Fingerprinting the Universe: Error in Error-bars





Dipak Munshi Institute of Astronomy - Cambridge



Planck Consortium Organizational Structure and the TLA soup

- * WG = Working Groups, e.g, PS, NG, Secondaries, Comp. Sep. etc.
- \star L2/L3 = Level -2/3

.

- ***** DPC = Data Processing Centre
- ***** LFI = Low Frequency Instrument
- ***** HFI = High Frequency Instrument
- * CPAC,LPAC,MPAC = Cambridge/London/Munich Planck Analysis Centre
- ***** DIPAK = Suggestions Wellcome for suitable bacronyms!

All simulations used in this talk are generated by CTP Working Group and the results obtained are for the Power Sepectrum Estimation Exercise This is a beauty contest for various teams and this exercise will eventually be a part of the larger E2E tests.



Figure 1: "No you cant be the First Author Albert!".

Sociology vs. Science: Authorship and Data rights

CMBology

- \star Generation of seed perturbations is Quantum Mechanical
- * Constrains Early Universe Models : Inflationary Dynamics amplitudes, spectral index and the running Evolution of Perturbations is General Relativistic
- ***** Constrains Geometry and Dynamics of the Universe
- Physics is Linear and Clean separation of Early Universe parametrisation and LSS can be achieved At small scale (high l) non-linearity is induced by SZ, kSZ, lensing etc.

Focus of this talk is Model testing and NOT Model building

From TOD to Precision Cosmology

* Step-0: Data from the Satelite: missing data, glitches, satellite pointing variation, cooler noise and other systematics



* Step -1: Map Making: Accuracy vs. Speed: To find nearly OPTIMAL but FAST techniques - Destriping vs Generalised Least Square Techniques

$$\mathbf{d} = \mathbf{A}\mathbf{s} + \mathbf{n} \tag{1}$$

d is time ordered data. s = signal at the pixel where detector is pointing. n = associated noise. $\langle n \rangle = 0$. $\langle nn^t \rangle = N$. Optimal Solution:

$$(\mathbf{A}^{\mathrm{T}}\mathbf{N}^{-1}\mathbf{A})\hat{\mathbf{s}} = \mathbf{A}^{\mathrm{T}}\mathbf{N}^{-1}\mathbf{d} \qquad \mathbf{S} = (\mathbf{A}^{\mathrm{T}}\mathbf{N}^{-1}\mathbf{A})^{-1}$$
(2)

For Planck size data various clever algorithm needs to be employed to tackle the size of the data.

* Step -2: Component Separation : Aim : Frequency Maps to Component Maps, Removal of non-CMB contamination using Frequency Information, Diffuse Galactic or Extra galactic Point Sources for T as well as Q and U maps - Techniques include ICA, MaxEnt. Two different appraoch: Foreground Removal vs Component Separation. Difficult for polarisation.



- * Step -3: Power Spectrum Estimation : optimal C_l s and associated error covariance Direct Inversion, Likelihood, Hybrid, Bayesian Sampling. Effective data compression Step.
- * Step -4A: Cosmological Parameter Estimation : Translating Cls to Cosmology and associated errorbars Grid, MCMC, various sampling techniques
- * Step -4B: Cosmological Parameter Estimation : Removing degeneracy in cosmo parameters using External data sets.
- * Step -5: Cosmological Parameter Estimation : Removing degeneracy by through use of higher order statistics

Aim is not to lose any Information atleast to keep the various steps as loseless as possible and to add information from External data sets and the higher order statistics to break degeneracies arising from CMB data alone i.e. -TO REDUCE THE ERROR-BARS-

Mission Specifications



- * Mass 1800 kg at launch
- * Dimensions 4.2 m high, 4.2 m maximum diameter
- * Launcher Ariane-5 from Guiana Space Centre
- * Launch Date 31st October 2008
- * Mission Lifetime 21 months nominal from end of commission phase
- * Wavelength Microwave: 25 GHz to 1 Thz (HFI : 83 GHz 1 THz); (LFI : 27 77 GHz)

Mission Goal



- \star Determine the Precise Primordial Fluctuation Spectrum
- * Test Inflation/Primodial Gravity Waves
- \star Statistics of the CMB Anisotropies
- ***** Small-scale Anisotropies and Reionization
- ***** Small-Scale Anisotropies and Galaxy Clustering
- * Sunyaev-Zel'dovich Effect in Clusters of Galaxies

Bit of Spin and All that Stuff

$$\Delta T(\hat{n}) = \sum_{lm} [a_{lm}] [Y_{lm}(\hat{n})]; \quad \hat{n} = (\theta, \phi)$$

$$\begin{aligned} Q(\hat{n}) - iU(\hat{n}) &= \sum_{lm} [a_{2,lm}] \quad [_2Y_{lm}(\hat{n})] \\ Q(\hat{n}) + iU(\hat{n}) &= \sum_{lm} [a_{-2,lm}] \quad [_{-2}Y_{lm}(\hat{n})] \end{aligned}$$

 $Y_{lm}(\hat{n})$ are usual spherical harmonics and $_2Y_{lm}(\hat{n})$ are their spinorial counterparts.

$$\pm P(\hat{n}) = Q(\hat{n}) \pm iU(\hat{n}) \tag{4}$$

$$\bar{\nabla}_{\pm 2} P(\hat{n}) = \sum_{lm} \left[\frac{(l+2)!}{(l-2)!} \right]^{1/2} [a_{\pm 2;lm}] \left[Y_{lm}(\hat{n}) \right]$$
(5)

$$E(\hat{n}) = \sum_{lm} a_{lm}^E Y_{lm}(\hat{n}) \tag{6}$$

$$B(\hat{n}) = \sum_{lm} a_{lm}^E Y_{lm}(\hat{n}) \tag{7}$$

$$a_{lm}^E = \frac{1}{2}(a_{2,lm} + a_{2,lm}) \tag{8}$$

$$a_{lm}^B = \frac{1}{2}(a_{2,lm} - a_{2,lm}) \tag{9}$$

Maps, Harmonics a_{lm} and Recovery of Power Spectra C_l

(3) Maps in Real Space and Harmonic Domain : Choice of Pixelisation Schemes

$$\Delta T(\theta_i, \phi_i) = \frac{1}{4\pi} \sum_{l}^{lmax} Y_{lm}(\theta, \phi) (a_{lm}^T + a_{lm}^N)$$
(10)

- \star Decomposition of Strokes parameters $Q\pm iU$ can be done interms of the spin harmonics ${}_{\pm 2}Y_{lm}.$
- * They are *spin-2* objects. In contrast scalers e.g. temperature anisotropy is a *spin-0* object.
- * Scalars or spinorial maps are constructed using a particular pixelisation schemes.
- * All pixelisation schemes are mathematically equivalent and employ different point sets to discritisation of the sphere.
- * But some can have distinct advantage on others as regards to computational speed and efficiency of data retivial.





- * Healpix uses hierachical organisation of pixels, Resolution parameter $N_{side} = 2^n$. Number of pixels $N_{pix} = 12Nside^2$. $l_{max} = 2 * N_{side}$. Nearly 12 million pixel to cover the surface of the sphere with 3.4 arcmin resolution (Planck resolution $\tilde{5}$ ') correspond to $N_{side} = 1024$, $l_{max} = 2048$
- * Highly optimised and portable library for various utilites exists including fast spherical transforms. http://healpix.jpl.nasa.gov/
- * Scalars or spinorial maps are constructed using a particular pixelisation schemes. http://www.mrao.cam.ac.uk/projects/cpac/igloo/, http://www.glesp.nbi.dk/
- * All pixelisation schemes are mathematically equivalent
- * But some can have distinct advantage on others as regards to computational speed and efficiency of data retivial.





"To err is human but to really f*** things up you need (super) computers "- Overheard

Correlation Function and Power Spectra

$$\xi_{ij} = \langle \Delta T(\theta_i, \phi_i) \Delta T(\theta_j, \phi_j) \rangle = \frac{1}{4\pi} \sum_{l} (2l+1) C_l b_l^2 P_l(Cos\theta_{ij}) + \sigma_0^2 \delta_{ij}$$
(11)

- * Most often the *sky coverage* is only partial.
- \star Incomplete sky coverage means ps estimates or C_l s are *correlated*. Binning is done to reduce the
- ***** errors and sometime it is necesarry to *deconvolve* the estimates.
- * Instrumental beam smoothes the maps. Typically its assumed to be a gaussian window function.
- * Pixelisation causes further smoothing. However beam smoothing is much coarser than *pixelisation* effects.
- * Pixel scale or the resolution of the maps limits to what l_{max} CMB spectrum can be recovered.
- ***** Often results are noise dominated much before this is achieved.
- * Various Estimates are available in extracting power spectra of CMB maps
- * MC based approaches can recover the input C_{ls} by inverting a "Transfer Function". Essentially taking a inverse problem approach. No clear technique to compute the associated "error covariance" matrix
- * Transfer function encodes beam, cut-sky, noise.
- * Likelihood based estimates recover input C_{ls} by maximising the likelihood of obtaining the map for a given set of C_{ls} .
- * Likelihood based approaches can provide Fisher Matrix based error estimates

Quadratic Maximum Likelihood Analysis

Degraded Lo Res Map



Likelihood :
$$\mathcal{L}(\mathcal{C}_{l}^{T}, \mathcal{C}_{l}^{E}, \mathcal{C}_{l}^{B}, \mathcal{C}_{l}^{x}|D) \propto \frac{1}{\det(\mathcal{C})^{1/2}} \exp(-\mathcal{D} \ \mathcal{C}^{-1} \ \mathcal{D}^{t})$$
 (12)

DataVector:
$$\mathcal{D} = (T(\theta_i), Q(\theta_j), U(\theta_k))$$
 (13)

$$\begin{array}{lll} \text{CovarianceMatrix}: & \mathcal{C} = \langle \mathcal{D}\mathcal{D}^{t} \rangle = \begin{pmatrix} \langle \mathbf{T}\mathbf{T}^{t} \rangle & \langle TQ^{t} \rangle & \langle TU^{t} \rangle \\ \langle \mathbf{T}\mathbf{Q}^{t} \rangle & \langle QQ^{t} \rangle & \langle QU^{t} \rangle \\ \langle \mathbf{T}\mathbf{U}^{t} \rangle & \langle UQ^{t} \rangle & \langle UU^{t} \rangle \end{pmatrix} \quad (14)$$

DerivativeMatrix:
$$\frac{dlogL}{d\mathcal{C}^{\S}_{\uparrow}} = \mathcal{D}^{T}(\mathcal{C}^{-1})^{T}(\partial_{l}^{x}\mathcal{C})(\mathcal{C}^{-1})\mathcal{D} - \operatorname{tr}(\mathcal{C}^{-1}\partial_{l}^{x}C)$$
(15)

FisherMatrix:
$$\left[\mathcal{F}_{ll'}^{xy}\right]^{-1} = \langle \delta C_l^x \delta C_{l'}^y \rangle; \quad x, y = T, E, B, X$$
 (16)

QMLEstimates:
$$\delta C_l^x = -\frac{1}{2} \left[\mathcal{F}_{ll'}^{xy} \right]^{-1} \frac{\partial f}{\partial C_{l'}^y}$$
 (17)

For simplified noise model $\langle n_i n_j \rangle = \delta_{ij} \sigma_{pix}^2$ and all sky coverage analytical results provide useful guidance.

$$\langle \delta C_l^2 \rangle = \frac{2}{(2l+1)} \left(C_l^2 + \Omega_{pix} \sigma_{pix}^2 \right)$$
(18)

- * One step Multidimensional Newton -Raphson root finder
- \star Effective Loseless Data Compression step $\mathcal{D} \rightarrow \mathcal{C}_l$
- ***** Brute Force technique can reach Nside = 32
- * Provides Clean Separation of Electric and Magnetic components from Q and U maps
- \star Inversion of C^{-1} can be done using Conjugate Gradient Technique
- * Trace calculations can be done using Monte-Carlo Calculations
- * Can be replaced by more efficient Exact likelihood sampling by Gibbs Sampling algorithms







- * High Resolution Noise Covariance Matrix (diagonal) when degraded generates off-diagoal terms.
- \star A large number of MC maps from Springtide/Madam are needed to achieve the level of required accuracy.

Pseudo C_l Analysis

- \star The measured convolved \tilde{C}_l can be espressed as a linear combination of input fiducial $C_l.$
- \star The estimator \hat{C}_l therefore is unbiased for realistic noise model and for arbitrary sky coverage

$$\tilde{C}_{l} = \frac{1}{2l+1} \sum_{m=-l}^{l} a_{lm} a_{lm}^{*}$$
(19)

$$\hat{C}_{l}^{T} = \left[M_{ll'}^{TT}\right]^{-1} \tilde{C}_{l}^{T}
\hat{C}_{l}^{e} + \hat{C}_{l}^{b} = \left[M_{ll'}^{e+b}\right]^{-1} (\tilde{C}_{l}^{e} + \tilde{C}_{l}^{b})
\hat{C}_{l}^{e} - \hat{C}_{l}^{b} = \left[M_{ll'}^{e+b}\right]^{-1} (\tilde{C}_{l}^{e} + \tilde{C}_{l}^{b})
\hat{C}_{l}^{X} = \left[M_{ll'}^{XX}\right]^{-1} \tilde{C}_{l}^{X}$$
(20)

- * Very fast but non-optimal
- ***** Analytical estimates for variance is only possible for the case of Temperature
- * Analytical estimates for variance for Polarisation is only possible for noise dominated regime
- \star Various weighting schemes can reduce the variance or the scatter associated with the estimator
- \star A weighting scheme can often reduce the variance associated with estimation
- ***** Equal Weights per pixels are near optimal in low I regime
- \star At high I the inverse variance per pixels is typically used







Mask





Hivon E. et. al., ApJ (567) 2, 2002 PolSpice, Xfaster, Xspect, Romaster, Madspec etc.



Comparison of Scatter in QML and PCL estimators

.



$$\Sigma_T(l_1, l_2, \tilde{W}_{l_3}) = \sum_{l_3} \frac{(2l_3 + 1)}{4\pi} \tilde{W}_{l_3} \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix}$$
(21)

 W_{l_3} is the window function and encodes information about sky-coverage, noise and depends on scanning strategy.

$$\langle \Delta \tilde{C}_{l}^{T} \Delta \tilde{C}_{l'}^{T} \rangle = 2C_{l}^{T} C_{l'}^{T} \Sigma_{T}(l, l', W_{l}^{(2)}) + \Sigma_{T}(l, l', W_{l}^{TT}) + 4(C_{l}^{T} C_{l'}^{T})^{1/2} \Sigma_{T}(l, l', W_{l_{3}}^{(2)T})$$
⁽²²⁾

- \star The window $W_l^{(2)}$ corresponds to finite sky coverage and is simply the power spectrum of the mask (apodised/unapodised). It plays a dominant role at small I
- \star The window W_l^{TT} corresponds to the noise distribution over the sky. Its dominant at high I
- \star The window W_l^{2T} corresponds to the cross between noise and partial sky coverage: Its always subdominant but not negligible.
- \star The deconvolution of the covariance matrix can be done using the matrices M.
- \star Similar results can be obtained for E, B, X
- \star At high -I can be roughly approximated by an f_{sky} approximation
- \star The approximations are more accurate for T than for E and B or X.



- \star The approximations effectively treats E and B as spin-0 field.
- \star More accurate treatment involves correction terms that depend gradients of spin-2 field and are computationally demanding.
- \star However as Planck is noise mostly dominated for E and B the results are very good approximations.
- \star The noise is assumed to be white and presence of large correlated component can lead to break down of analytical approximations.
- \star With effective destriping the results are in very good agreement with MC results.

$$\langle \delta \hat{C}_l \delta \hat{C}_{l'} \rangle = M_{ll''}^{-1} \langle \delta \tilde{C}_{l''} \delta \tilde{C}_{l'''} \rangle M_{l'''l}^{-1,t}$$
⁽²³⁾



Equal and Inverse variance Weights: Construction of Hybrid Estimator

Motivations for a hybrid Estimator C_l^{hy} :

- \star PCL estimators with Equal weight per-pixels Cl^{eq} are near optimum for Signal dominated regime (low I).
- \star The PCL estimators with Inverse weight C_l^{inv} per-pixel are near optimum at noise dominated regime (highl).

Computation of a hybrid Estimator

- \star The PCL estimates with Equal weight per-pixels C_l^{eq} are computed along with their covariances
- \star The PCL estimates with Inverse weight C_l^{inv} are computed next and the covariances.
- \star The cross-covariances are computed for equal and inverse weights
- \star Two estimates are combined to minimize the following $\chi^2.$

$$\chi^{2} = \sum_{l,l'} \sum_{\alpha,\beta} (\hat{C}_{l}^{\alpha} - \hat{C}_{l}^{hy}) F_{ll'}^{\alpha\beta} (\hat{C}_{l}^{\beta} - \hat{C}_{l}^{hy})$$
(24)



A Toy Model If we assume we only have mode in the sky l. The hybrid covariance of estimated hybrid C_l are given by:

$$\frac{1}{\sigma_{hy}^2} = \frac{1}{\sigma_{eq}^2} + \frac{1}{\sigma_{inv}^2} + \frac{2}{\sigma_{cross}^2}$$
(25)

The estimated C_l s are given by:

$$\frac{Cl^{hy}}{\sigma_{hy}^2} = \frac{C_l^{eq}}{\sigma_{eq}^2} + \frac{C_l^{inv}}{\sigma_{inv}^2} + \frac{C_l^{eq}}{\sigma_{cross}^2} + \frac{C_l^{inv}}{\sigma_{cross}^2}$$
(26)

The specific weighting scheme that is typically uses is:

$$w_i = \frac{1}{\frac{1}{N_i} + \frac{\epsilon}{\langle N \rangle}} \tag{27}$$

For $N_i >> \langle N \rangle$ i.e. for signal dominated regime we get $w_i = \langle N \rangle$ and for $N_i << \langle N \rangle$ we have $w_i = N_i$.

Efstathiou G., Mon.Not.Roy.Astron.Soc. 349 (2004) 603 Efstathiou G., Mon.Not.Roy.Astron.Soc. 370 (2006) 343



CTP - Planck Working Group Tests

- ***** Pase1a All Sky , No-noise Simulations.
- * Phase1b: Galactic cuts and Point source masks. Noise is diagonal and uncorrelated and in-pixel noise covariance matrix is very close to diagonal. MC simulations can be carried out by using Gaussian and Uncorrelated noise-approximations.
- * Phase2-symmetric beam : The noise is no-longer uncorrelated. In addition to the Gaussian noise there is a correlated component.
- * Phase2-assymetric beam: Beam Assymmetry is introduced. Phase2 maps are generated using realistic map-making algorithm GLS or Destriping techniques.
- * Phase-3: Full focal plane simulations which incorporates all detector assemblies of all channels so the errors from Component Separations are also taken in to account. Most realistic simulations till date.



Likelihood Results for Phase1b: Red is full-sky, Black is result from PCI simulation with Wishart distribution with pcls -covariances

see http://cosmologist.info/cosmomc/CMBLike.html for more details

Finally : Likelihood and Cosmology with CMB

Exact likelihood

$$lnP(\hat{C}|C) = \sum_{l} -\frac{(2l+1)}{2} \left(ln\left(\frac{C_l}{\hat{C}_l}\right) + \frac{C_l}{\hat{C}_l} \right)$$
(28)

$$lnP(\hat{C}|C) = \sum_{l} -\frac{(2l-1)}{2} ln\hat{C}_{l} - \frac{(2l-1)}{2} \left(ln\left(\frac{C_{l}}{\hat{C}_{l}}\right) + \frac{C_{l}}{\hat{C}_{l}} \right)$$
(29)

Approximate likelihood Usual suspects: Simplest Approximation is Gaussian! It is known to be biased (Bond Jaffe Knox)

Gaussian :
$$\ln P_{Gauss}(\hat{\mathbf{C}}|\mathbf{C}) \propto \exp\left[-\frac{1}{2}(\hat{C}-C)^T \mathbf{S}(\hat{C}-C)\right]$$
 (30)

Next Step is Lognormal! It is biased but in the opposite sense to the Gaussian Approx.

OffsetLognormal :
$$\ln P_{Offset-lognormal}(\hat{\mathbf{C}}|\mathbf{C}) \propto Exp\left[-\frac{1}{2}(\hat{z}-z)^T \mathbf{M}(\hat{z}-z)\right]$$
 (31)

A weighted sum of Gaussian and Lognormal likelihood used by WMAP team.

WMAP :
$$\ln P_{WMAP}(\hat{\mathbf{C}}|\mathbf{C}) = \frac{1}{3} ln P_{Gauss}(\hat{\mathbf{C}}|\mathbf{C}) + \frac{2}{3} ln P_{LN}(\hat{\mathbf{C}}|\mathbf{C})$$
 (32)

More complicated models exist including the Hybrid extension to Likelihood, Equal Variance Likelihood and Xfaster approximations to Likelihood. inc also include simpler models such as Inverse Gamma approximation.



Figure 2: Two Parameter Estimation Exercise from Phase2 simulations: Figure Courtesy: Eric Hivon, Planc Working Group (CTP).



Figure 3: Two Parameter Estimation Exercise from Phase2 simulations: Spice vs Hybrid



Cross - Correlating with External data sets : Life beyond PS estimation

Philosophy: Cross-Correalting two data sets can often provide insights which no individual data can.

Caution : Cross-Correlating World atlas with WMAP sky gives nonzero cross-correlation.

Cross correlation using the QML Estimator

Assuming statistical isotropy for both Φ^{α} and Φ^{β} data vectors, We can express the individual covariances and the cross-covariance in terms of the respective power spectra.

$$\boldsymbol{\Phi} = (\Phi^{\alpha}(\theta_i), \Phi^{\beta}(\theta_j)) \tag{33}$$

$$\mathbf{C} = \boldsymbol{\Phi}^{\mathrm{T}} \boldsymbol{\Phi} = \begin{pmatrix} \langle \Phi^{\alpha} \Phi^{\alpha} \rangle & \langle \Phi^{\alpha} \Phi^{\beta} \rangle \\ \langle \Phi^{\beta} \Phi^{\alpha} \rangle & \langle \Phi^{\beta} \Phi^{\beta} \rangle \end{pmatrix}$$
(34)



$$F_{\ell\ell'}^{\delta\delta'} = \langle \frac{\partial^2 \mathcal{L}}{\partial \mathcal{C}_{\ell}^{\alpha} \partial \mathcal{C}_{\ell'}^{\beta}} \rangle; \quad \delta, \delta' \in \alpha, \beta, \chi$$
(35)

The Fisher matrix denoted here by ${f F}$ is the expectation value of the curvature matrix.

Cross correlation using the PCL Estimator Similarly the covariance of C_l^x for cross- power-spectrum in pcl estimation can be expressed as:

$$\langle \delta \mathcal{C}_{\ell}{}^{x} \delta \mathcal{C}_{\ell'}{}^{x} \rangle = \sum_{L} \left\{ (\mathcal{C}_{\ell}{}^{a}) (\mathcal{C}_{\ell}{}^{b}) \frac{1}{2L+1} \sum_{M} |(w)_{LM}^{a}|^{2} |(w)_{LM}^{b}|^{2} + (\mathcal{C}_{\ell}{}^{a}) \frac{1}{2L+1} \sum_{M} |(w)_{LM}^{a}|^{2} |\sigma_{a}^{2}(w)_{LM}^{a}|^{2} \right\}$$
(36)

$$+(\mathcal{C}_{\ell}^{\ b})\frac{1}{2L+1}\sum_{M}|(w)_{LM}^{a}|^{2}|\sigma_{b}^{2}(w)_{LM}^{b}|^{2}+(\mathcal{C}_{\ell}^{\ x})\}\frac{1}{2L+1}\sum_{M}|\sigma_{a}^{2}(w)_{LM}^{a}|^{2}|\sigma_{b}^{2}(w)_{LM}^{b}|^{2}\left(\begin{array}{cc}L&\ell&\ell'\\0&0&0\end{array}\right)^{2},$$
(37)



Non-Gaussianities and other Non-Elephants

- ***** Work in Progress
- * Low Priority task at the Moment. Work is in progress to incorporate realistic noise and mask for already existing Estimators
- $\star~$ Non-Gaussian sky simulations are now complete
- \star Effect of Lensing and recovery of Lensed cls is being attempted

"Theory without Practice is Bullsh*t "-Unknown "Practice without theory is cow-dung "-Famous Politician



"In theory there is no difference between theory and practice But, in practice, there is." - Jan L. A. van de Snepscheut