{\boldsymbol{a}_{\ell m}\}|\boldsymbol{C}_\ell)\mathrm{d}\Re\{\boldsymbol{a}_{\ell m}\} = \frac{1}{\pi^{3/2}|\boldsymbol{C}|^{1/2}} \exp\left\{ -\Re\{\boldsymbol{a}_{\ell m}\} \boldsymbol{C}^{-1}_\ell \Re\{\boldsymbol{a}_{\ell m}\} \right\} , \tag{6}$$

where

$$\boldsymbol{a}_{\ell m} = \begin{pmatrix} a^T_{\ell m} \\ a^E_{\ell m} \\ a^B_{\ell m} \end{pmatrix} \tag{7}$$

and

$$\boldsymbol{C}_\ell = \begin{pmatrix} C^{TT}_\ell & C^{TE}_\ell & 0 \\ C^{TE}_\ell & C^{EE}_\ell & 0 \\ 0 & 0 & C^{BB}_\ell \end{pmatrix} , \tag{8}$$

Similarly for the imaginary part.

Combining together all the values of m, we find, for a particular ℓ:

$$P(\hat{\boldsymbol{C}}_\ell|\boldsymbol{C}_\ell) \propto |\hat{\boldsymbol{C}}_\ell|^{\frac{2\ell-3}{2}} |\boldsymbol{C}_\ell|^{-\frac{2\ell+1}{2}} \exp\left\{ -\frac{2\ell+1}{2} \mathrm{Tr}\left(\hat{\boldsymbol{C}}_\ell \boldsymbol{C}^{-1}_\ell \right) \right\} . \tag{9}$$

where

$$\hat{C}^{XX'}_\ell = \sum_{m=-\ell}^{\ell} \frac{(a^X_{\ell m})^* a^{X'}_{\ell m}}{2\ell+1} , \tag{10}$$

and the normalisation is independent of $\boldsymbol{C}_\ell$ and $\hat{\boldsymbol{C}}_\ell$. In data analysis, the measured power spectrum $\hat{\boldsymbol{C}}_\ell$ is a fixed quantity, hence the dependence of the likelihood on this value is generally dropped. In this case, up to a constant, we can write the log-likelihood as

$$-2\ln P(\hat{\boldsymbol{C}}_\ell|\boldsymbol{C}_\ell) = (2\ell+1)\left(\ln|\boldsymbol{C}_\ell| + \mathrm{Tr}\left(\hat{\boldsymbol{C}}_\ell \boldsymbol{C}^{-1}_\ell \right) \right) , \tag{11}$$

i.e., the inverse Wishart distribution. If we consider only one measurement, e.g., one T-mode or B-mode, we can write the likelihood function

Exact Likelihood

as (Bond, Jaffe, & Knox 2000)

$$-2\ln P(\hat{C}_\ell|C_\ell) = (2\ell+1)\left(\ln\left(\frac{C_\ell}{\hat{C}_\ell}\right) + \frac{\hat{C}_\ell}{C_\ell}\right), \tag{12}$$

i.e., the inverse Gamma distribution, where C_ℓ is the theoretical value of C_ℓ^{TT} (or C_ℓ^{BB}) and $\hat{C}_\ell$ is the measured value.

We can write an exact expression for the likelihood function for our measured power spectrum $\hat{C}_\ell$ given the true underlying power spectrum C_ℓ, which is a function of cosmological parameters. Since this likelihood is usually considered in the context of data analysis, it is common to regard the measured $\hat{C}_\ell$ as fixed quantities, and to write the likelihood as

$$\ln P(\hat{C}|C) = \sum_\ell -\frac{(2\ell+1)}{2}\left(\ln\left(\frac{C_\ell}{\hat{C}_\ell}\right) + \frac{\hat{C}_\ell}{C_\ell}\right) + \text{constant}, \tag{13}$$

where the constant depends on $\hat{C}_\ell$. For fixed $\hat{C}_\ell$, this function peaks at $C_\ell = \hat{C}_\ell$. However if we wish to consider the likelihood as a function of $\hat{C}_\ell$ then it is necessary to include the $\hat{C}_\ell$-dependence of the likelihood, in which case it should be written as

$$\ln P(\hat{C}|C) = \sum_\ell \frac{(2\ell-1)}{2}\ln\hat{C}_\ell - \frac{(2\ell+1)}{2}\left(\ln C_\ell + \frac{\hat{C}_\ell}{C_\ell}\right) + \text{constant}. \tag{14}$$

For a fixed underlying power spectrum C_ℓ, this function peaks at $\hat{C}_\ell = \left(\frac{2\ell-1}{2\ell+1}\right)C_\ell$.

This is adequate for a full-sky experiment with an infinitely narrow beam and no instrumental noise. For a partial or "cut" sky it is necessary to account for the correlations between the $\hat{C}_\ell$ that are introduced. In addition, real experiments always have non-uniform noise, and must estimate bandpowers ($\hat{C}_B$) rather than individual $\hat{C}_\ell$. We need to find an appropriate likelihood function that includes the correct correlations and accounts properly for noise.

Likelihoods

Hybrid Likelihood

- *High-l:*
    - Approximations considered such as:
        - Gaussian, Lognormal, Offset Lognormal, Equal Variance, WMAPLike, Offset Lognormal bandpower (Gaussian for TE and Offset Lognormal for TT, EE, BB), SCR Likelihood (1/3 Like), Bond,Jaffe,Knox; Verde et al; Contaldi; Smith,Challinor,Rocha - used by all at CTP
        - These approximations exist for historical reasons:
            - » The Gaussian approx. was shown to be bias-low at low-l - the uncertainties are a function of the variable Cl (not the case for a Gaussian dist) - upward fluctuations are given less weight
            - » The Offset Lognormal corrects this by using a function which uncertainties are independent of the variable Cl and uses a offset x, computed from the noise PS and beam to correct at high-l
            - » However Offset Lognormal shown to be biased-high at low-l - WMAP likelihood corrects this by considering a weight combination of Gaussian and Offset Lognormal
            - » However the WMAP likelihood shown to be biased when Non-Gaussianity of the signal (such as that from lensing) is included - SCR like corrects for this bias
    - Issues: how to account properly for the Temperature and Polarization cross Power Attempts:
        - Xfaster Likelihood - Contaldi, Rocha
        - HL Likelihood - Hamimeche, Lewis

Likelihood for high l

- Approximations :

Bond,Jaffe,Knox

Verde et al.

Contaldi

...

  - Gaussian

$$\mathcal{P}_{Gauss}(\hat{\mathbf{C}}|\mathbf{C}) \propto exp\left\{-\frac{1}{2}(\hat{\mathbf{C}}-\mathbf{C})^T\mathbf{S}^{-1}(\hat{\mathbf{C}}-\mathbf{C})\right\}$$

  - *Lognormal, Offset Lognormal*

$$\mathcal{P}_{LN}(\hat{\mathbf{C}}|\mathbf{C}) \propto exp\left\{-\frac{1}{2}(\hat{\mathbf{z}}-\mathbf{z})^T\mathbf{M}(\hat{\mathbf{z}}-\mathbf{z})\right\}$$

$$z_\ell = \ln(C_\ell + x_\ell)$$
$$M_{\ell\ell'} = (C_\ell + x_\ell)S^{-1}_{\ell\ell'}(C_{\ell'} + x_{\ell'})$$

  - *Equal Variance*

$$\ln\mathcal{L} = -\frac{1}{2}G\left[e^{-(z-\hat{z})} - (1-(z-\hat{z}))\right]$$

$$z = \ln(q_b + q_b^N)$$
$$\sigma_z = \sqrt{F^{-1}_{bb'}/(q_b + q_b^N)}$$
$$G = \left[e^{-\sigma_z} - (1-\sigma_z)\right]^{-1}$$

  - *WMAPLike*

$$\ln\mathcal{P}_{WMAP}(\hat{\mathbf{C}}|\mathbf{C}) = \frac{1}{3}\ln\mathcal{P}_{Gauss} + \frac{2}{3}\ln\mathcal{P}_{LN}$$

  - *Offset Lognormal bandpower (Gaussian for TE and Offset Lognormal for TT, EE, BB)*
  - *SCR Likelihood (1/3 Like)* *Gaussian on* $\hat{x} = \hat{C}_l^{1/3}$

Smith, Challinor, Rocha

- Attempts to properly account for the Temperature and Polarization
  - *Xfaster Likelihood*

$$\ln L = -\frac{1}{2}\sum_\ell g_\ell(2\ell+1)\left[\frac{C_\ell^{obs}}{(\tilde{C}_\ell + \langle N_\ell\rangle)} + \ln\left(\tilde{C}_\ell + \langle N_\ell\rangle\right)\right]$$

Rocha, Contaldi

  - *HL Likelihood*

$$\ln L(C_l \mid \hat{C}_l) = -\frac{1}{2}\frac{2l+1}{2}\sum_i [g(D_{l,ii})]^2 = \frac{2l+1}{2}Tr[g(D_l^2)] \leftrightarrow with \leftrightarrow g(x) = sign(x-1)\sqrt{2(x-\ln(x)-1)}$$

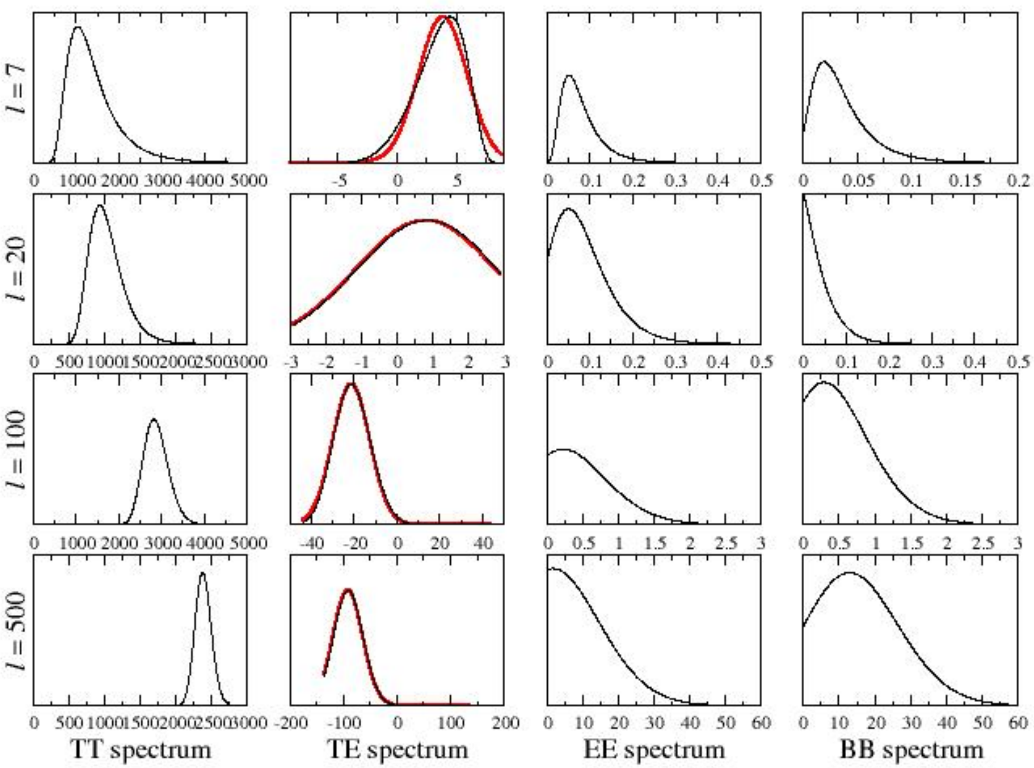
Hamimeche, Lewis

Comparing Likelihoods - Exact full-sky likelihood

Joint (T,P) exact full-sky like

$$\ln P(\hat{C}_l \,|\, C_l) = -\frac{2l+1}{2}\left(\ln|C_l| + Tr\left(\hat{C}_l\, C_l^{-1}\right)\right)$$

Slices through the Exact full-sky likelihood for Simulations



Exact full-sky like for one component, say T:

$$\ln P(\hat{C}_l \,|\, C_l) = -\frac{2l+1}{2}\left(\ln\left(\frac{C_l}{\hat{C}_l}\right) + \frac{\hat{C}_l}{C_l}\right)$$

Gaussian likelihood (mode and variance of Gaussian like for TE depends on TT and EE - not enough to compute products of independent TT,TE,EE likelihoods - the trick is to find a way of coupling these components reliably)

Likelihood at low-l

- **Pixel based Likelihood**

  Considering a Gaussian assumption for the Likelihood of the observed data (might it be, T or $a_{\ell m}$):

$$L(\mathbf{a}|\mathbf{p}) = \frac{1}{2\pi^{N/2}|\mathbf{C}|^{1/2}} \exp\left(-\frac{1}{2}\mathbf{a}.\mathbf{C}^{-1}.\mathbf{a}^t\right)$$

  where, $\mathbf{C}$ is the covariance of the data, and $\mathbf{p}$ is the set of model parameters, with $\mathbf{C}(\mathbf{p}) = \mathbf{S}(\mathbf{p}) + \mathbf{N}$.

  For single dish, full sky observations an isotropic signal is diagonal in the Spherical Harmonic space and can be described by a m-averaged power Spectrum C_ℓ on each multipole: $S_{\ell m,\ell' m'} = \delta_{\ell\ell'}\delta_{mm'}C_\ell$, the noise is generally not diagonal.

- **Blackwell-Rao Likelihood**
  - The CMB Gibbs sampler provides samples from the joint distribution $P(s, C_\ell|d)$, where s is a *full sky* map consistent with the data. The collection of all s covers the set of all consistent skies.
  - For each of these full-sky maps, the exact likelihood is trivial to compute:

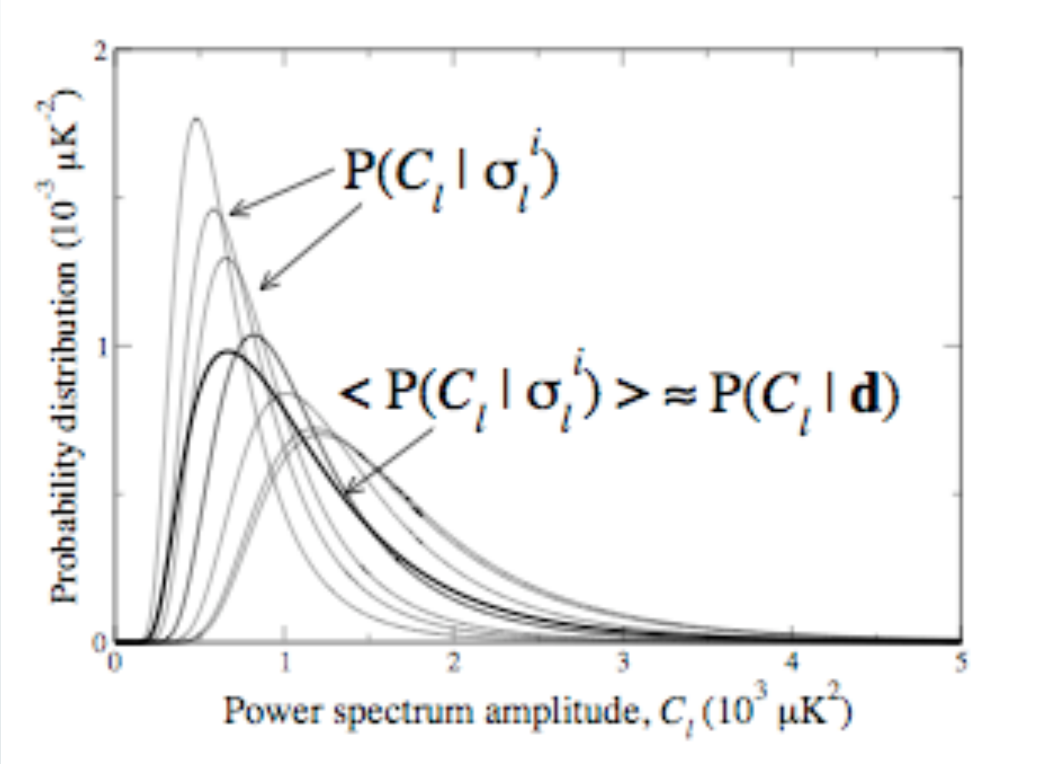
$$-2lnL_i(C_l) = \sum_l (2l+1)[\sigma_l/C_l + ln(C_l/\sigma_l) + ln(\sigma_l)]$$

  where σ_l is the power spectrum of s.
  - The Blackwell-Rao estimator is simply the average over these individual likelihoods for all samples generated by the Gibbs sampler:

$$L_{BR}(C_l) = < L_i(C_l) >_{gibbs}$$

- Pixel based likelihood
- BfLike – Brute force likelihood
- Blackwell-Rao Likelihood or fitting (Gibbs)
- Teasing
- …..

BR-Likelihood



High-l: XFaster Likelihood

- Assume the alms in the cut-sky follow the same distribution of the full sky - expression follows nicely from the exact full sky inverse Wishart distribution

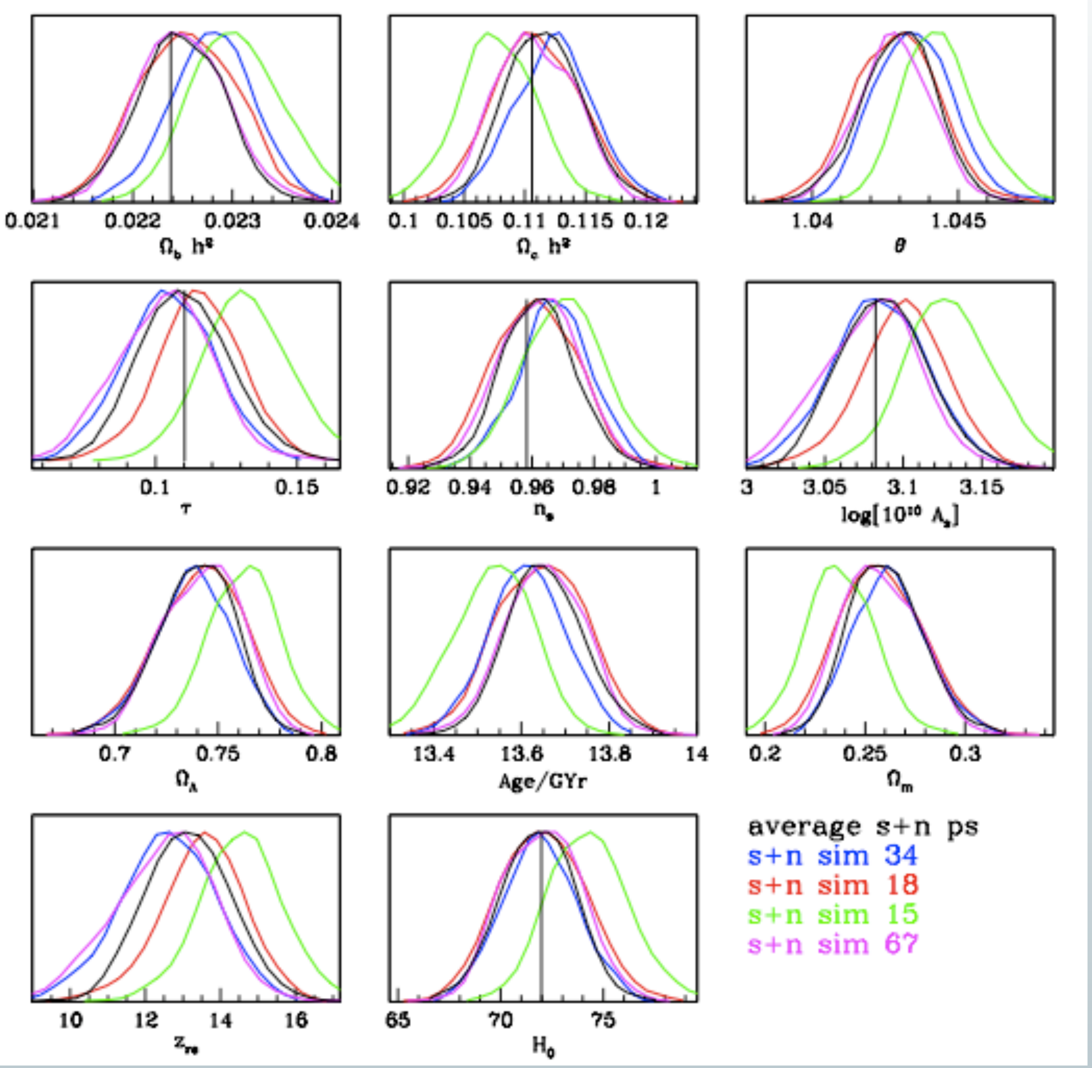
- Estimated in single ℓ – provides a straight path from maps (via their pseudo-Cls) to parameters bypassing the band-power spectra step

$$\ln L = -\frac{1}{2}\sum_\ell g_\ell(2\ell+1)\left[\frac{C_\ell^{obs}}{(\tilde{C}_\ell + \langle N_\ell\rangle)} + \ln\left(\tilde{C}_\ell + \langle N_\ell\rangle\right)\right]$$

$$\ln L = -\frac{1}{2}\sum_\ell g_\ell(2\ell+1)\left(Tr\left(\tilde{D}_\ell^{obs}\left(\tilde{D}_\ell + \langle\tilde{N}_\ell\rangle\right)^{-1}\right) + \ln\left|\tilde{D}_\ell + \langle\tilde{N}_\ell\rangle\right|\right)$$

Rocha, Contaldi, Colombo, Bond, Gorski & Lawrence, sub. MNRAS 2010
Rocha et al., 2010

Likelihoods validation via parameters



CosmoMC

MCMC sampler

estimate cosmological parameters and residual foreground parameters

Hybrid Likelihood

- Low-l
  - **Commander** – Gibbs sampling
- High-l
  - Spectra-based:
    - **CamSpec –** Baseline
    - **PLIK**
  - Map-based:
    - **XFcmb - XFaster**

Planck High-l Likelihoods
CamSpec

- Estimate pseudo-Cls

$$\tilde{C}_\ell^{ij} = \frac{1}{2\ell+1}\sum_m \tilde{T}_{\ell m}^{i}\tilde{T}_{\ell m}^{j\dagger},$$

only cross-spectra is used

- Estimate the deconvolved spectra:

$$\hat{C}^{T_{ij}} = M_{ij}^{TT}\,\tilde{C}^{T_{ij}}.$$

$$\hat{\mathcal{M}} = (\hat{X} - \langle\hat{X}\rangle)(\hat{X} - \langle\hat{X}\rangle)^T.$$

with

$$\hat{X} = \mathrm{Vec}(\hat{C})$$

- The deconvolved spectra is efficiently combined within a frequency pair after a small recalibration factor, taking into account respective beam transfer functions and noise levels; the Covariance matrix is computed for a fiducial model

- Estimate Likelihood as a Gaussian: $p = e^{-S}$ with $S = \frac{1}{2}\left(\hat{X} - X\right)^T \hat{\mathcal{M}}^{-1}\left(\hat{X} - X\right).$

$$\hat{X} = (\hat{C}_\ell^{100\times100}, \hat{C}_\ell^{143\times143}, \hat{C}_\ell^{217\times217}, \hat{C}_\ell^{143\times217}),$$

coupled to a parametric model of the CMB and FG power spectra

- Calibration, beam uncertainties and instrumental noise

$$B^{ij}(\ell) = B^{ij}_{\mathrm{mean}}(\ell)\exp\left(\sum_{k=1}^{n_{\mathrm{modes}}} g_k^{ij}E_k^{ij}(\ell)\right),$$

construct a Gaussian posterior dist. of beam eigenmodes from the associated Covariance

- Noise pseudo spectra estimated from half-ring difference maps + noise rms /pixel

Planck High-l Likelihoods
PLIK

Start from the full-sky exact likelihood for a Gaussian signal, which for $\mathbf{N}_{map}$ detector maps is given by:

$$p(\text{maps}|\theta) \propto \exp -\{\sum_{\ell}(2\ell+1)\mathcal{K}(\hat{\mathbf{C}}_{\ell}, \mathbf{C}_{\ell}(\theta))\},$$

K(**A**,**B**) - Kullback divergence between two **n**-variate zero-mean Gaussian distributions with covariance matrices **A** and **B**.

$$\mathcal{K}(\mathbf{A}, \mathbf{B}) = \frac{1}{2}\left[\mathrm{tr}(\mathbf{A}\mathbf{B}^{-1}) - \log\det(\mathbf{A}\mathbf{B}^{-1}) - n\right].$$

Bin the power spectra in such a way that off-diagonal terms of the covariance due to sky cuts are negligible

$$p(\text{maps}|\theta) \propto \exp -\mathcal{L}(\theta), \quad \text{with} \quad \mathcal{L}(\theta) = \sum_{q=1}^{Q} n_q \mathcal{K}(\hat{\mathbf{C}}_q, \mathbf{C}_q),$$

The Plik bin width is Δl= 9 from l = 100 to l = 1503; Δl= 17 to l= 2013; Δl= 33 to l_{max} = 2508. This ensures that correlations between any two bins are smaller than 10 %.

Binned Likelihood approximation - Computational speed, and it agrees well with the primary likelihood - well suited for performing an extensive suite of robustness tests + instrumental effects can be investigated quickly - assess the agreement between pairs of detectors within a frequency channel, such as individual detector calibrations and beam errors.

Jointly estimate the noise together with all other parameters using both auto and cross-spectra – then fix the noise estimates, and use the fiducial Gaussian approximation to explore the remaining free parameters excluding the autospectra, optionally including only specific data combinations

Planck High-l Likelihoods XFcmb

Band powers estimated with XFaster for each of the CMB maps generated by

SMICA, Commander-Ruler, NILC, SEVEM

XFaster: an approximation to the iterative, Maximum likelihood, quadratic band power estimator based on a diagonal approximation to the quadratic Fisher matrix estimator

$$\tilde{C}_\ell = \sum_b q_b \tilde{C}^S_{b\ell} = \sum_b \left(\frac{1}{2} \sum_{b'} \mathcal{F}^{-1}_{bb'} \sum_\ell (2\ell+1) g_\ell \frac{\tilde{C}^S_{b'\ell}}{(\tilde{C}_\ell + \langle \tilde{N}_\ell \rangle)^2})(\tilde{C}^{obs}_\ell - \langle \tilde{N}_\ell \rangle) \right) \tilde{C}^S_{b\ell}$$

$$\mathcal{F}_{bb'} = \frac{1}{2} \sum_\ell (2\ell+1) g_\ell \frac{\tilde{C}^S_{b\ell} \tilde{C}^S_{\ell b'}}{(\tilde{C}_\ell + \langle \tilde{N}_\ell \rangle)^2}$$

The iterative scheme starts from a flat spectrum model - the result is a band power spectrum and the associated Fisher matrix (hence uncertainty of the band powers)

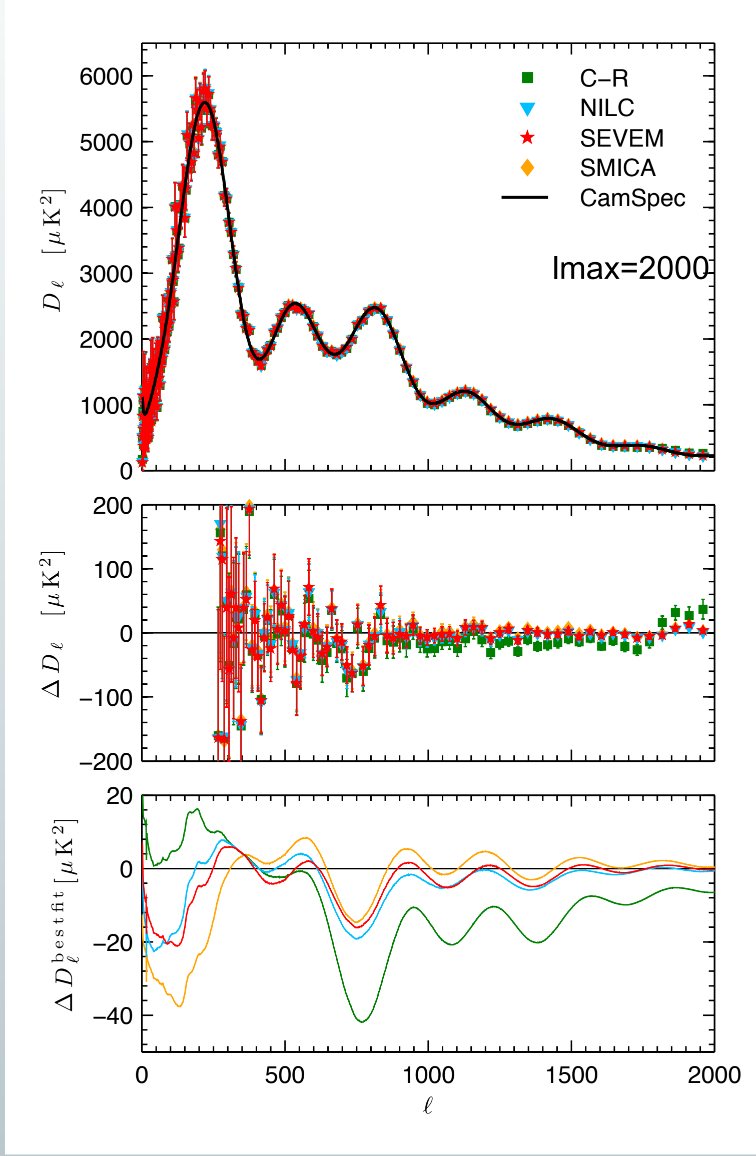
Use a Gaussian Correlated likelihood and a MCMC sampler and PICO for 70<l<2000

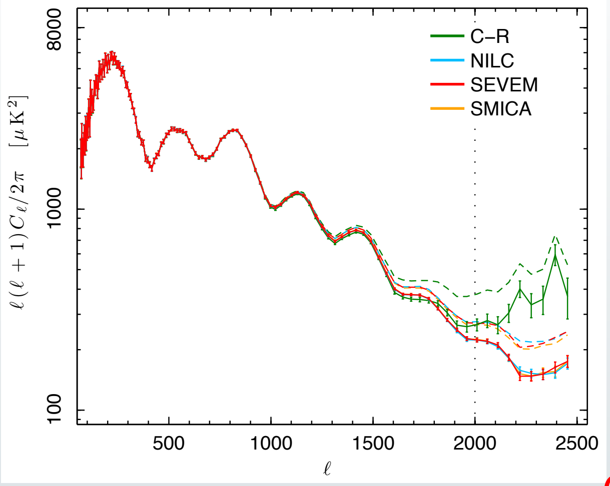
- 6 cosmological parameters
- A_{ps} - the amplitude of a Poisson component , $C_l = A_{ps}$=constant
- Acl - the amplitude of a clustered component with shape:

D_l at l = 3000 in units of μK^2

$$D_\ell = \ell(\ell+1)C_\ell / 2\pi \propto \ell^{0.8}$$

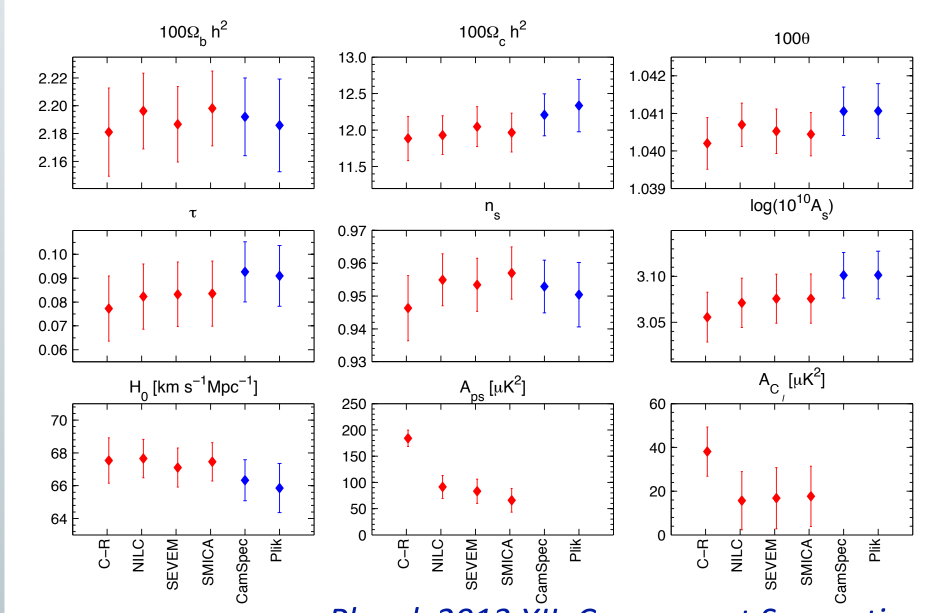
Planck - an example





XFaster

CosmoMC



- *Planck 2013 XII, Component Separation*
- *Planck 2013 XV, Likelihood*

Appendix

Graça Rocha
JPL/Caltech
graca@caltech.edu

Vth INPE Advance Course on Astrophysics, INPE, September 2013