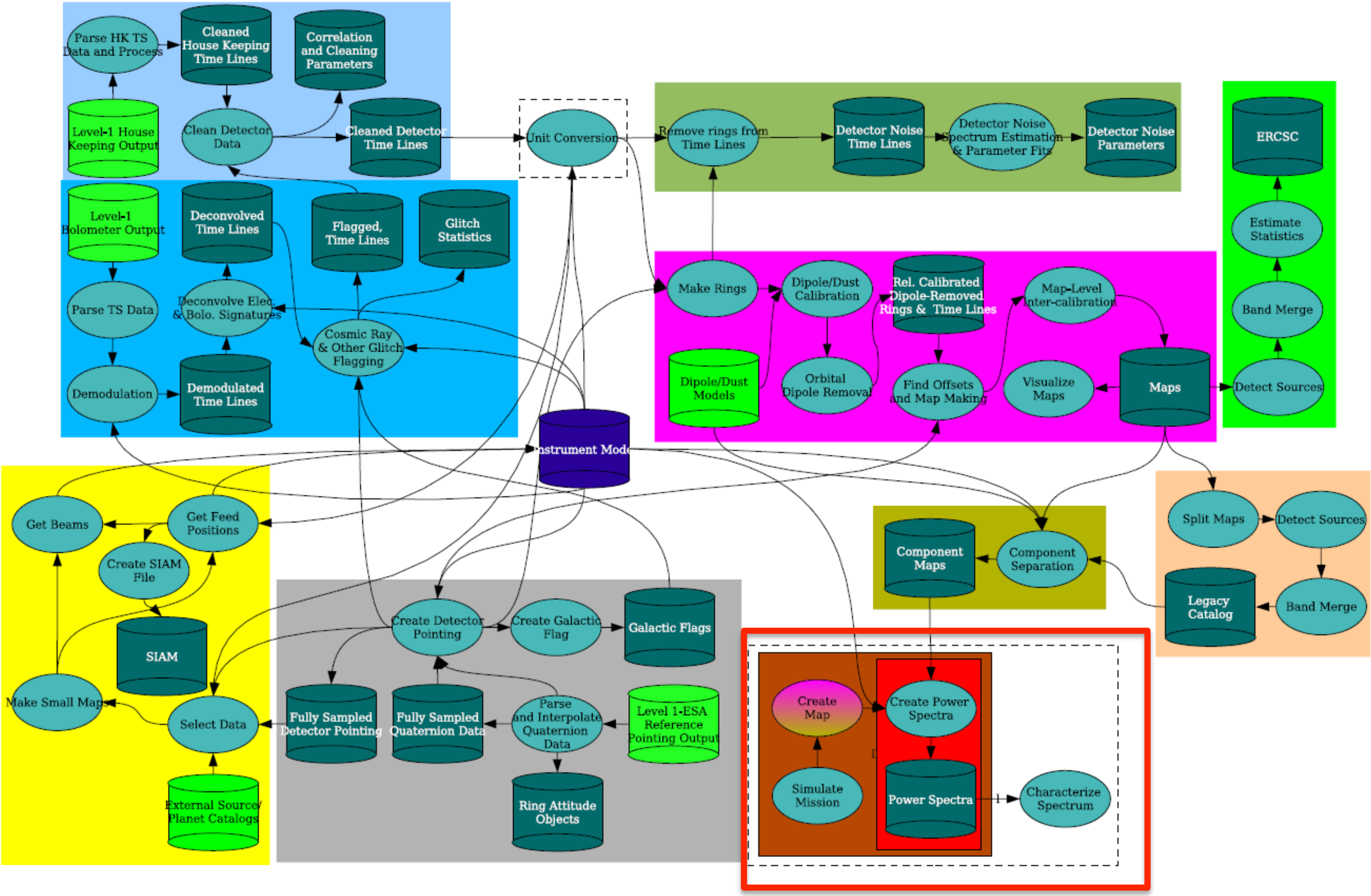
CMB DATA PROCESSING

- Spatial agency point of view:
  - provide a set of frequency channel maps
  - provide 'clean, calibrated' time-ordered data

- Cosmologist's point of view:
  - provide a likelihood, parameterized by power spectra
  - provide a clean, well-characterized CMB map (e.g for NG studies)
  - Don't want to hear about systematic effects !

- What we have to do:
  - Comply with spatial agency (always wise, they have the money)
  - Provide a likelihood to cosmologists
  - Provide 'some' CMB map (characterization ? Not necessarily intermediate step from above problem… see component separation)
  - Systematic effects:
    - Learn about the instrument parameters from calibration phases and flight data (iterative process…)
    - Validate the noise model (stationarity, null tests, etc.)
    - Propagate systematics residuals to likelihood somehow (hard, not always possible…)

A CMB pipeline sketch example: Planck HFI (FM)

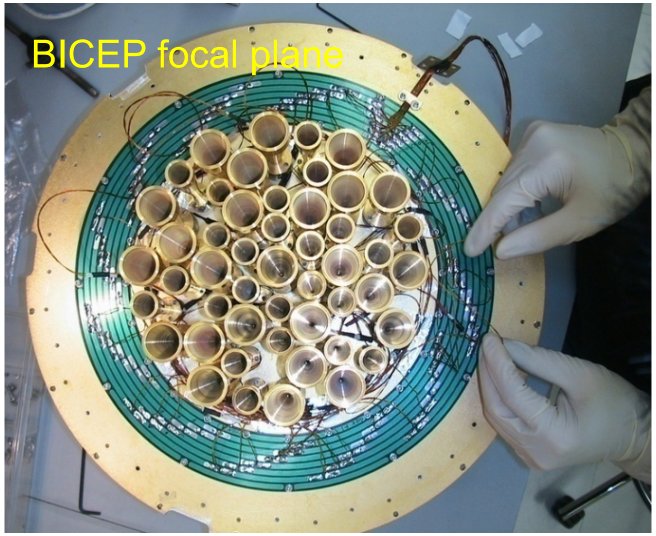


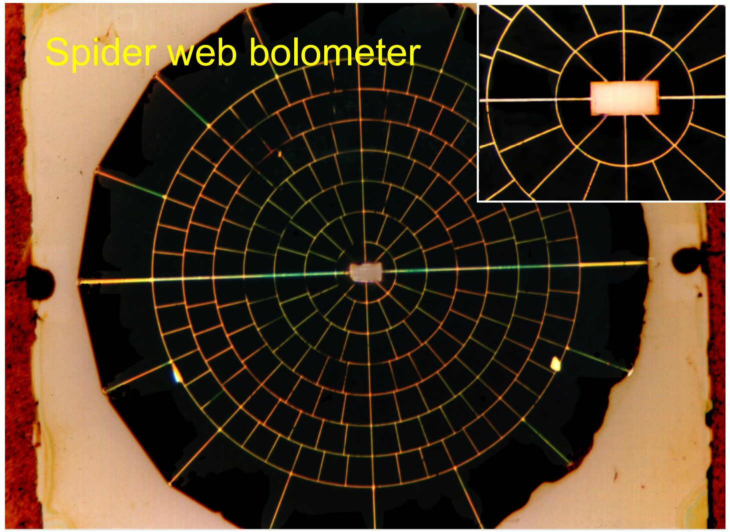
CMB data processing: plan of the lectures

- Cartography
  - Data model
  - From timelines to maps: optimal estimation
  - Polarized maps estimation
  - Noise models: the case of Planck (destripers)
  - Open questions: sub-pixel modelling, beam asymetry, non-stationarity
- Power spectrum estimation: towards a hybrid philosophy ?
  - Maximum likelihood (large scales)
  - Pseudo spectra (small scales, fast heuristic weighting)
  - Bayesian posterior samplers (Gibbs, HMC, PMC)
  - Bayesian posterior approximations: the Gaussian copula case
  - Open questions: towards a hybrid likelihood ?

SCALAR MAPMAKING

CMB imaging: scanning experiments





$$d_t = \int_{t-\tau}^{t} f(t') \int_{\nu-\Delta\nu}^{\nu+\Delta\nu} Z(\nu') \int B_\nu(\mathbf{n}(t').\mathbf{n}')T(\nu',\mathbf{n}')dt'd\nu'd^2\mathbf{n}' + \epsilon_t$$

Time-response
(detector + elec

and scanning strategy

Detector noise

Simplified linear model (pixelized sky)

$$d_t = A_{tp}T_p + \epsilon_t$$

Imagers: map-making

$$P(\theta|d,I) \propto P(d|\theta,I)P(\theta|I)$$

BAYES theorem

$$d_t = s_t + n_t = \sum_p A_{tp}T_p + n_t$$

Linear data model

$$\begin{aligned} P(T_p|N,d,I) &\propto P(d|T_p,N,I) \\ &\propto |2\pi C_N|^{-1/2} \times \exp\left[-\frac{1}{2}\sum_{pp'}(T_p-\bar{T}_p)C^{-1}_{N,pp'}(T_{p'}-\bar{T}_{p'})\right] \end{aligned}$$

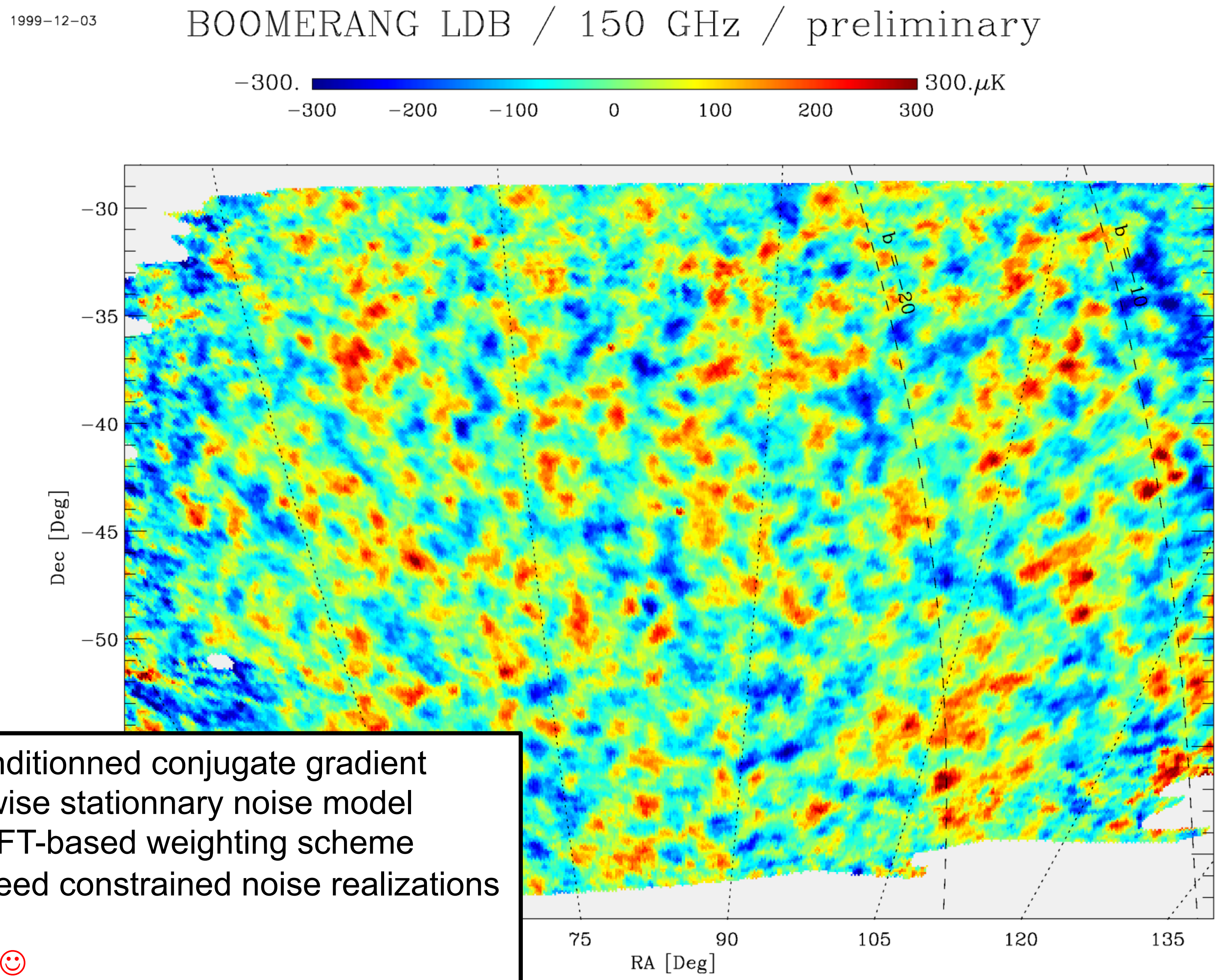
Uniform signal prior

$$\begin{aligned} \bar{T} &= (A^T N^{-1} A)^{-1} A^T N^{-1} d \\ C_N &= (A^T N^{-1} A)^{-1} \end{aligned}$$

Sufficient statistics
Covariance matrix of the map

Huge linear system to solve: use iterative methods (PCG) + FFTs

Exemple: Boomerang 1998



- Preconditionned conjugate gradient
- Piecewise stationnary noise model
  - FFT-based weighting scheme
  - need constrained noise realizations

- I did it ☺

POLARISED MAPMAKING

Imagers: polarised map-making

$$d_t^{(i)} = \frac{1}{2} P_{tp}^{(i)} \left[T_p + \cos(2\alpha_t^{(i)}) Q_p + \sin(2\alpha^{(i)}) U_p \right] + n_t^{(i)}$$

One polarised detector (i)

Let us consider n measurements of the same pixel, indexed by their angle α

$$\begin{aligned} s_\alpha &= \frac{1}{2}[I + Q\cos(2\alpha) + U\sin(2\alpha)] \\ \mathbf{d} &= \mathbf{AT} + \mathbf{n} \\ \mathbf{T} &= (I, Q, U)^T \end{aligned}$$

$$\mathbf{A} = \frac{1}{2} \begin{pmatrix} 1 & \cos 2\alpha_1 & \sin 2\alpha_1 \\ \vdots & \vdots & \vdots \\ 1 & \cos 2\alpha_p & \sin 2\alpha_p \\ \vdots & \vdots & \vdots \\ 1 & \cos 2\alpha_n & \sin 2\alpha_n \end{pmatrix}$$

$$\begin{aligned} \bar{\mathbf{T}} &= (\mathbf{A}^T \mathbf{N}^{-1} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{N}^{-1} \mathbf{d} \\ \langle \bar{\mathbf{T}} \bar{\mathbf{T}}^T \rangle &= \mathbf{V} = (\mathbf{A}^T \mathbf{N}^{-1} \mathbf{A})^{-1} \end{aligned}$$

ML solution

Polarisation: optimal configurations

$$\mathbf{V}^{-1} = \frac{1}{\sigma^2}\mathbf{X}, \qquad \mathbf{X} = \mathbf{A}^T\mathbf{A} = \frac{1}{4} \times$$

$$\begin{pmatrix} n & \sum_1^n \cos 2\alpha_p & \sum_1^n \sin 2\alpha_p \\ \sum_1^n \cos 2\alpha_p & \frac{1}{2}(n + \sum_1^n \cos 4\alpha_p) & \frac{1}{2}\sum_1^n \sin 4\alpha_p \\ \sum_1^n \sin 2\alpha_p & \frac{1}{2}\sum_1^n \sin 4\alpha_p & \frac{1}{2}(n - \sum_1^n \cos 4\alpha_p) \end{pmatrix}$$

General expression of the covariance matrix

Assume uncorrelated and equal variance measurements, look for optimal configuration of angles :

$$\mathbf{V}_0 = \sigma^2 \mathbf{X}_0{}^{-1}, \text{ with } \mathbf{X}_0{}^{-1} = \frac{4}{n}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

$$\alpha_p = \alpha_1 + (p-1)\frac{\pi}{n}, \; p = 1 \ldots n, \text{ with } n \geq 3$$

- Stokes parameters errors are uncorrelated
- Covariance determinant is minimized

Planck special case: destripers

- Specific observing strategy: ~45-60 redundant circles, then depointing
- Noise in phase-coadded data: mostly white + random offset
- Effect of N^{-1}: kill average of each ring
- Map-making reduces to chi-square fit of offset values using circles crossings

$$y = Pm + Fa + n.$$

F_j constant on ring j

⟹

Minimize w.r.t. a and m

$$\chi^2 = (y - Fa - Pm)^T C_n^{-1} (y - Fa - Pm).$$

$$m = (P^T C_n^{-1} P)^{-1} P^T C_n^{-1} (y - Fa).$$

White noise ⟹

Simple coadded map with offsets removed

$$F^T C_n^{-1} Z F a = F^T C_n^{-1} Z y.$$

$$Z = I - P(P^T C_n^{-1} P)^{-1} P^T C_n^{-1}.$$

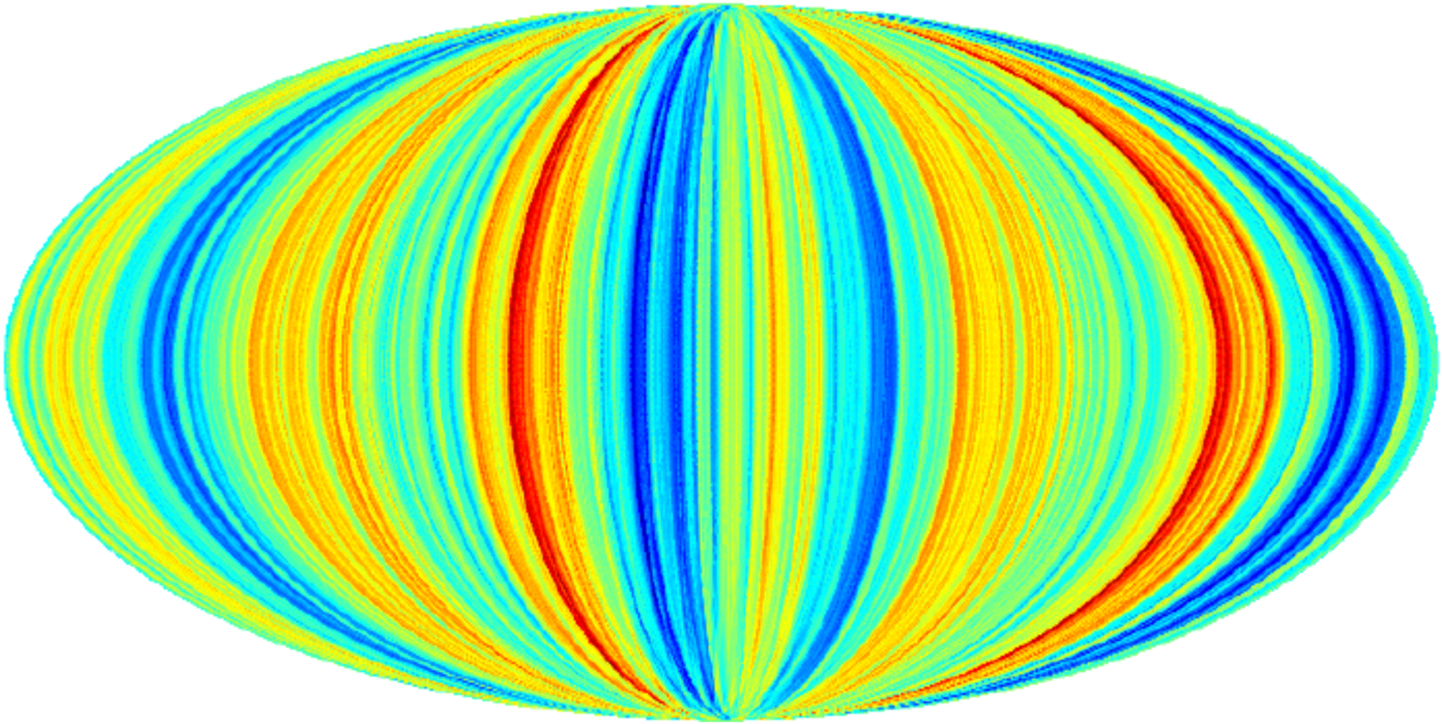
⟹

- Simple offset determination
- Z removes average value in pixel

Generalizes to more complex baselines F (Fourier modes, polynomials, etc.)

Delabrouille 98, Revenu et al. 00, Maino et al. 02, Keihanen et al. 04

Visual aspect of stripes



- Can dominate the signal at large scales if not taken care of !
- See Efstathiou 07 for power spectra of stripes or stripes residuals

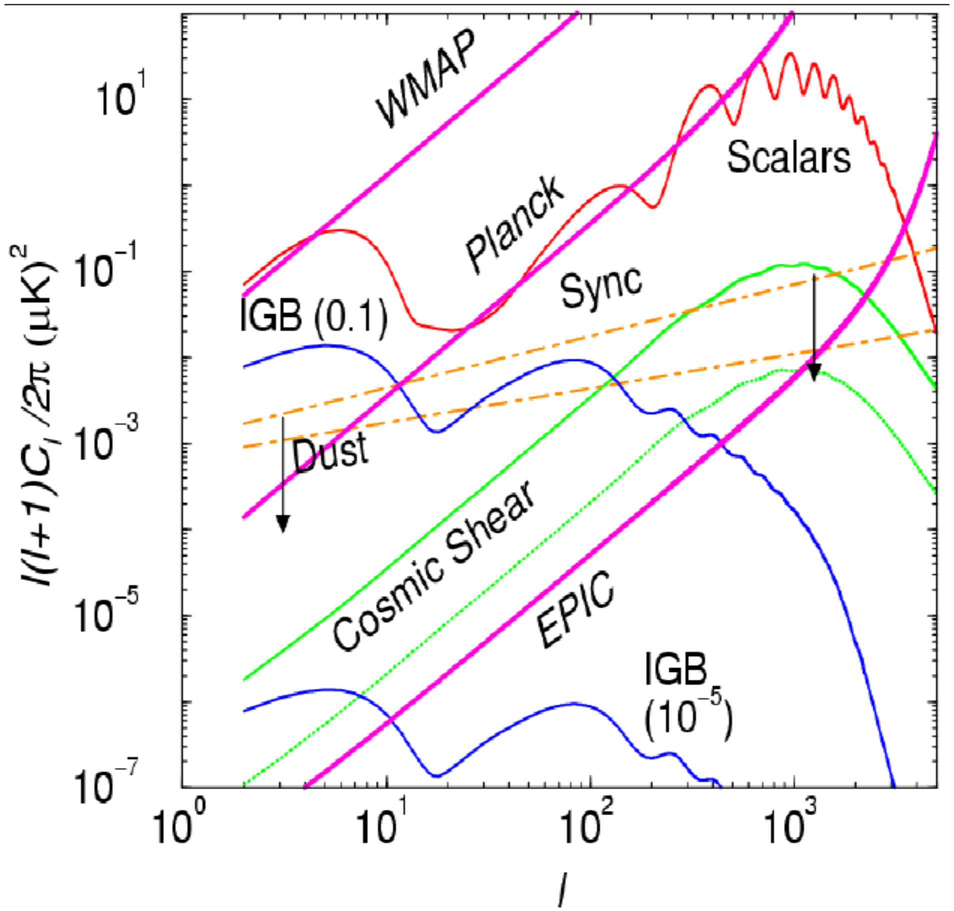
Maino et al. 08

What remains to be done for mapmaking

- Mapmaking methods described so far do not adress beam asymmetries
  - Case of circular beams: resulting beam is approximately homogeneous and circular on the map: easy to propagate to power spectra
  - Effect can in principle be propagated to power spectrum estimation (see work of T. Souradeep & S. Mitra)
  - Beam deconvolution: boosts small scale noise: needs regularization !
  - Far side lobe decomposition using all sky convolver: see Armitage & Wandelt 04 (still an open problem at small scales)
  - Sub-pixel modeling: signal variations inside pixels (important for Planck, especially polarization measurements)

Systematic deviations from the idealized case: impact on B modes

- Reminder: B-mode polarization signal is (expected to be) very tiny !
- All systematic effects can create artificial B-mode from E or T ☹



Detection chain systematics

Polarization tensor:

$$\mathbf{P} = C\langle \mathbf{E}\mathbf{E}^\dagger \rangle$$

$$= \Theta \mathbf{I} + Q\boldsymbol{\sigma}_3 + U\boldsymbol{\sigma}_1 + V\boldsymbol{\sigma}_2$$

Jones matrix: electric field transmission

$$\mathbf{E}_{\rm out} = \mathbf{J}\mathbf{E}_{\rm in}\,.$$

$$\mathbf{P}_{\rm out} = \mathbf{J}\mathbf{P}_{\rm in}\mathbf{J}^\dagger\,.$$

$$\hat{\mathbf{P}}_{\rm in} = \hat{\mathbf{J}}^{-1}\mathbf{P}_{\rm out}(\hat{\mathbf{J}}^\dagger)^{-1}$$

$$= (\hat{\mathbf{J}}^{-1}\mathbf{J})\mathbf{P}_{\rm in}(\hat{\mathbf{J}}^{-1}\mathbf{J})^\dagger$$

Error in Jones matrix determination:

$$\hat{\mathbf{J}}^{-1}\mathbf{J} = \begin{pmatrix} 1+g_1 & \epsilon_1 e^{i\phi_1} \\ \epsilon_2 e^{-i\phi_2} & (1+g_2)e^{i\alpha} \end{pmatrix}$$

- g_1, g_2: gains (fluctuations)
- ε_1, ε_2: cross-talk amplitudes
- ϕ_1, ϕ_2: cross-talk phases
- α: phase delay

$$\langle E_1 E_1^* \rangle - \langle E_2 E_2^* \rangle \rightarrow \delta Q = (g_1 + g_2)Q - (\epsilon_2 \cos\phi_2 - \epsilon_1 \cos\phi_1)U + (g_1 - g_2)\Theta$$

Hu, Hedman & Zaldarriaga 03

Gain error — Rotation error — Leakage error

Beam & pointing errors

Symmetry (spin) conditions imply local errors of the form:

$$\delta[Q \pm iU](\hat{\mathbf{n}};\sigma) = \sigma \mathbf{p}(\hat{\mathbf{n}}) \cdot \nabla [Q \pm iU](\hat{\mathbf{n}};\sigma) + \sigma [d_1 \pm i d_2](\hat{\mathbf{n}}) [\partial_1 \pm i \partial_2] \Theta(\hat{\mathbf{n}};\sigma) + \sigma^2 q(\hat{\mathbf{n}}) [\partial_1 \pm i \partial_2]^2 \Theta(\hat{\mathbf{n}};\sigma) + \cdots,$$

Beam errors, different for each polarized component:

$$B(\hat{\mathbf{n}};\mathbf{b},e) = \frac{1}{2\pi\sigma^2(1-e^2)} \exp\left[-\frac{1}{2\sigma^2} \left(\frac{(n_1-b_1)^2}{(1+e)^2} + \frac{(n_2-b_2)^2}{(1-e)^2} \right) \right],$$

Hu, Hedman & Zaldarriaga 03

Identifying error terms

Q measurement → $B(\hat{\mathbf{n}};\mathbf{b}_a, e_a) - B(\hat{\mathbf{n}};\mathbf{b}_b, e_b)$

Mean pointing error

$$\sigma\mathbf{p} = (\mathbf{b}_a + \mathbf{b}_b)/2,$$

Differential pointing error

$$\sigma\mathbf{b}_d = (\mathbf{b}_a - \mathbf{b}_b)/2,$$

Mean ellipticity error

$$e_s = (e_a + e_b)/2,$$

Differential ellipticity error

$$q = (e_a - e_b)/2,$$

At first order in the errors :

$$\hat{Q}(\hat{\mathbf{n}};\sigma) = \int d\hat{\mathbf{n}}' B(\hat{\mathbf{n}}') \left\{ Q(\hat{\mathbf{n}} + \hat{\mathbf{n}}' + \sigma\mathbf{p}) + \left[\left(\frac{\mathbf{b}_d \cdot \hat{\mathbf{n}}'}{\sigma} \right) + \frac{q}{\sigma^2}(n_2'^2 - n_1'^2) \right] \Theta(\hat{\mathbf{n}} + \hat{\mathbf{n}}') \right\},$$

$$\approx Q(\hat{\mathbf{n}};\sigma) + \sigma\mathbf{p} \cdot \nabla Q(\hat{\mathbf{n}};\sigma) + \sigma\mathbf{b}_d \cdot \nabla\Theta(\hat{\mathbf{n}};\sigma) + \sigma^2 q [\partial_1^2 - \partial_2^2] \Theta(\hat{\mathbf{n}};\sigma),$$

Hu, Hedman & Zaldarriaga 03

Impact on polarization B-modes

Measurement chain errors

Negligible beam size – varying coherence length

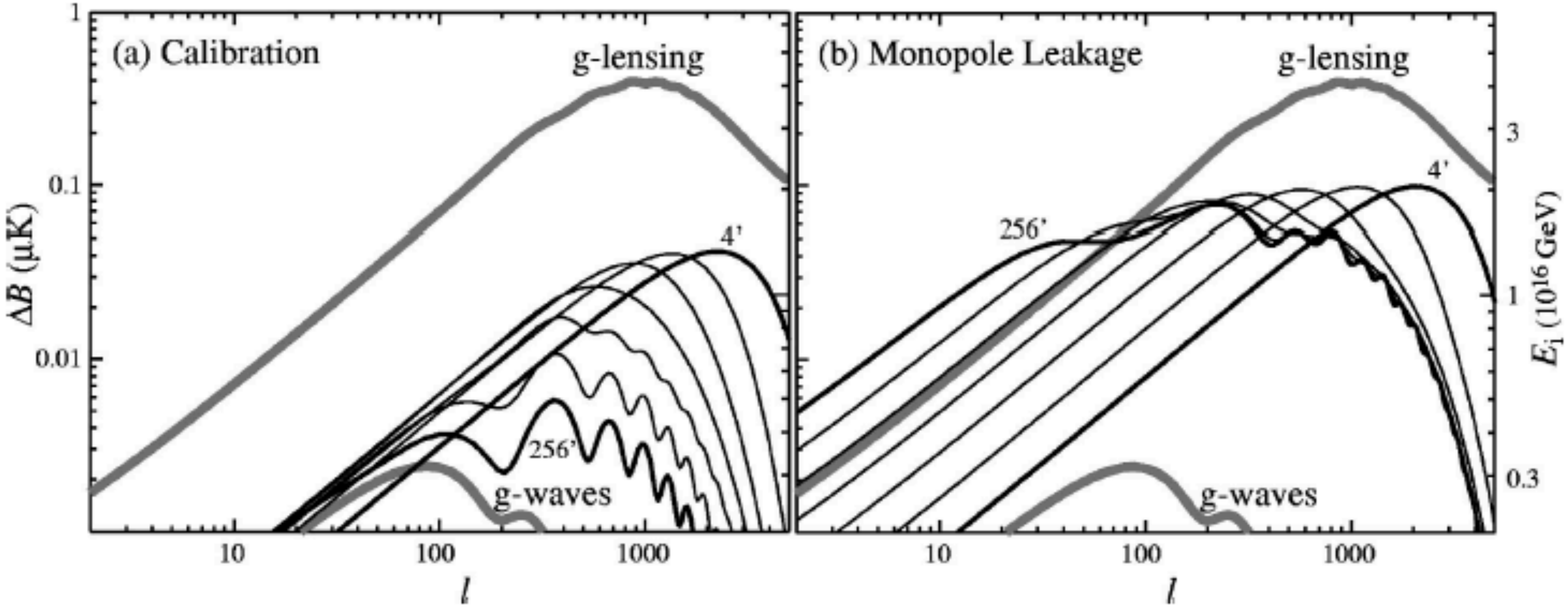


FIG. 3. Coherence dependence of B-mode contamination (a) for calibration a with rms $A_a = 10^{-2}$; (b) for monopole leakage γ_a, γ_b with $A_{\gamma_a} = A_{\gamma_b} = 10^{-3}$ added in quadrature. The beam scale is full width at half maximum (FWHM)$=(8 \ln 2)^{1/2}\sigma = 1'$ to remove beam effects and the FWHM coherence $(8 \ln 2)^{1/2}\alpha$ is stepped from $256'$ to $4'$ in factors of 2. Other effects follow the trend of calibration errors, not monopole leakage. For a coherence large compared with the CMB acoustic peaks, B contamination picks up their underlying structure. Here and in the following figures, the gravitational lensing and minimum detectable gravitational wave ($E_i = 3.2 \times 10^{15}$ GeV) B modes are shown for reference (thick shaded lines). The scaling with E_i of the peak in the B-mode spectrum is shown on the right hand axis.

Hu, Hedman & Zaldarriaga 03

Impact on B-modes (cont.)

Beam & pointing errors

Varying beam size (and coherence length accordingly)

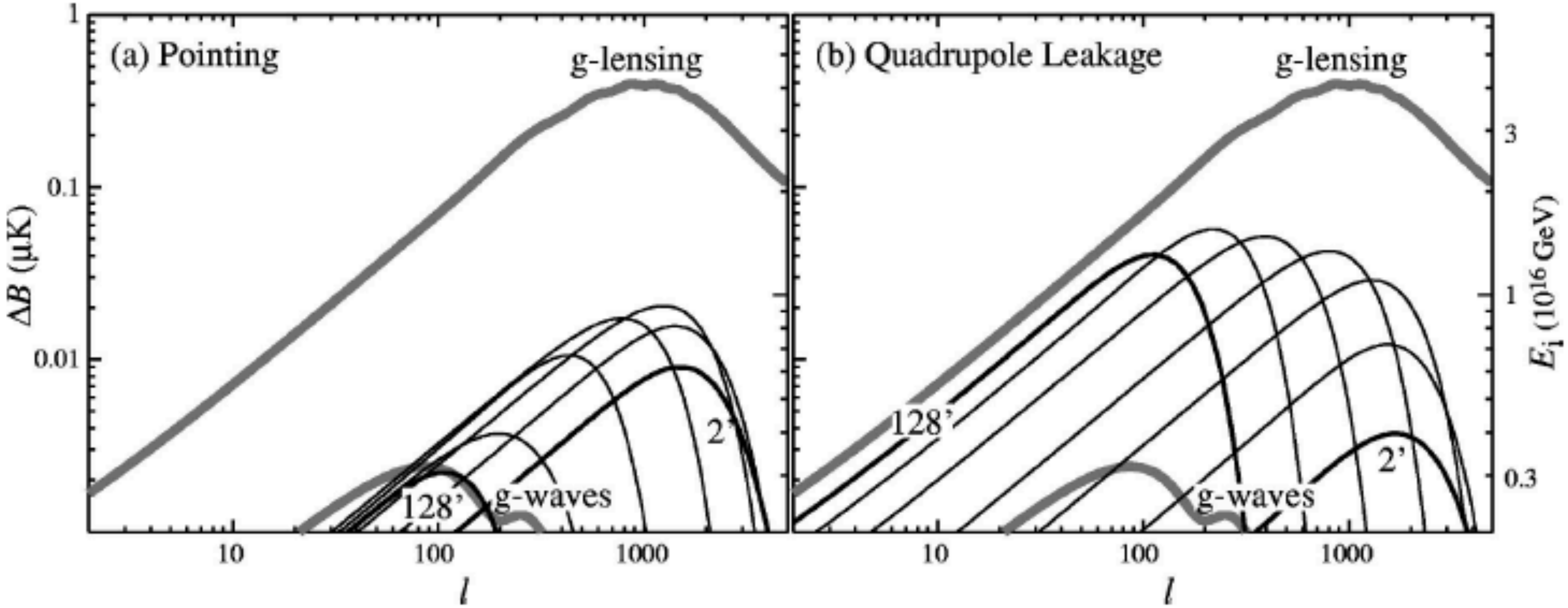


FIG. 4. Beam dependence of B-mode contamination for (a) pointing with a rms $A_{p_a}=A_{p_b}=10^{-2}$ (in units of the Gaussian beamwidth) added in quadrature; (b) quadrupole leakage with a rms $A_q=0.002$ (in units of differential beam ellipticity). The coherence α is set to $\max[\sigma, 10'/(8\ln 2)^{1/2}]$ and the beam is stepped from 128′ to 2′ in factors of 2.

Hu, Hedman & Zaldarriaga 03

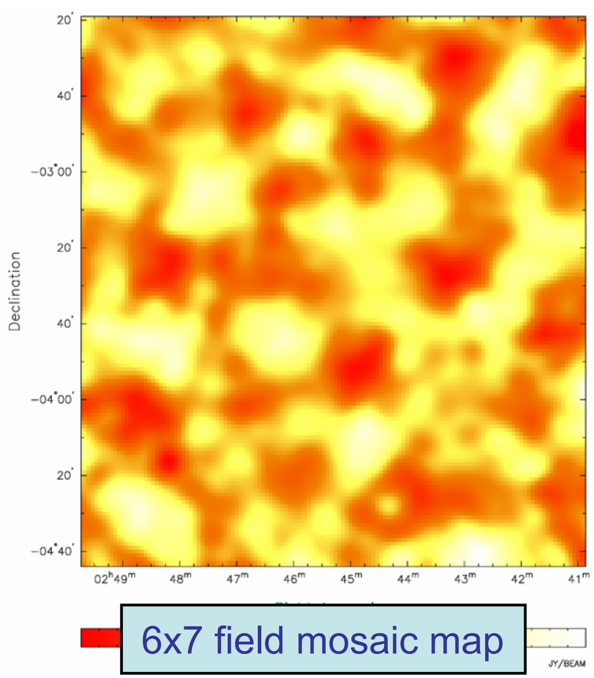
CMB data processing: Interferometers Application to CBI data

The case of interferometers

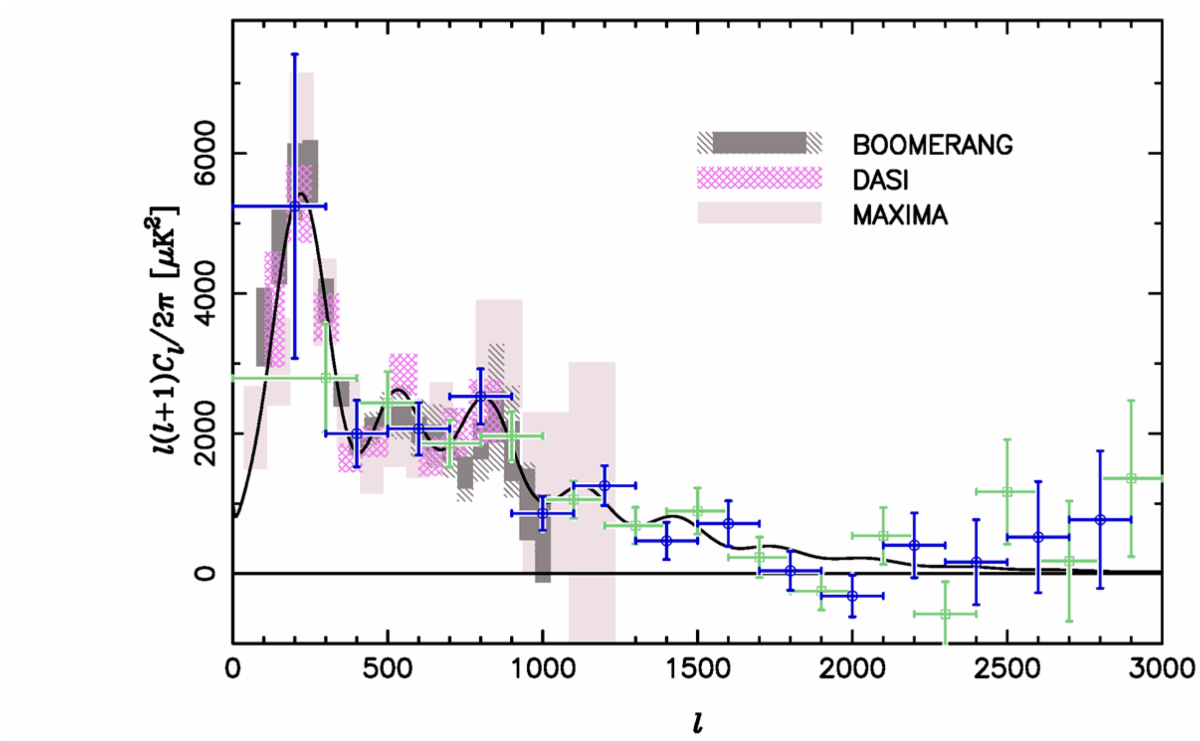
… and the diffusion scale

The CBI example:
- Atacama desert
- Interferometer (13 antennas)
- 10 frequency bands (26-36 GHz)
- Noise properties simpler (no drift scan)
- Ground effects, point sources …

(May 2002, as VSA)



6x7 field mosaic map



Also BIMA …

Interferometers: data model

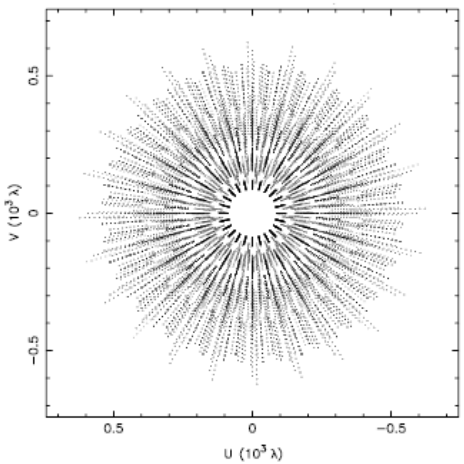
Visibilities: sample the convolved UV space:

$$V_{y_m}^T(\boldsymbol{u},\nu) = \frac{\partial B_\nu}{\partial T}\int d^2x\, A(\boldsymbol{x}-\boldsymbol{y}_m,\nu)\Delta T(\boldsymbol{x})e^{i2\pi\boldsymbol{u}\cdot\boldsymbol{x}}$$

$$V_{y_m}^T(\boldsymbol{u},\nu) = \frac{\partial B_\nu}{\partial T}\int d^2w\, \tilde{A}(\boldsymbol{w},\nu)e^{i2\pi\boldsymbol{w}\cdot\boldsymbol{y}_m}\Delta\tilde{T}(\boldsymbol{u}-\boldsymbol{w})$$

Idem for Q and U Stokes parameters

RL and LR baselines give (Q±iU)



UV coverage of a single pointing of CBI (10 freq. bands)
(Pearson et al. 2003)

Relationship between (Q,U) and (E,B) in UV (flat) space

$$\Delta T(\boldsymbol{x}) = \int d^2u\,\Delta\tilde{T}(\boldsymbol{u})e^{-i2\pi\boldsymbol{u}\cdot\boldsymbol{x}}\ ,$$
$$Q(\boldsymbol{x}) = \int d^2u\left[\tilde{E}(\boldsymbol{u})\cos 2\phi_{\boldsymbol{u}} - \tilde{B}(\boldsymbol{u})\sin 2\phi_{\boldsymbol{u}}\right]e^{-i2\pi\boldsymbol{u}\cdot\boldsymbol{x}}$$
$$U(\boldsymbol{x}) = \int d^2u\left[\tilde{E}(\boldsymbol{u})\sin 2\phi_{\boldsymbol{u}} + \tilde{B}(\boldsymbol{u})\cos 2\phi_{\boldsymbol{u}}\right]e^{-i2\pi\boldsymbol{u}\cdot\boldsymbol{x}}$$

Visibilities correlation matrix

$$M_{mn}^{ij} \equiv \left\langle V_{\boldsymbol{y}_m}^X(\boldsymbol{u}_i,\nu_i)V_{\boldsymbol{y}_n}^{Y*}(\boldsymbol{u}_j,\nu_j)\right\rangle = \frac{\partial B_{\nu_i}}{\partial T}\frac{\partial B_{\nu_j}}{\partial T}\int d^2w\,\tilde{A}_{\boldsymbol{y}_m}(\boldsymbol{u}_i-\boldsymbol{w},\nu_i) \times \tilde{A}_{\boldsymbol{y}_n}^*(\boldsymbol{u}_j-\boldsymbol{w},\nu_j)\mathscr{S}_{XY}(\boldsymbol{w})\ ,$$

$$\mathscr{S}_{TT}(\boldsymbol{w}) = \boldsymbol{S}_{TT}(w)$$
$$\mathscr{S}_{TQ}(\boldsymbol{w}) = \boldsymbol{S}_{TE}(w)\cos 2\phi_{\boldsymbol{w}}$$
$$\mathscr{S}_{TU}(\boldsymbol{w}) = \boldsymbol{S}_{TE}(w)\sin 2\phi_{\boldsymbol{w}}$$
$$\mathscr{S}_{QQ}(\boldsymbol{w}) = \boldsymbol{S}_{EE}(w)\cos^2 2\phi_{\boldsymbol{w}} + S_{BB}(w)\sin^2 2\phi_{\boldsymbol{w}}$$
$$\mathscr{S}_{QU}(\boldsymbol{w}) = \boldsymbol{S}_{EE}(w)\cos 2\phi_{\boldsymbol{w}}\sin 2\phi_{\boldsymbol{w}} - S_{BB}(w)\sin 2\phi_{\boldsymbol{w}}\cos 2\phi_{\boldsymbol{w}}$$
$$\mathscr{S}_{UU}(\boldsymbol{w}) = \boldsymbol{S}_{EE}(w)\sin^2 2\phi_{\boldsymbol{w}} + S_{BB}(w)\cos^2 2\phi_{\boldsymbol{w}}$$

Pixelisation in UV/pixel space

- Redundant measurements in UV-space
- Possibility to compress the data ~w/o loss

$$\boldsymbol{V}^{\rm tod} = \boldsymbol{A}\boldsymbol{V}^{\rm pix} + \boldsymbol{n}\,.$$

$$V^{\rm tod}_{\boldsymbol{y}(t_i)}(\boldsymbol{u}_k) = \sum_{p,l} A_{(ik)(pl)}\, V^{\rm pix}_{\boldsymbol{y}_p}(\boldsymbol{u}_l) + n_{\boldsymbol{y}(t_i)}(\boldsymbol{u}_k)$$

Least squares solution

$$\tilde{\boldsymbol{V}}^{\rm pix} = \boldsymbol{W}\boldsymbol{V}^{\rm tod}.$$

$$\boldsymbol{W} = \left(\boldsymbol{A}^T\boldsymbol{N}_t^{-1}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^T\boldsymbol{N}_t^{-1}$$

For an NGP pointing matrix:

$$\tilde{V}^{\rm pix}_{\boldsymbol{y}_p}(\boldsymbol{u}_l) = \frac{\sum_{i\in p,k\in l} V^{\rm tod}_{\boldsymbol{y}(t_i)}(\boldsymbol{u}_k)/\sigma^2_{ik}}{\sum_{i\in p,k\in l} 1/\sigma^2_{ik}}$$

Resultant noise matrix

$$(\boldsymbol{N}_p)_{(pl)(p'l')} = \left(\frac{1}{\sum_{i\in p,k\in l} 1/\sigma^2_{ik}}\right)\delta_{pp'}\delta_{ll'}$$

- Hobson and Maisinger 2002
- Myers et al. 2003
- Park et al. 2003

Use in conjonction with an ML estimator

$$\mathscr{L}(\{\mathscr{C}_b\}) = \frac{1}{(2\pi)^{N_p/2}|\boldsymbol{C}|^{1/2}}\exp\left(-\frac{\boldsymbol{V}^T\boldsymbol{C}^{-1}\boldsymbol{V}}{2}\right)$$

Newton-like iterative maximisation

$$\delta\mathscr{C}_b = \sum_{b'}(\boldsymbol{F}^{-1})_{bb'}\frac{\partial\ln\mathscr{L}}{\partial\mathscr{C}_{b'}} = \frac{1}{2}\sum_{b'}(\boldsymbol{F}^{-1})_{bb'}\,{\rm Tr}\left[(\boldsymbol{V}\boldsymbol{V}^T - \boldsymbol{C})\boldsymbol{C}^{-1}\frac{\partial\boldsymbol{S}}{\partial\mathscr{C}_{b'}}\boldsymbol{C}^{-1}\right]$$

Fisher matrix

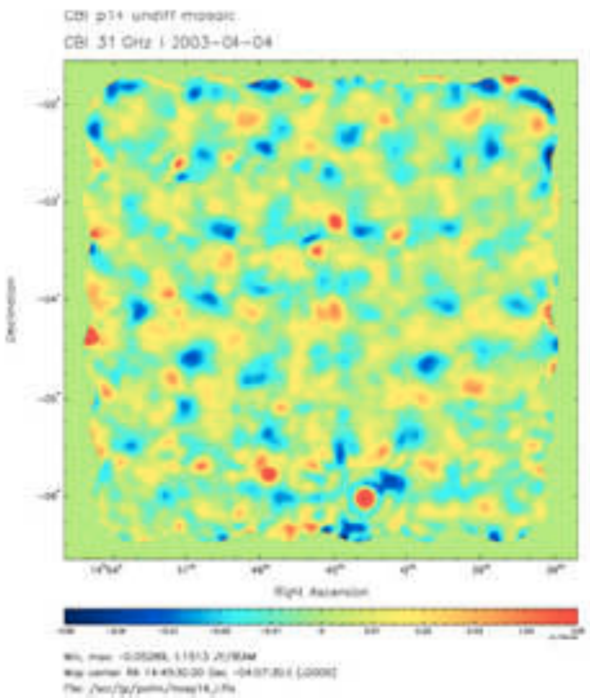
$$F_{bb'} \equiv -\left\langle\frac{\partial^2\ln\mathscr{L}}{\partial\mathscr{C}_b\partial\mathscr{C}_{b'}}\right\rangle = \frac{1}{2}{\rm Tr}\left(\boldsymbol{C}^{-1}\frac{\partial\boldsymbol{S}}{\partial\mathscr{C}_b}\boldsymbol{C}^{-1}\frac{\partial\boldsymbol{S}}{\partial\mathscr{C}_{b'}}\right)$$

Covariance derivatives for one visibility

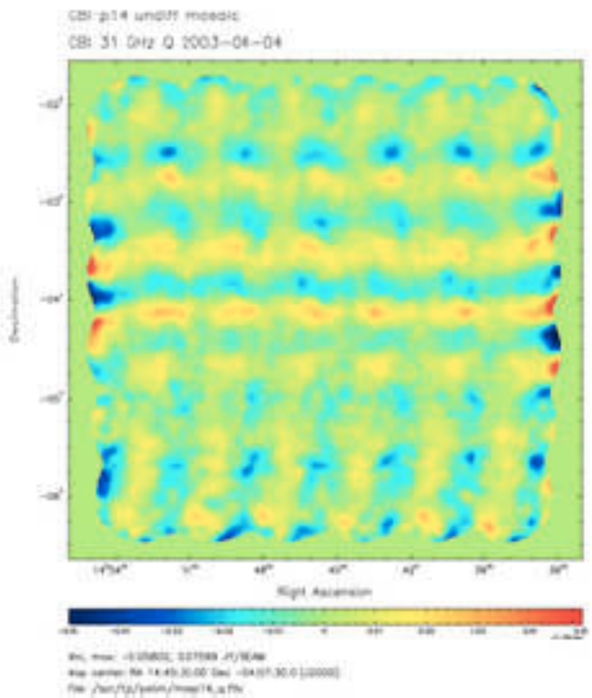
$$\frac{\partial M^{ij}_{mn}}{\partial\mathscr{C}^{TT}_b} = \frac{\partial B_{\nu_i}}{\partial T}\frac{\partial B_{\nu_j}}{\partial T}\int_0^{2\pi}\frac{d\theta_w}{2\pi}\int_{|\boldsymbol{u}_{b_1}|}^{|\boldsymbol{u}_{b_2}|}\frac{dw}{w} \times \tilde{A}_{\boldsymbol{y}_m}(\boldsymbol{u}_i - \boldsymbol{w},\nu_i)\tilde{A}^*_{\boldsymbol{y}_n}(\boldsymbol{u}_j - \boldsymbol{w},\nu_j)$$

Polarized mosaic observations: gd pick-up

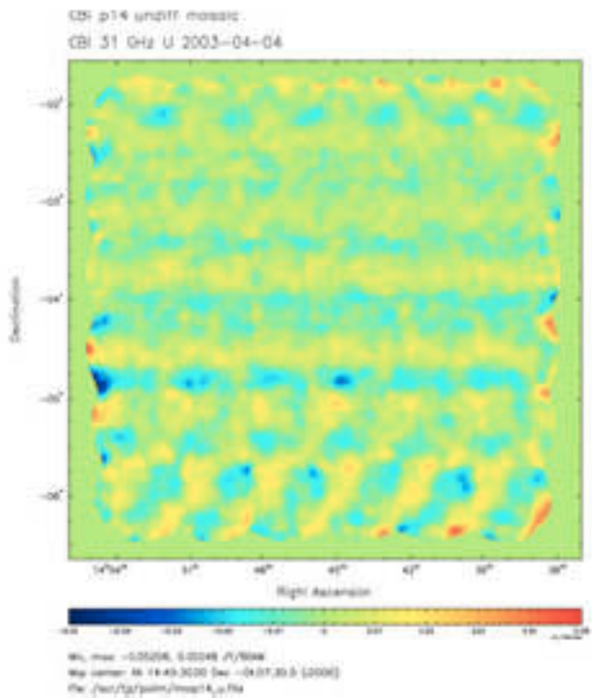
I



Q



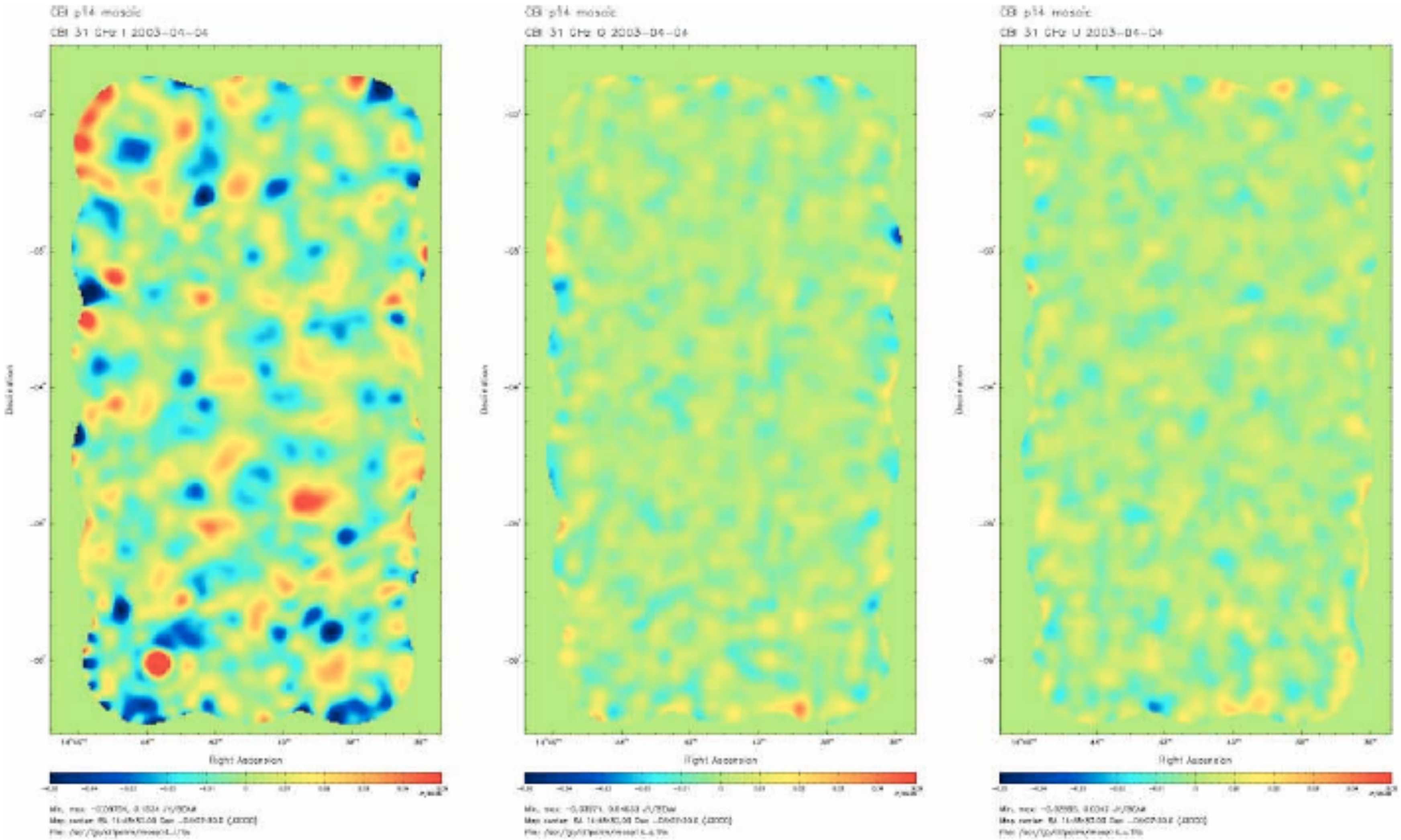
U



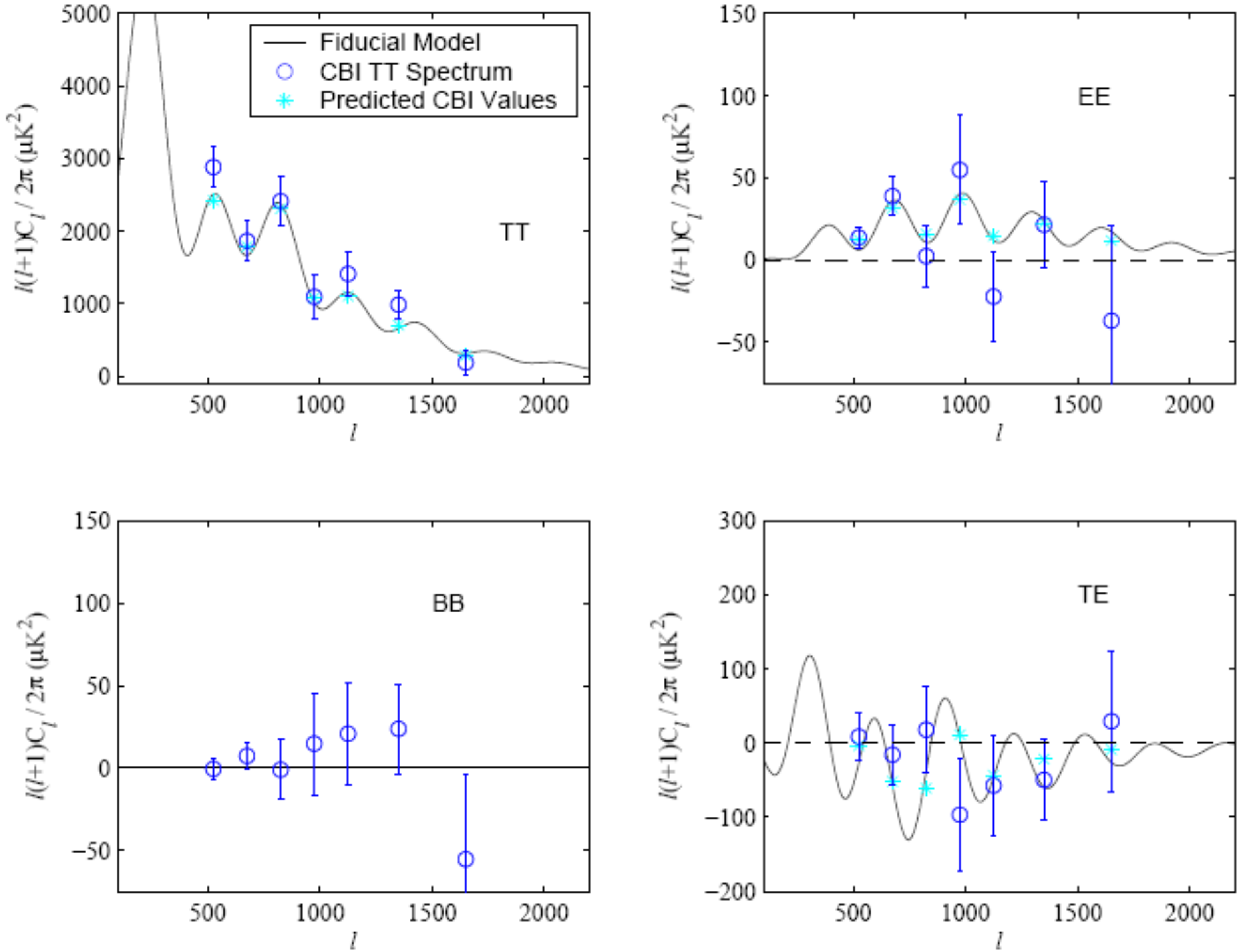
Polarization signal dominated by ground spillover: needs cleaning

Lead-trail differencing

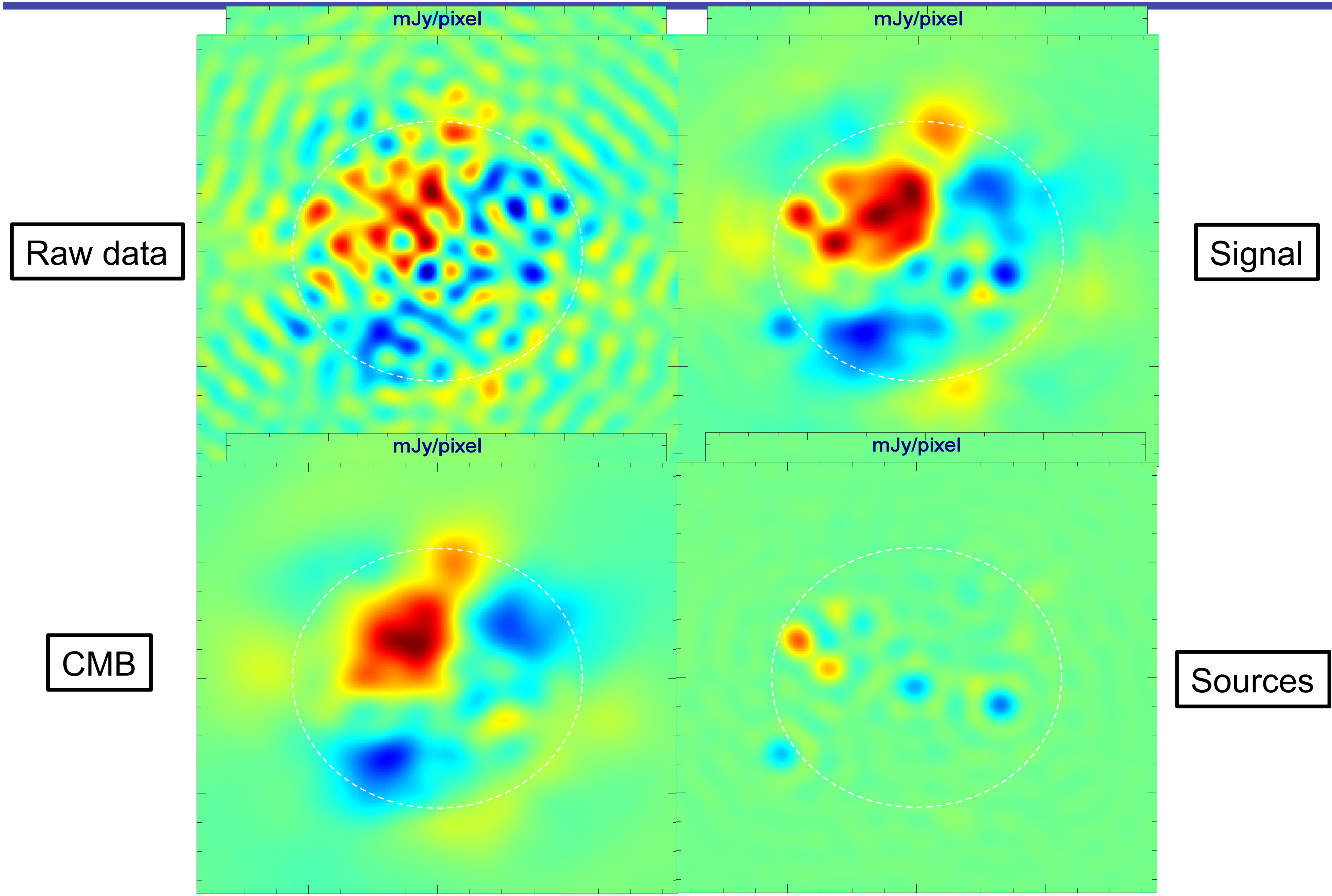
Differencing visibility measurements on fields separated by 9’ RA



Results of polarized mosaic observations



Linear (Wiener) filtering: application to imaging



SCALAR POWER SPECTRUM: MLE

Imagers: power spectrum

$$\langle T_p T_{p'} \rangle = C_{t,pp'} = \sum_\ell \frac{2\ell+1}{4\pi} C_\ell B_\ell^2 P_\ell(\hat{x}_p . \hat{x}_{p'})$$

Signal covariance matrix

$$P(C_\ell, T_p | d, I) \propto P(T_p | C_\ell, I) P(d | T_p, I)$$

BAYES again…

$$\begin{aligned} P(C_\ell | d, I) &= \int dT_p P(C_\ell, T_p | d, I) \\ &= P(C_\ell | I) \int dT_p P(T_p | C_\ell, I) P(d | T_p, I) \\ &= P(C_\ell | I) P(\bar{T}(d) | C_\ell, I) \\ &= P(C_\ell, I) \int dT_p |2\pi C_N|^{-1/2} |2\pi C_T|^{-1/2} \\ &\quad \times \exp\left[-\frac{1}{2} \sum_{pp'} (T_p C_{T,pp'}^{-1} T_p' + (T_p - \bar{T}_p) C_{N,pp'}^{-1} (T_{p'} - \bar{T}_{p'}) \right] \end{aligned}$$

Marginalize over the map

$$P(\bar{T}_p | C_\ell, I) = |2\pi (C_T + C_N)_{pp'}|^{-1/2} \exp\left[-\frac{1}{2} \sum_{pp'} \bar{T}_p (C_T + C_N)_{pp'}^{-1} \bar{T}_{p'} \right]$$

TO BE MAXIMIZED WITH RESPECT TO POWER SPECTRUM

Imagers: power spectrum (cont.)

$$\ln P(C+\delta C) = \ln P(C) + \sum_{\ell} \frac{\partial \ln P}{\partial C_\ell}\delta C_\ell + \frac{1}{2}\sum_{\ell,\ell'} \frac{\partial^2 \ln P}{\partial C_\ell \partial C'_\ell}\delta C_\ell \delta C_{\ell'} + \dots$$

Second order Taylor expansion

$$\partial C_\ell = -\sum_{\ell'} \left\langle \frac{\partial^2 \ln P}{\partial C_\ell \partial C_{\ell'}} \right\rangle^{-1} \frac{\partial \ln P}{\partial C_{\ell'}}$$

PSEUDO-NEWTON (FISHER)

$$\frac{\partial \ln P}{\partial C_\ell} = \frac{1}{2}\mathrm{Tr}\left[(\bar{T}\bar{T}^T - C)((C_T + C_N)^{-1}\frac{\partial C_T}{\partial C_\ell}(C_T + C_N)^{-1})\right]$$

$$\left\langle \frac{\partial \ln P}{\partial C_\ell \partial C_{\ell'}} \right\rangle = \frac{1}{2}\mathrm{Tr}((C_T+C_N)^{-1}\frac{\partial C_T}{\partial C_\ell}(C_T+C_N)^{-1}\frac{\partial C_T}{\partial C_{\ell'}})$$

For each iteration and each band, N_{pix}^3 operation scaling !!

‘Quadratic Maximum Likelihood’

- If some knowledge of the power spectrum is available, one may side-step iterations of the pseudo-Newton to compute an optimal estimator (i.e. lossless)
- The estimator is then quadratic in the data…
- Note that in principle you need the answer to get the answer !!
- Suppose noisy experiment, with no beam (easy to generalize though)

$$\widehat{C}_\ell = \mathbf{x}^t \mathbf{E}^\ell \mathbf{x} - b_\ell$$

$$\mathbf{C} \equiv \langle \mathbf{x}\mathbf{x}^t \rangle = \mathbf{N} + \sum_\ell \mathbf{P}^\ell \mathrm{C}_\ell,$$

Unbiased

$$b_\ell = \mathrm{tr}\, \mathbf{N}\mathbf{E}^\ell$$

$$\langle \widehat{C}_\ell \rangle = \sum_{\ell'} \mathbf{W}_{\ell\ell'} \mathrm{C}_{\ell'},$$

$$\mathbf{W}_{\ell\ell'} \equiv \mathrm{tr}\, \mathbf{P}^{\ell'} \mathbf{E}^\ell.$$

$$\mathbf{W}_{\ell\ell} = 1$$

Minimum variance, under constraint

$$\mathbf{V}_{\ell\ell'} = 2\mathrm{tr}\, [\mathbf{C}\mathbf{E}^\ell \mathbf{C}\mathbf{E}^{\ell'}]$$

$$L = tr[\mathbf{C}\mathbf{E}^\ell \mathbf{C}\mathbf{E}^{\ell'} - \lambda_\ell (\mathbf{P}^\ell \mathbf{E}^\ell \delta_{\ell\ell'} - \delta_{\ell\ell'})]$$

$$\mathbf{E}^\ell = \frac{1}{2} \mathbf{F}^{-1}_{\ell\ell'} tr[\mathbf{C}^{-1} \mathbf{P}^{\ell'} \mathbf{C}^{-1}]$$

$$\mathbf{F}_{\ell\ell'} = tr[\mathbf{C}^{-1} \mathbf{P}^\ell \mathbf{C}^{-1} \mathbf{P}^{\ell'}]$$

- Knowing C means knowing power spectrum !
- Still a very time-consuming process…
- Fisher matrix for these parameters not always invertible
- Generalizes easily to polarization data

Tegmark 97

Application to COBE data

State of the art in 1997….
Do you see a peak ? ☺

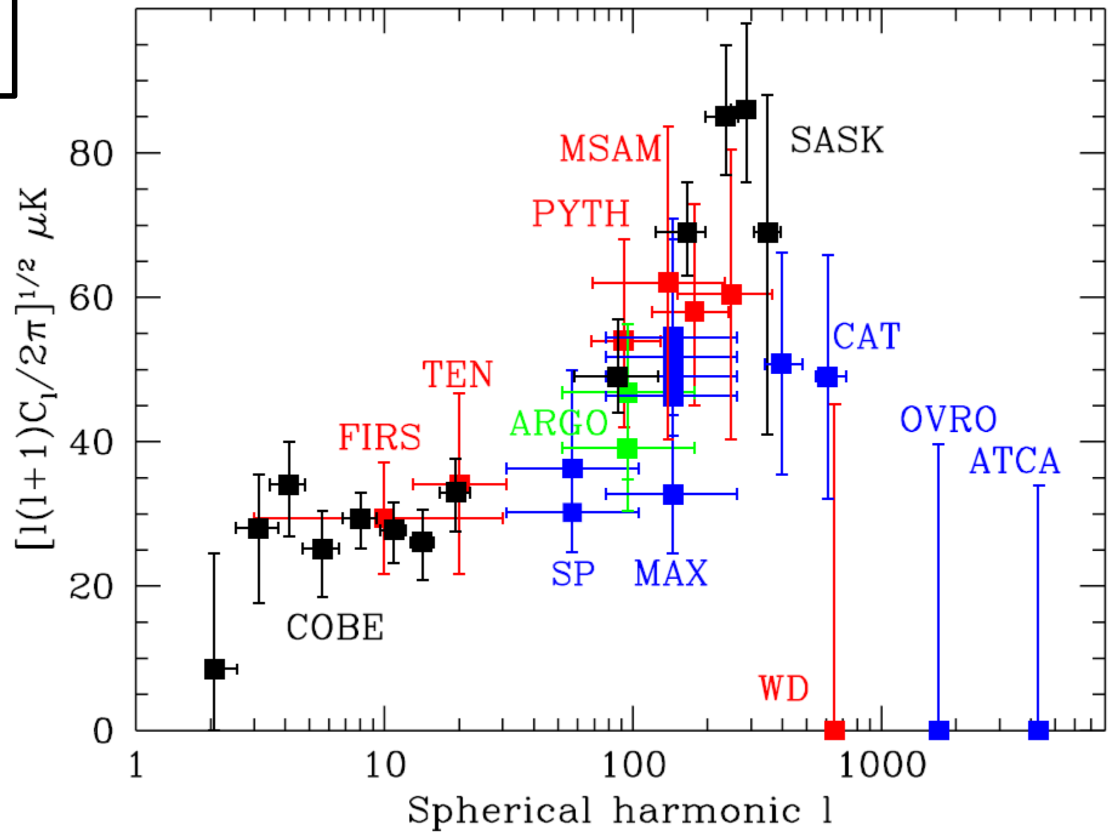


FIG. 8. The power spectrum observed by COBE/DMR binned into 8 bands and compared with other experiments.

Tegmark 97

SCALAR MASTER

Imagers: too many pixels !

➔ New (fast) analysis methods needed

$$\tilde{a}_{lm} = \int d\boldsymbol{u} \Delta T(\boldsymbol{u}) W(\boldsymbol{u}) Y^{*}_{lm}(\boldsymbol{u})$$
$$= \sum_{l'm'} a_{l'm'} \int d\boldsymbol{u} Y_{l'm'}(\boldsymbol{u}) W(\boldsymbol{u}) Y^{*}_{lm}(\boldsymbol{u})$$
$$= \sum_{l'm'} a_{l'm'} K_{lml'm'}[W] ,$$

- Fast harmonic transforms
- Heuristically weighted maps

$$K_{l_1 m_1 l_2 m_2} \equiv \int d\boldsymbol{u} Y_{l_1 m_1}(\boldsymbol{u}) W(\boldsymbol{u}) Y_{l_2 * m_2}(\boldsymbol{u})$$
$$= \sum_{l_3 m_3} w_{l_3 m_3} \int d\boldsymbol{u} Y_{l_1 m_1}(\boldsymbol{u}) Y_{l_3 m_3}(\boldsymbol{u}) Y_{l_2 * m_2}(\boldsymbol{u})$$
$$= \sum_{l_3 m_3} w_{l_3 m_3} (-1)^{m_2} \left[\frac{(2l_1 + 1)(2l_2 + 1)(2l_3 + 1)}{4\pi} \right]^{1/2}$$
$$\times \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & -m_2 & m_3 \end{pmatrix} ,$$

Quite ugly at first sight !!

Imagers (cont.)

Power spectrum expectation value

$$\langle \tilde{C}_{l_1} \rangle \equiv \frac{1}{2l_1+1} \sum_{m_1=-l_1}^{l_1} \langle \tilde{a}_{l_1\, m_1} \tilde{a}_{l_1 *\, m_1} \rangle ,$$

$$= \frac{1}{2l_1+1} \sum_{m_1=-l_1}^{l_1} \sum_{l_2\, m_2} \sum_{l_3\, m_3} \langle a_{l_2\, m_2} a_{l_3 *\, m_3} \rangle K_{l_1\, m_1\, l_2\, m_2}[W] K_{l_1 *\, m_1\, l_3\, m_3}[W]$$

$$= \frac{1}{2l_1+1} \sum_{m_1=-l_1}^{l_1} \sum_{l_2} \langle C_{l_2} \rangle \sum_{m_2=-l_2}^{l_2} | K_{l_1\, m_1\, l_2\, m_2}[W] |^2 .$$

…simplifies, after summation over angles (m):

$$\langle \tilde{C}_{l_1} \rangle = \sum_{l_2} M_{l_1\, l_2} \langle C_{l_2} \rangle ,$$

$$M_{l_1\, l_2} = \frac{2l_2+1}{4\pi} \sum_{l_3} (2l_3+1) \mathcal{W}_{l_3} \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix}^2 .$$

Imagers: "Master" method

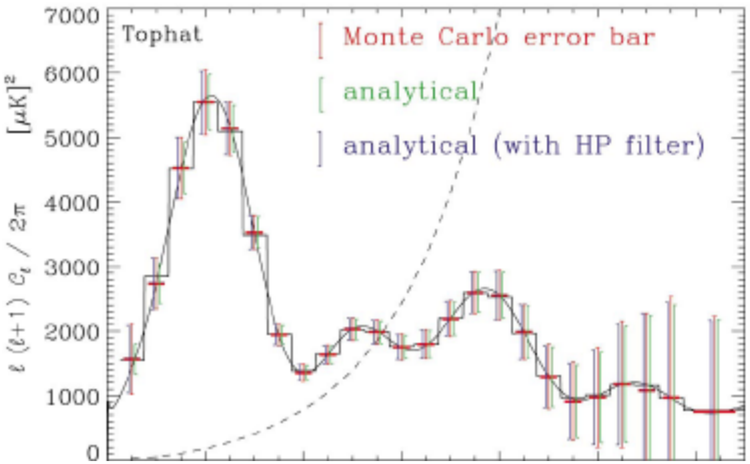
Finite sky coverage ➔ loss of spectral resolution ➔ need to regularize inversion

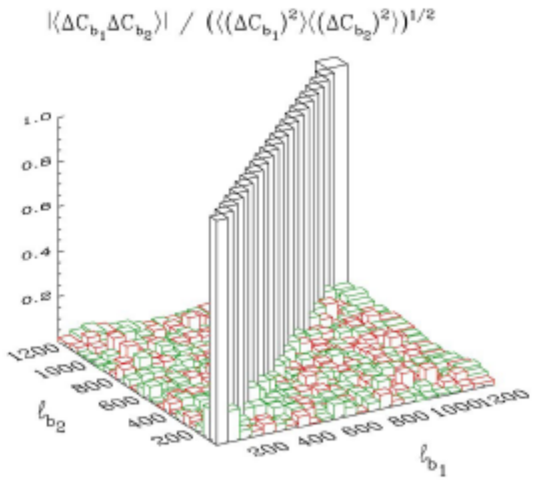
Spectral binning of the kernel

$$K_{bb'} = P_{bl} K_{ll'} Q_{l'b'} \ ,$$
$$= P_{bl} M_{ll'} F_{l'} B_{l'}^2 Q_{l'b'} \ .$$

Unbiased estimator

$$\hat{\mathscr{C}}_b = K_{bb'}^{-1} P_{b'l} (\tilde{C}_l - \langle \tilde{N}_l \rangle_{\mathrm{MC}})$$





MC estimation of covariance matrix of PS estimates

Works also for polarization (easier regularization on correlation function)

Quadratic estimators: covariances

Temperature case, reminders

$$\tilde{a}_{lm} = \int d\boldsymbol{u} \Delta T(\boldsymbol{u}) W(\boldsymbol{u}) Y^*_{lm}(\boldsymbol{u})$$
$$= \sum_{l'm'} a_{l'm'} \int d\boldsymbol{u} Y_{l'm'}(\boldsymbol{u}) W(\boldsymbol{u}) Y^*_{lm}(\boldsymbol{u})$$
$$= \sum_{l'm'} a_{l'm'} K_{lml'm'}[W] ,$$

$$\langle \tilde{C}_{l_1} \rangle = \sum_{l_2} M_{l_1 l_2} \langle C_{l_2} \rangle ,$$

$$M_{l_1 l_2} = \frac{2l_2 + 1}{4\pi} \sum_{l_3} (2l_3 + 1) \mathscr{W}_{l_3} \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix}^2 .$$

Edge-corrected estimators covariances in terms of pseudo-Cls covariances

$$\widehat{C}_\ell = M^{-1}_{\ell\ell'} \tilde{C}_\ell$$
$$\mathrm{Cov}(\widehat{C}_\ell, \widehat{C}_{\ell'}) = M^{-1}_{\ell\ell_1} \mathrm{Cov}(\tilde{C}_{\ell_1}, \tilde{C}_{\ell_2}) M^{-T}_{\ell_2\ell'}$$

As long as $M_{ll'}$ is invertible, same information content in edge-corrected Cls and pseudo-Cls

Pseudo-Cls estimators: cosmic variance

Forget noise for the moment, consider signal only:

$$\left\langle \Delta\tilde{C}^p_\ell \Delta\tilde{C}^p_{\ell'} \right\rangle = \frac{2}{(2\ell+1)(2\ell'+1)} \sum_{mm'} \sum_{\ell_1 m_1} \sum_{\ell_2 m_2} C_{\ell_1} C_{\ell_2} K_{\ell m \ell_1 m_1} K^*_{\ell' m' \ell_1 m_1} K^*_{\ell m \ell_2 m_2} K_{\ell' m' \ell_2 m_2}.$$

Case of high ells and/or almost full sky

$$\left\langle \Delta\tilde{C}^p_\ell \Delta\tilde{C}^p_{\ell'} \right\rangle = \tilde{V}_{\ell\ell'} \approx 2 C_\ell C_{\ell'} \Xi(\ell, \ell', \tilde{W}^{(2)}),$$

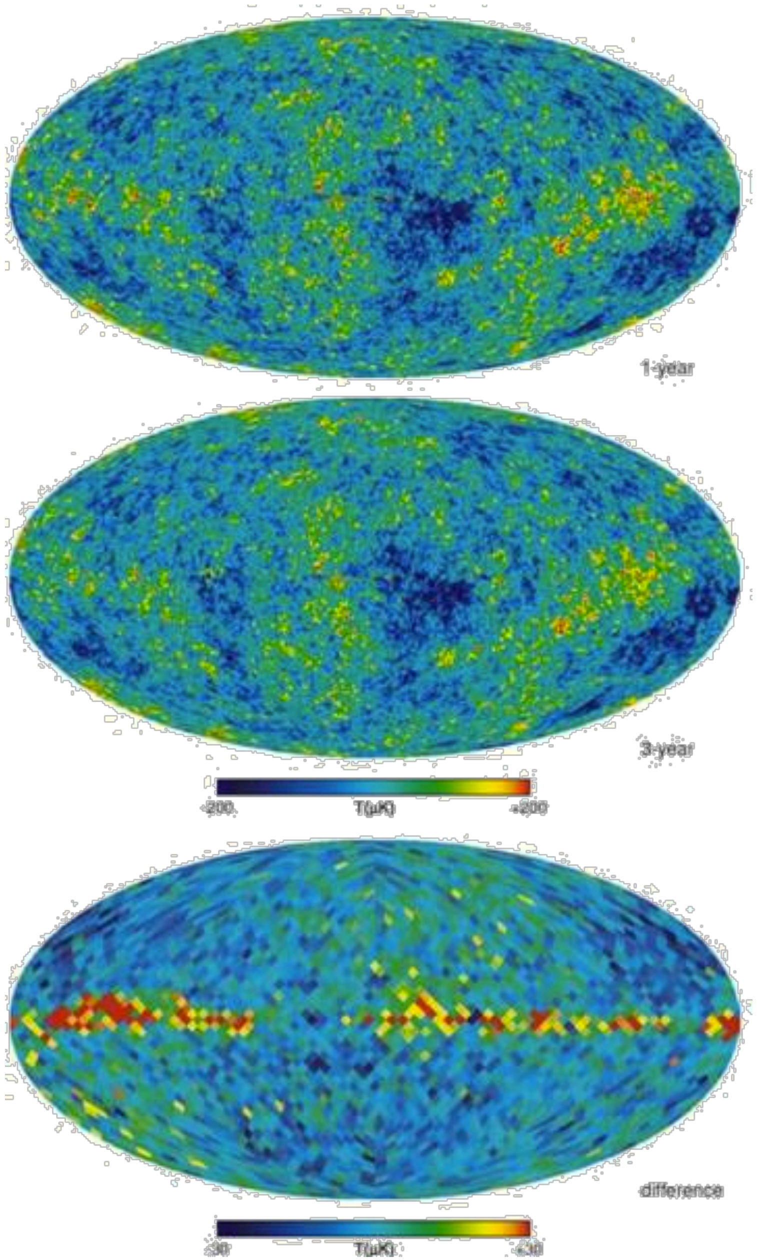
$$\Xi(\ell_1, \ell_2, \tilde{W}) = \sum_{\ell_3} \frac{(2\ell_3+1)}{4\pi} \tilde{W}_{\ell_3} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ 0 & 0 & 0 \end{pmatrix}^2$$

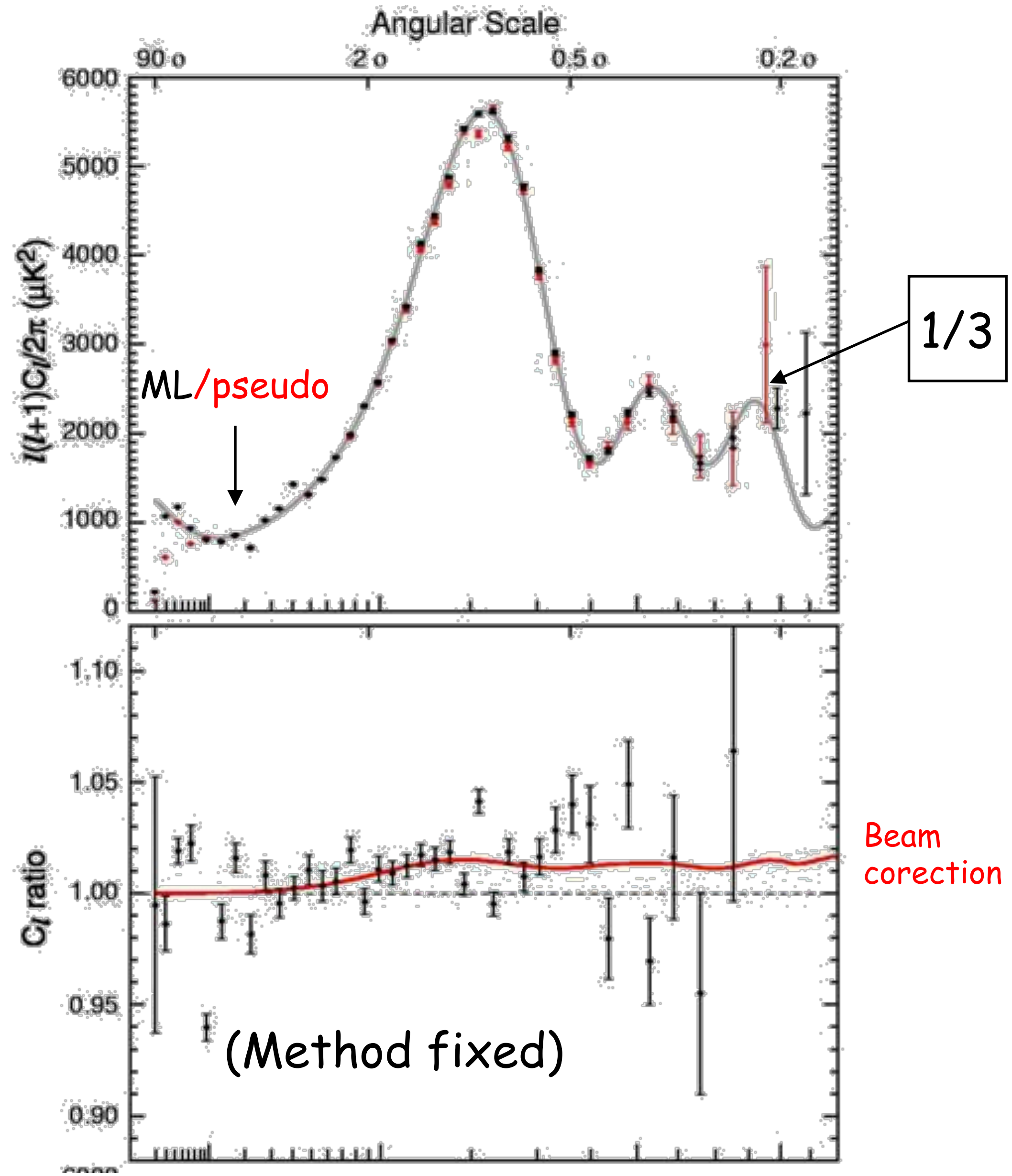
If simple weighting (zeros and ones)

$$\langle \Delta \widehat{C}^2_\ell \rangle \approx \frac{2}{(2\ell+1) f_{sky}} (C^2_\ell + N^2_\ell)$$

Same can be done for polarization, only more complicated …

Application: WMAP power spectrum





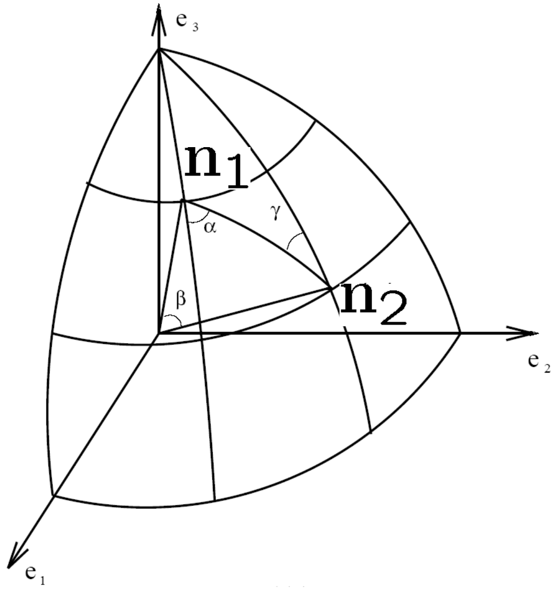
Imagers: polarised spectrum estimation

$$P(\mathbf{n}) = (Q \pm iU)(\mathbf{n}) = \sum_{\ell m} [E_{\ell m} \mp iB_{\ell m}]_{\mp 2}Y_{\ell m}$$

$$\langle E_{lm} E^*_{l'm'} \rangle = \delta_{ll'} \delta_{mm'} C_l^E$$

$$\langle B_{lm} B^*_{l'm'} \rangle = \delta_{ll'} \delta_{mm'} C_l^B$$

$$\langle E_{lm} T^*_{l'm'} \rangle = \delta_{ll'} \delta_{mm'} C_l^{TE}$$



$$D^{-1}(\phi_1, \theta_1, 0) D(\phi_2, \theta_2, 0) = D(\alpha, \beta, -\gamma)$$

$$D^l_{-ms}(\phi, \theta, 0) = (-1)^m \sqrt{\frac{4\pi}{2l+1}} \, {}_sY_{lm}(\hat{\boldsymbol{n}}).$$

$$D^l_{ss'}(\alpha, \beta, -\gamma) = \sum_m \frac{4\pi}{2l+1} \, {}_sY^*_{lm}(\hat{\boldsymbol{n}}_1)_{s'} Y_{lm}(\hat{\boldsymbol{n}}_2)$$

$$\langle \bar{P}(\hat{\boldsymbol{n}}_1) \bar{P}(\hat{\boldsymbol{n}}_2) \rangle = \sum_l \frac{2l+1}{4\pi} \left(C_l^E - C_l^B \right) d^l_{2-2}(\beta)$$

$$\langle \bar{P}^*(\hat{\boldsymbol{n}}_1) \bar{P}(\hat{\boldsymbol{n}}_2) \rangle = \sum_l \frac{2l+1}{4\pi} \left(C_l^E + C_l^B \right) d^l_{22}(\beta)$$

$$\langle T(\hat{\boldsymbol{n}}_1) \bar{P}(\hat{\boldsymbol{n}}_2) \rangle = \sum_l \frac{2l+1}{4\pi} C_l^{TE} d^l_{20}(\beta)$$

$$\bar{P}(\hat{\boldsymbol{n}}_1) \equiv e^{2i\alpha} P(\hat{\boldsymbol{n}}_1)$$

$$\bar{P}(\hat{\boldsymbol{n}}_2) \equiv e^{2i\gamma} P(\hat{\boldsymbol{n}}_2)$$

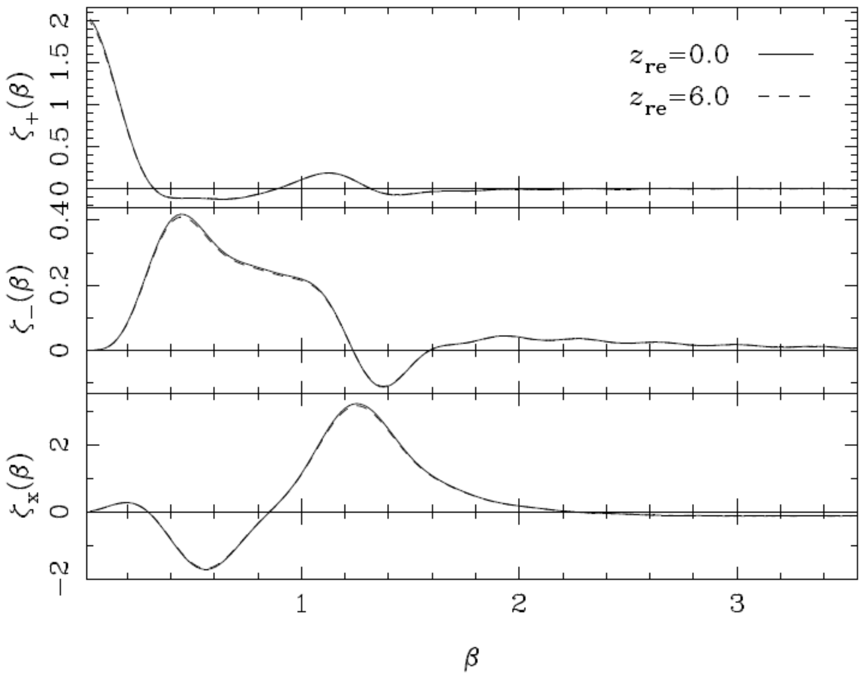
Stokes parameter in the great circle basis

Polarisation: correlation functions

$$\xi_-(\beta) \equiv \langle \bar{P}(\hat{\boldsymbol{n}}_1)\bar{P}(\hat{\boldsymbol{n}}_2)\rangle \qquad \langle \bar{Q}(\hat{\boldsymbol{n}}_1)\bar{Q}(\hat{\boldsymbol{n}}_2)\rangle = \frac{1}{2}[\xi_+(\beta) + \Re\xi_-(\beta)]$$

$$\xi_+(\beta) \equiv \langle \bar{P}^*(\hat{\boldsymbol{n}}_1)\bar{P}(\hat{\boldsymbol{n}}_2)\rangle \qquad \langle \bar{U}(\hat{\boldsymbol{n}}_1)\bar{U}(\hat{\boldsymbol{n}}_2)\rangle = \frac{1}{2}[\xi_+(\beta) - \Re\xi_-(\beta)]$$

$$\xi_X(\beta) \equiv \langle T(\hat{\boldsymbol{n}}_1)\bar{P}(\hat{\boldsymbol{n}}_2)\rangle \qquad \langle T(\hat{\boldsymbol{n}}_1)\bar{Q}(\hat{\boldsymbol{n}}_2)\rangle = \Re\xi_X(\beta)$$



$$\int_{-1}^{1} d^l_{mm'}(\beta) d^{l'}_{mm'}(\beta)\, \mathrm{d}\cos\beta = \frac{2}{2l+1}\delta_{ll'}$$

$$\begin{aligned} C_l^E - C_l^B - 2iC_l^{EB} &= 2\pi \int_{-1}^{1} \xi_-(\beta) d^l_{2\,-2}(\beta)\, \mathrm{d}\cos\beta, \\ C_l^E + C_l^B &= 2\pi \int_{-1}^{1} \xi_+(\beta) d^l_{2\,2}(\beta)\, \mathrm{d}\cos\beta, \\ C_l^{TE} + iC_l^{TB} &= 2\pi \int_{-1}^{1} \xi_X(\beta) d^l_{2\,0}(\beta)\, \mathrm{d}\cos\beta. \end{aligned}$$

Polynomials in cos(β): integrate exactly with Gauss-Legendre quadrature

Polarisation: (fast) CF estimators

Heuristic weighting (w_P,w_T):

$$\hat{C}_+(\psi) = A_P(\psi)\int \mathrm{d}\hat{\boldsymbol{n}}_1\mathrm{d}\hat{\boldsymbol{n}}_2\,[\delta(\hat{\boldsymbol{n}}_1\cdot\hat{\boldsymbol{n}}_2-\cos\psi) \times w_P(\hat{\boldsymbol{n}}_1)w_P(\hat{\boldsymbol{n}}_2)\bar{P}^*(\hat{\boldsymbol{n}}_1)\bar{P}(\hat{\boldsymbol{n}}_2)]$$

$$\hat{C}_-(\psi) = A_P(\psi)\int \mathrm{d}\hat{\boldsymbol{n}}_1\mathrm{d}\hat{\boldsymbol{n}}_2\,[\delta(\hat{\boldsymbol{n}}_1\cdot\hat{\boldsymbol{n}}_2-\cos\psi) \times w_P(\hat{\boldsymbol{n}}_1)w_P(\hat{\boldsymbol{n}}_2)\bar{P}(\hat{\boldsymbol{n}}_1)\bar{P}(\hat{\boldsymbol{n}}_2)]$$

$$\hat{C}_X(\psi) = A_X(\psi)\int \mathrm{d}\hat{\boldsymbol{n}}_1\mathrm{d}\hat{\boldsymbol{n}}_2\,[\delta(\hat{\boldsymbol{n}}_1\cdot\hat{\boldsymbol{n}}_2-\cos\psi) \times w_T(\hat{\boldsymbol{n}}_1)w_P(\hat{\boldsymbol{n}}_2)T(\hat{\boldsymbol{n}}_1)\bar{P}(\hat{\boldsymbol{n}}_2)]$$

Normalization: correlation function of the weights

$$\frac{1}{A_P(\psi)} = \int \mathrm{d}\hat{\boldsymbol{n}}_1\mathrm{d}\hat{\boldsymbol{n}}_2\,[\delta(\hat{\boldsymbol{n}}_1\cdot\hat{\boldsymbol{n}}_2-\cos\psi) \times w_P(\hat{\boldsymbol{n}}_1)w_P(\hat{\boldsymbol{n}}_2)]$$

$$\frac{1}{A_X(\psi)} = \int \mathrm{d}\hat{\boldsymbol{n}}_1\mathrm{d}\hat{\boldsymbol{n}}_2\,[\delta(\hat{\boldsymbol{n}}_1\cdot\hat{\boldsymbol{n}}_2-\cos\psi) \times w_T(\hat{\boldsymbol{n}}_1)w_P(\hat{\boldsymbol{n}}_2)]$$

Using

$$\sum_{l\geqslant\max(|m|,|n|)}\frac{2l+1}{2}d^l_{mn}(\beta)d^l_{mn}(\psi) = \delta(\cos\beta-\cos\psi)$$

for m=n=2 involves

$$w_P(\hat{\boldsymbol{n}}_1)\bar{P}^*(\hat{\boldsymbol{n}}_1)d^l_{22}(\beta)\bar{P}(\hat{\boldsymbol{n}}_2)w_P(\hat{\boldsymbol{n}}_2) = \tilde{P}^*(\hat{\boldsymbol{n}}_1)D^l_{22}(\alpha,\beta,-\gamma)\tilde{P}(\hat{\boldsymbol{n}}_2)$$

with $\tilde{P}(\hat{\boldsymbol{n}}) \equiv w_P(\hat{\boldsymbol{n}})P(\hat{\boldsymbol{n}})$ **Weighted polarization field**

Using

$$\tilde{P}(\hat{\boldsymbol{n}}) = \sum_{lm}(\tilde{E}_{lm}-i\tilde{B}_{lm})_{-2}Y_{lm}(\hat{\boldsymbol{n}})$$

$$\tilde{P}^*(\hat{\boldsymbol{n}}) = \sum_{lm}(\tilde{E}_{lm}+i\tilde{B}_{lm})_{+2}Y_{lm}(\hat{\boldsymbol{n}})$$

We get

$$\hat{C}_+(\psi) = 2\pi A_P(\psi)\sum_{lm} d^l_{22}(\psi)|\tilde{E}_{lm}+i\tilde{B}_{lm}|^2$$

Polarisation: (fast) CF and PS estimators

Define the pseudo-Cls estimates:

$$\tilde{C}_l^E \equiv \frac{1}{2l+1}\sum_m |\tilde{E}_{lm}|^2$$

$$\tilde{C}_l^B \equiv \frac{1}{2l+1}\sum_m |\tilde{B}_{lm}|^2$$

$$\tilde{C}_l^{EB} \equiv \frac{1}{2l+1}\sum_m \tilde{E}_{lm}\tilde{B}^*_{lm} = \frac{1}{2l+1}\sum_m \tilde{B}_{lm}\tilde{E}^*_{lm}$$

$$\tilde{C}_l^{TB} \equiv \frac{1}{2l+1}\sum_m \tilde{T}_{lm}\tilde{B}^*_{lm} = \frac{1}{2l+1}\sum_m \tilde{B}_{lm}\tilde{T}^*_{lm}$$

These can be computed using fast SPH transforms in $O(n_{pix}^{3/2})$ (compare to $o(n_{pix}^{3})$ scaling of ML…)

$$\hat{C}_+(\psi) = 2\pi A_P(\psi)\sum_l (2l+1)d^l_{22}(\psi)(\tilde{C}_l^E + \tilde{C}_l^B)$$

$$\hat{C}_-(\psi) = 2\pi A_P(\psi)\sum_l (2l+1)d^l_{2\,-2}(\psi)(\tilde{C}_l^E - \tilde{C}_l^B - 2i\tilde{C}_l^{EB})$$

$$\hat{C}_X(\psi) = 2\pi A_X(\psi)\sum_l (2l+1)d^l_{2\,0}(\psi)(\tilde{C}_l^{TE} - i\tilde{C}_l^{TB})$$

$$\frac{1}{A_P(\psi)} = 2\pi \sum_{l\geqslant 0}(2l+1)P_l(\cos\psi)w_{P,l}$$

$$w_{P,l} = \frac{1}{2l+1}\sum_m |w_{P,lm}|^2$$

If CF measured at all angles: integrate with GL quadrature

$$C_l^E - C_l^B - 2iC_l^{EB} = 2\pi\int_{-1}^{1}\xi_-(\beta)d^l_{2\,-2}(\beta)\,\mathrm{d}\cos\beta,$$

$$C_l^E + C_l^B = 2\pi\int_{-1}^{1}\xi_+(\beta)d^l_{2\,2}(\beta)\,\mathrm{d}\cos\beta,$$

$$C_l^{TE} + iC_l^{TB} = 2\pi\int_{-1}^{1}\xi_X(\beta)d^l_{2\,0}(\beta)\,\mathrm{d}\cos\beta.$$

– – Assuming parity invariance

Polarisation: CF estimators on finite surveys

Incomplete measurement of correlation function: apodizing function f(β):

$$\langle \hat{C}_l^E \pm \hat{C}_l^B \rangle = \sum_{l'} {}_{\pm 2}K_{ll'} (C_{l'}^E \pm C_{l'}^B)$$

$${}_{\pm 2}K_{ll'} \equiv \frac{2l'+1}{2} \int f(\beta) d^l_{2\,\pm 2}(\beta) d^{l'}_{2\,\pm 2}(\beta)\, \mathrm{d}\cos\beta$$

$${}_{2}K_{ll'} = \frac{2l'+1}{2} \sum_L (2L+1) f_L \begin{pmatrix} l & l' & L \\ 2 & -2 & 0 \end{pmatrix}^2$$

$$f(\psi) = \sum_{l \geqslant 0} \frac{2l+1}{2} f_l P_l(\cos\psi)$$

Normalization of the window functions

$$\sum_{l'} {}_2K_{ll'} = \sum_{L \geqslant 0} \frac{2L+1}{2} f_L = f(0)$$

$$\sum_{l'} {}_{-2}K_{ll'} = \int f(\beta) \csc^2(\beta/2) d^l_{2\,-2}(\beta)\, \mathrm{d}\cos\beta$$

Results in E/B modes leakage

$$\langle \hat{C}_l^E \rangle = \sum_{l'} ({}_{+}K_{ll'} C_{l'}^E + {}_{-}K_{ll'} C_{l'}^B)$$

$$\langle \hat{C}_l^B \rangle = \sum_{l'} ({}_{-}K_{ll'} C_{l'}^E + {}_{+}K_{ll'} C_{l'}^B)$$

$${}_{\pm}K_{ll'} \equiv ({}_2K_{ll'} \pm {}_{-2}K_{ll'})/2$$

Cut-sky effects: E-B mixing

- Mixing occurs from line integrals on the border
- Define STF windows that project out E contribution
- This can be achieved by SVD of coupling matrix
- For each m, 2 modes are lost

$$B'_W \equiv -2 \int_S \mathrm{d}S\, W^* \epsilon^b{}_c \nabla^c \nabla^a \mathcal{P}_{ab},$$

$$B'_W = \sqrt{2} \int_S \mathrm{d}S\, W_B^{ab*} \mathcal{P}_{ab}$$

$$- 2 \oint_{\partial S} \mathrm{d}l^a \left(\epsilon^b{}_a W^* \nabla^c \mathcal{P}_{cb} - \epsilon^b{}_c \nabla^c W^* \mathcal{P}_{ab}\right)$$

- Separation is done at the map level
- Block-diagonal structure of coupling allows to gain CPU time for azimuthally symmetric patches
- Pixel effects can be important if no quadrature sampling … (e.g. Bunn et al. 2002)

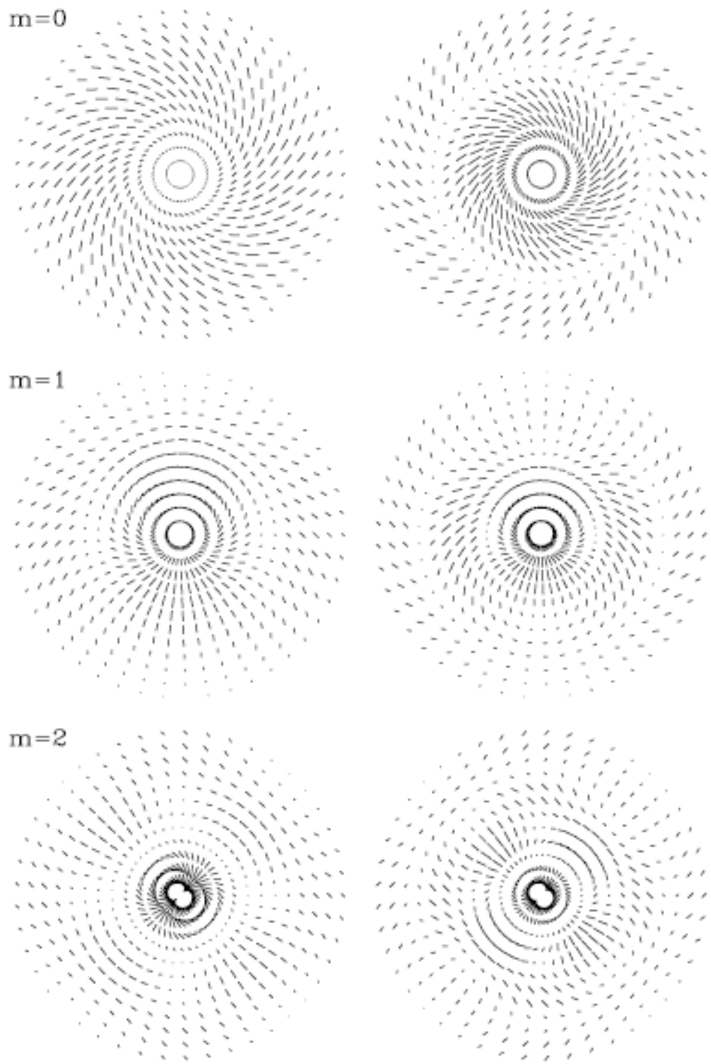


FIG. 1: The real space window functions for an azimuthally symmetric sky patch with $\theta < 10°$, evaluated in the frame where the signal is diagonal so the leftmost window produces the largest signal for that m.

Lewis, Challinor, Turok 2001

E-B mixing: statistical separation

- Use integrals of the Stokes correlations functions over observed angular range to construct pure E and B statistics
- Originally derived for lensing (Crittenden et al. 2002)
- Generalized to the sphere (Chon et al. 2004) and coupled to fast, edge-corrected estimation of correlation functions

OR

- Use the coupling kernels of polarised pseudo-Cls (Hansen & Gorski 2003)
- Generalise MASTER (or FASTER) method
- Regularised (binned) inversion of coupling kernel
- This was used in the B03 data processing

$$(\hat{C}_l^E \pm \hat{C}_l^B) = \sum_{l'} K_{\pm 2}(l, l')(C_{l'}^E + C_{l'}^B)$$

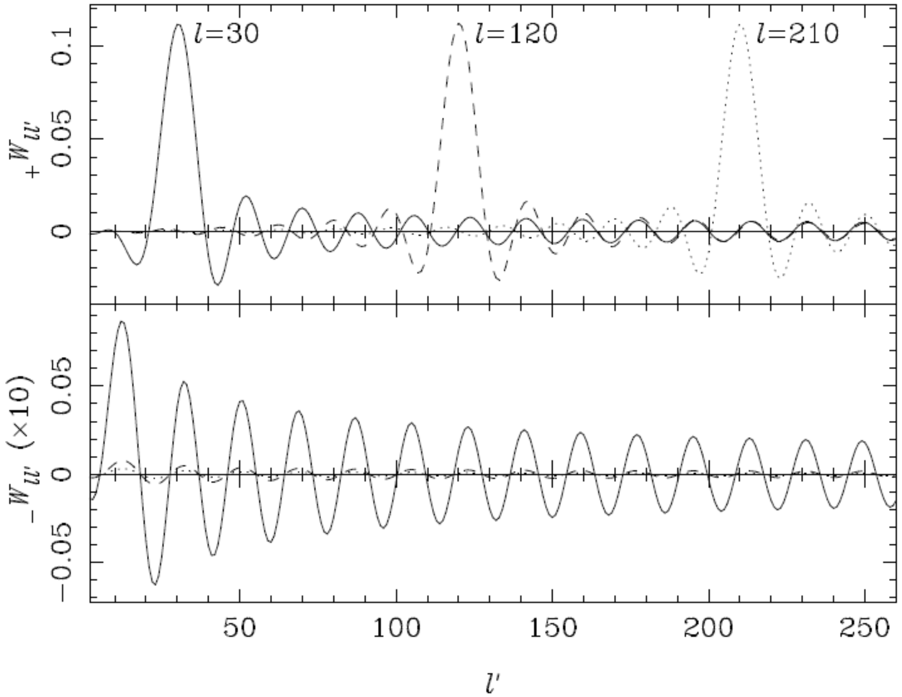
$$K_{\pm 2}(l, l') = \frac{2l'+1}{2} \sum_L (\pm 1)^{l+l'+L} (2L+1) W_L \begin{pmatrix} l & l' & L \\ 2 & -2 & 0 \end{pmatrix}^2$$

Fast decoupled, edge-corrected estimators of polarized spectra available

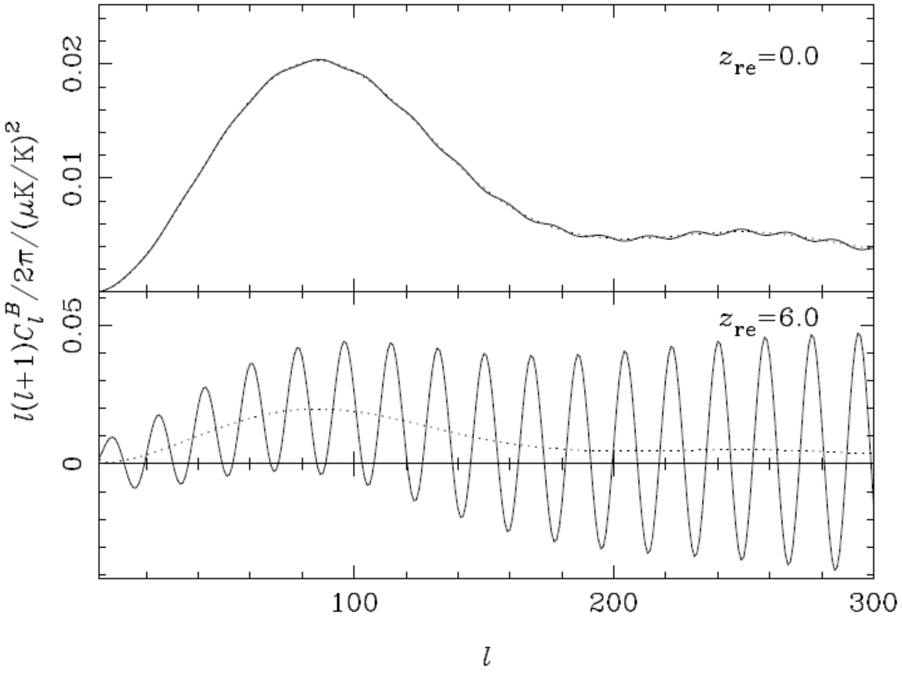
- E-B separation only in the mean !
- E-mode cosmic variance leaks into B-mode variance
- Only valid for sufficiently large surveys (Challinor & Chon 2005)

Polarisation: E/B coupling of cut-sky

Leakage window functions (not normalized)



Recovered BB spectra (dots)

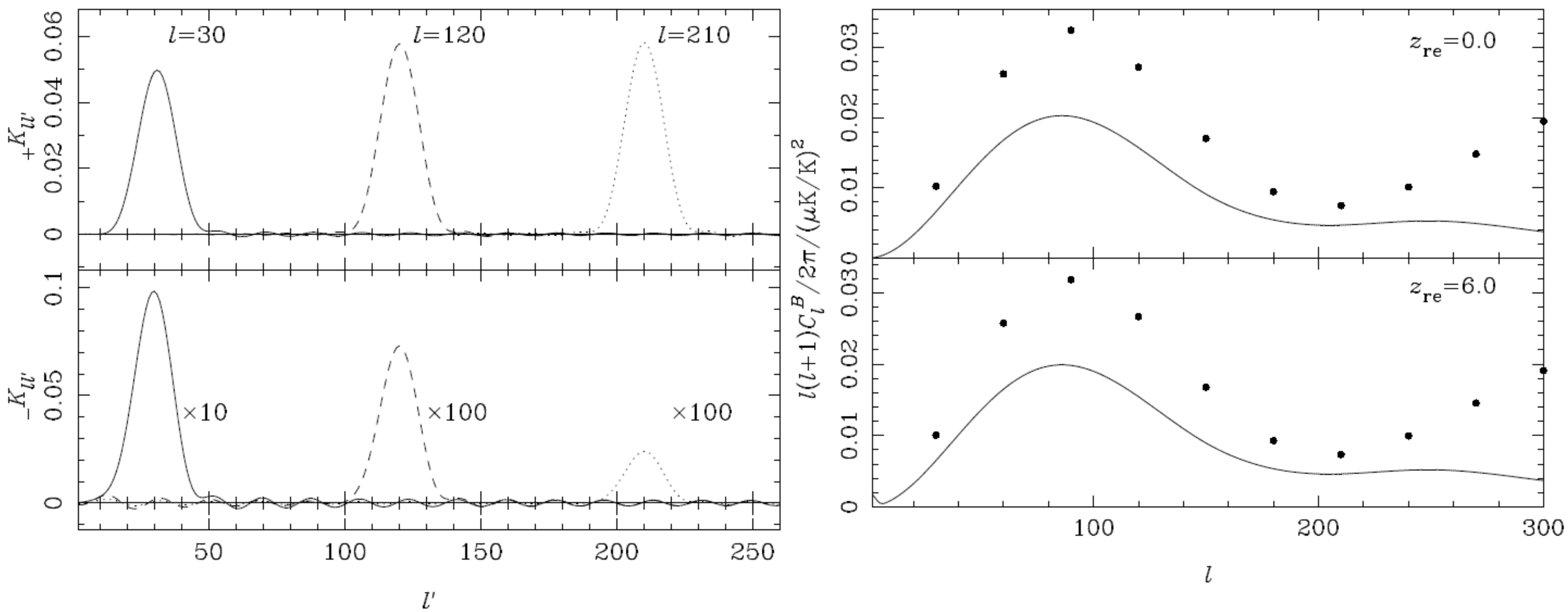


No correlation function information over β_{max}=20±

Polarisation: E/B coupling of cut-sky

Leakage window functions (not normalized)

Recovered BB spectra (dots)



No correlation function over
Gaussian apodization

$$\beta_{max} = 20^\circ$$
$$\sigma = 10^\circ$$

Polarisation: E/B leakage correction

Define:

$$\xi(\beta) \equiv \sum_l \frac{2l+1}{4\pi}(C_l^E + C_l^B) d^l_{2\,-2}(\beta)$$

Then:

$$\frac{1}{2}[\xi(\beta) - \Re\xi_-(\beta)] = \sum_l \frac{2l+1}{4\pi} C_l^B d^l_{2\,-2}(\beta)$$

$$\hat{C}_l^B = 2\pi \int \frac{1}{2}[\hat{\xi}(\beta) - \Re\hat{\xi}_-(\beta)] f(\beta) d^l_{2\,-2}(\beta)\, \mathrm{d}\cos\beta$$

As a function of ξ_+

$$\xi(\beta) = \int_{-1}^{1} \mathrm{d}\cos\beta'\, \xi_+(\beta') \sum_l \frac{2l+1}{2} d^l_{2\,-2}(\beta) d^l_{22}(\beta')$$

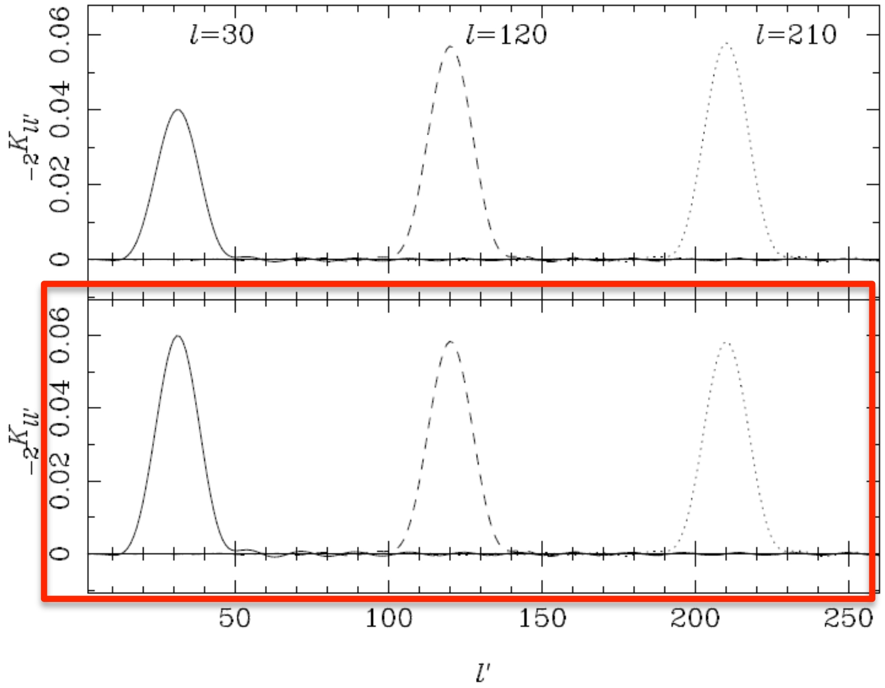
$$d^l_{2\,-2}(\beta) = d^l_{22}(\beta) - \frac{2(2+\cos\beta)}{\sin^4(\beta/2)} \int_0^\beta \tan^3(\beta'/2) d^l_{22}(\beta')\, \mathrm{d}\beta' + \frac{2}{\sin^2(\beta/2)} \int_0^\beta \sec^3(\beta'/2)\sin(\beta'/2) d^l_{22}(\beta')\, \mathrm{d}\beta'$$

$$\begin{aligned}\xi(\beta) &= \xi_+(\beta) + \frac{1}{\sin^2(\beta/2)} \int_{\cos\beta}^{1} \mathrm{d}\cos\beta'\, \xi_+(\beta') \sec^4(\beta'/2) \\ &\quad - \frac{2(2+\cos\beta)}{\sin^4(\beta/2)} \int_{\cos\beta}^{1} \mathrm{d}\cos\beta'\, \xi_+(\beta') \frac{\tan^3(\beta'/2)}{\sin\beta'}\end{aligned}$$

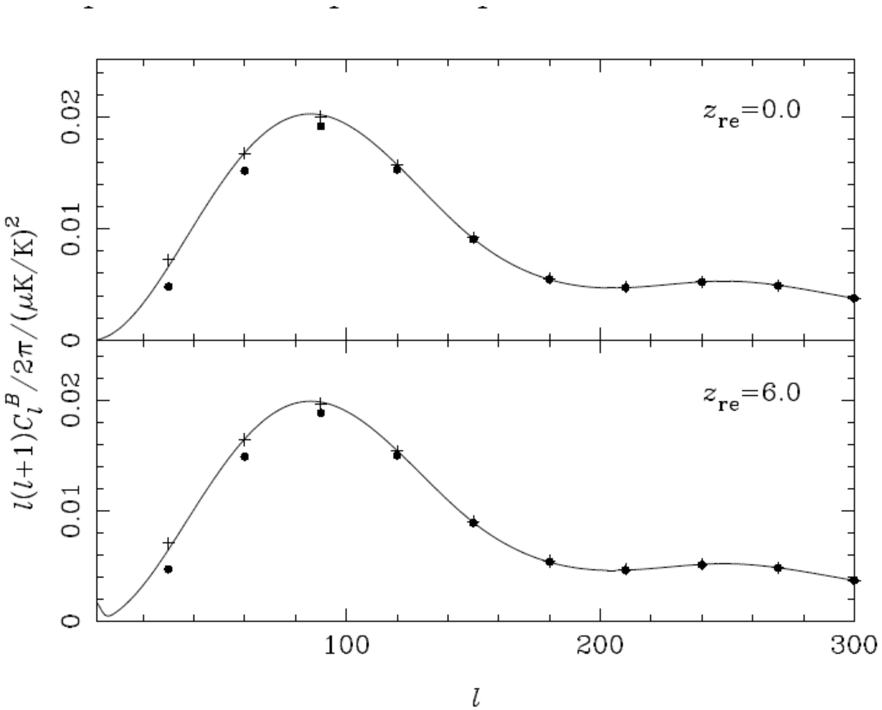
We have obtained pure E and B spectra (in the mean)

$$\langle \hat{C}_l^B \rangle = \sum_{l'} {}_{-2}K_{ll'} C_{l'}^B \qquad \langle \hat{C}_l^E \rangle = \sum_{l'} {}_{-2}K_{ll'} C_{l'}^E$$

Results on small survey simulation



Renormalized windows



- Solid line: input BB spectrum
- bullets: recovered BB (unnormalized)
- + : recovered BB (normalized)

Polarized case: harmonic point of view

$$(Q \pm \mathrm{i}U)(\hat{\boldsymbol{n}}) = \sum_{lm} (E_{lm} \mp \mathrm{i}B_{lm})_{\mp 2}Y_{lm}(\hat{\boldsymbol{n}}).$$

$$\tilde{E}_{lm} \pm \mathrm{i}\tilde{B}_{lm} = \sum_{(lm)'} {}_{\pm 2}I_{(lm)(lm)'}\left[E_{(lm)'} \pm \mathrm{i}B_{(lm)'}\right]$$

$${}_{\pm 2}I_{(lm)(lm)'} = \int \mathrm{d}\hat{\boldsymbol{n}}\, w(\hat{\boldsymbol{n}})_{\pm 2}Y_{(lm)'}(\hat{\boldsymbol{n}})_{\pm 2}Y^{*}_{lm}(\hat{\boldsymbol{n}}).$$

$${}_{\pm 2}I_{(lm)(lm)'} = \sum_{LM} (-1)^m w_{LM} \sqrt{\frac{(2L+1)(2l+1)(2l'+1)}{4\pi}} \begin{pmatrix} l & l' & L \\ m & -m' & -M \end{pmatrix} \begin{pmatrix} l & l' & L \\ \mp 2 & \pm 2 & 0 \end{pmatrix}$$

$$\tilde{E}_{lm} = \sum_{(lm)'} \left[{}_{+}I_{lm(lm)'}E_{(lm)'} + \mathrm{i}\,{}_{-}I_{(lm)(lm)'}B_{(lm)'}\right],$$

$$\tilde{B}_{lm} = \sum_{(lm)'} \left[{}_{+}I_{(lm)(lm)'}B_{(lm)'} - \mathrm{i}\,{}_{-}I_{(lm)(lm)'}E_{(lm)'}\right],$$

$${}_{+}I_{(lm)(lm)'} \equiv \frac{1}{2}\left({}_{+2}I_{(lm)(lm)'} + {}_{-2}I_{(lm)(lm)'}\right),$$

$${}_{-}I_{(lm)(lm)'} \equiv \frac{1}{2}\left({}_{+2}I_{(lm)(lm)'} - {}_{-2}I_{lm(lm)'}\right)$$

Challinor & Chon 05
Hansen et al. 03

Pseudo-spectra: continued

$$\langle \tilde{C}_l^E \rangle = \sum_{l'} \left(P_{ll'} C_{l'}^E + M_{ll'} C_{l'}^B \right), \qquad \langle \tilde{C}_l^B \rangle = \sum_{l'} \left(M_{ll'} C_{l'}^E + P_{ll'} C_{l'}^B \right) \qquad \langle \tilde{C}_l^{EB} \rangle = \sum_{l'} (P_{ll'} - M_{ll'}) C_{l'}^{EB}$$

$$P_{ll'} \equiv \frac{1}{2l+1} \sum_{mm'} \left| {}_{+}I_{(lm)(lm)'} \right|^2 = \frac{2l'+1}{8\pi} \sum_L (2L+1) w_L [1 + (-1)^K] \begin{pmatrix} l & l' & L \\ -2 & 2 & 0 \end{pmatrix}^2$$

$$M_{ll'} \equiv \frac{1}{2l+1} \sum_{mm'} \left| {}_{-}I_{(lm)(lm)'} \right|^2 = \frac{2l'+1}{8\pi} \sum_L (2L+1) w_L [1 - (-1)^K] \begin{pmatrix} l & l' & L \\ -2 & 2 & 0 \end{pmatrix}^2$$

$$K \equiv l + l' + L$$

$$\sum_{l'} P_{ll'} + M_{ll'} = \sum_L \frac{2L+1}{4\pi} w_L = w^{(2)} f_{\rm sky}$$

$$\sum_{l'} M_{ll'} \sim \frac{1}{2\pi} \int \frac{|\nabla w|^2}{l(l+1)} \, \mathrm{d}\hat{\boldsymbol{n}} \qquad (l \gg l_{\max}).$$

$$4\pi w^{(i)} f_{\rm sky} \equiv \int w^i(\hat{\boldsymbol{n}}) \, \mathrm{d}\hat{\boldsymbol{n}}$$

Recovered spectra (E/B statistical separation)

$$\hat{C}_l^E + \hat{C}_l^B = \sum (P + M)^{-1}_{ll'} \left(\tilde{C}_l^E + \tilde{C}_l^B\right)$$

$$\hat{C}_l^E - \hat{C}_l^B = \sum_{l'} (P - M)^{-1}_{ll'} \left(\tilde{C}_l^E - \tilde{C}_l^B\right)$$

Only invertible if correlations functions are measured over [0,π], otherwise need regularization (e.g. binning)

$$\sum_{l'} \left(\tilde{P}^{-1}_{ll'} P_{l'l''} + \tilde{M}^{-1}_{ll'} M_{l'l''}\right) = {}_{-2}\bar{K}_{ll''},$$

$$\sum_{l'} \left(\tilde{P}^{-1}_{ll'} M_{l'l''} + \tilde{M}^{-1}_{ll'} P_{l'l''}\right) = 0.$$

Using (apodized) correlation functions is a way to find pseudo-inverses with specific properties

Edge-corrected, E/B decoupled estimators:

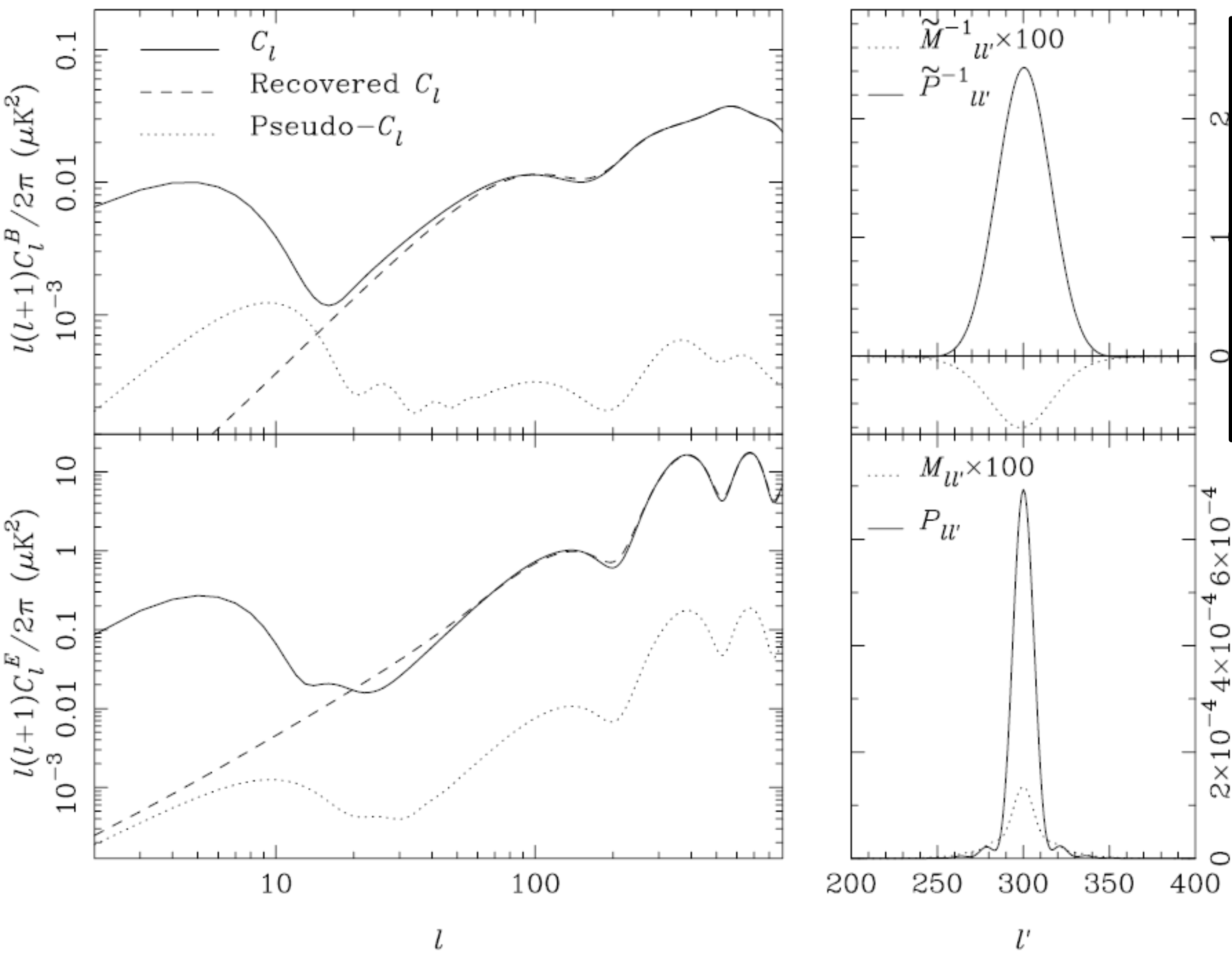
$$\hat{C}_l^E = \sum_{l'} \left(\tilde{P}^{-1}_{ll'} \tilde{C}^E_{l'} + \tilde{M}^{-1}_{ll'} \tilde{C}^B_{l'}\right), \qquad \hat{C}_l^B = \sum_{l'} \left(\tilde{P}^{-1}_{ll'} \tilde{C}^B_{l'} + \tilde{M}^{-1}_{ll'} \tilde{C}^E_{l'}\right)$$

$$\langle \hat{C}_l^E \rangle = \sum_{l'} {}_{-2}\bar{K}_{ll'} C^E_{l'}$$

and idem for B modes

$${}_{-2}K_{ll'} \equiv \frac{2l'+1}{2} \int f(\beta)\, d^l_{2\,-2}(\beta)\, d^{l'}_{2\,-2}(\beta)\, \mathrm{d}\cos\beta$$

A small survey exemple:15 degree radius coverage



Conservative regularization (apodization) hence poor spectral resolution of recovered spectra

Figure 1. Left: power spectra for B (top) and E (bottom; solid lines) compared with the mean pseudo-C_ls (dotted lines) and the recovered power spectra (dashed lines). Right: the bottom panel shows representative window functions $P_{ll'}$ (solid lines) and $M_{ll'}$ (dotted lines) that give the mean pseudo-C_ls on convolving with the true C_ls; the top panel shows the pseudo-inverses $\tilde{P}^{-1}_{ll'}$ (solid lines) and $\tilde{M}^{-1}_{ll'}$ (dotted lines) that when convolved with the mean pseudo-C_ls remove the effect of E–B mixing. Note that $M_{ll'}$ and $\tilde{M}^{-1}_{ll'}$ have been multiplied by a factor of 100 for clarity. The weight function applied to the map is uniform inside a circle of 10° radius, with cosine apodization out to 15°. To obtain the pseudo-inverses, a Gaussian apodization of 4° HWHM is applied to the correlation functions.

Exact covariance properties

$$\mathrm{cov}\left(\tilde{C}^E_l, \tilde{C}^E_{l'}\right) = \frac{2}{(2l+1)(2l'+1)} \sum_{mm'} \left| \left\langle \tilde{E}_{lm} \tilde{E}^*_{(lm)'} \right\rangle \right|^2.$$

$$\mathrm{cov}\left(\tilde{C}^E_l, \tilde{C}^E_{l'}\right) = \frac{2}{(2l+1)(2l'+1)} \sum_{mm'} \left| \sum_{LM} {}_{+}I_{(lm)(LM)} {}_{+}I^*_{(lm)'(LM)} C^E_L + {}_{-}I_{(lm)(LM)} {}_{-}I^*_{(lm)'(LM)} C^B_L \right|^2$$

$$\mathrm{cov}\left(\tilde{C}^B_l, \tilde{C}^B_{l'}\right) = \frac{2}{(2l+1)(2l'+1)} \sum_{mm'} \left| \sum_{LM} {}_{+}I_{(lm)(LM)} {}_{+}I^*_{(lm)'(LM)} C^B_L + {}_{-}I_{(lm)(LM)} {}_{-}I^*_{(lm)'(LM)} C^E_L \right|^2$$

$$\mathrm{cov}\left(\tilde{C}^E_l, \tilde{C}^B_{l'}\right) = \frac{2}{(2l+1)(2l'+1)} \sum_{mm'} \left| \sum_{LM} {}_{+}I_{(lm)(LM)} {}_{-}I^*_{(lm)'(LM)} C^E_L + {}_{-}I_{(lm)(LM)} {}_{+}I^*_{(lm)'(LM)} C^B_L \right|^2$$

Exemple: BB covariance, with or without E leakage

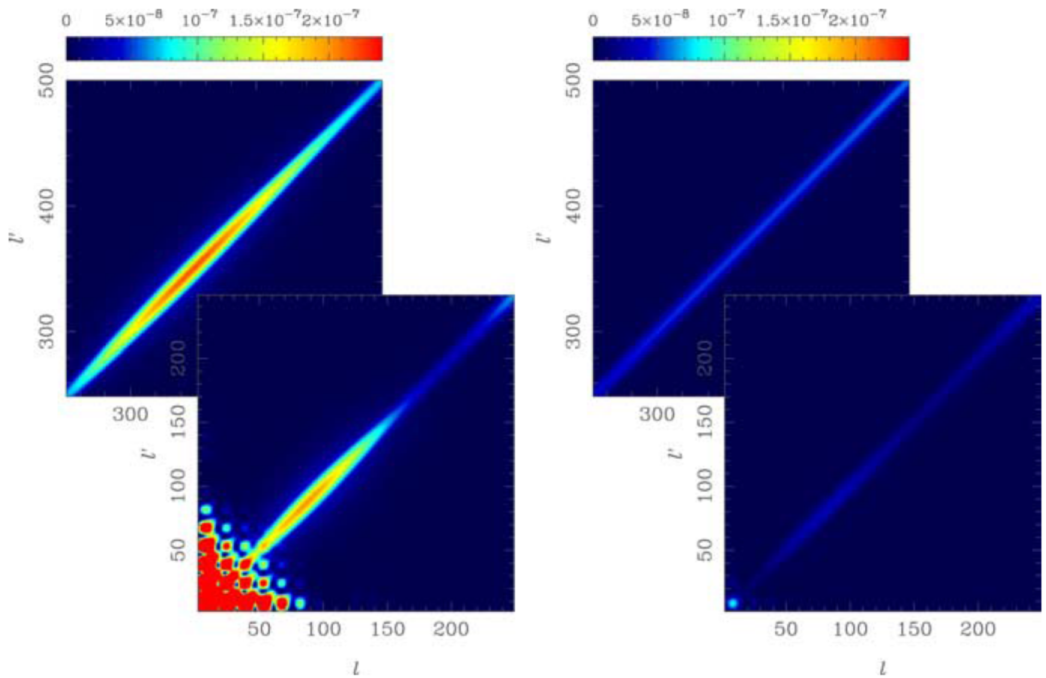


Figure 5. Blocks of the covariance matrix $l(l+1)l'(l'+1)\mathrm{cov}(\tilde{C}^B_l, \tilde{C}^B_{l'})$ for a 15°-radius region with weighting as in Fig. 1. The exact covariance matrix is shown on the left-hand side, and the contribution from B modes is shown on the right (i.e. with C^E_l set to zero).

Approximations: first stage

For sufficiently smooth windows

- $\eth^2 w_2 Y_{\ell m}$ neglected compared to e.g. $\eth w \eth\ {}_2Y_{\ell m}$
- Spectra taken out from inner convolution

Spin raising/lowering operators

$$ {}_2Y_{lm} = \sqrt{\frac{(l-2)!}{(l+2)!}}\eth^2 Y_{lm}, \qquad {}_{-2}Y_{lm} = \sqrt{\frac{(l-2)!}{(l+2)!}}\bar{\eth}^2 Y_{lm} $$

$$ \mathrm{cov}\left(\tilde{C}_l^E, \tilde{C}_{l'}^E\right) = \frac{1}{4\pi}\sum_{LM}\Bigg\{[1+(-1)^K]\Bigg|\sqrt{C_l^E C_{l'}^E}(w^2)_{LM}\begin{pmatrix} l & l' & L \\ -2 & 2 & 0\end{pmatrix} + \frac{2\left(\sqrt{C_l^E C_{l'}^E}-\sqrt{C_l^B C_{l'}^B}\right)}{\sqrt{l(l+1)l'(l'+1)}}\left[|\eth w|^2_{LM}\begin{pmatrix} l & l' & L \\ -1 & 1 & 0\end{pmatrix} + \mathcal{E}_{LM}\begin{pmatrix} l & l' & L \\ -1 & -1 & 2\end{pmatrix}\right]\Bigg|^2\Bigg\} + \frac{1}{\pi}\left[\frac{\sqrt{C_l^E C_{l'}^E}-\sqrt{C_l^B C_{l'}^B}}{\sqrt{l(l+1)l'(l'+1)}}\right]^2 \sum_{LM}[1-(-1)^K]|\mathcal{B}_{LM}|^2\begin{pmatrix} l & l' & L \\ -1 & -1 & 2\end{pmatrix}^2 $$

$$ {}_2(\eth w)^2_{lm} = \mathcal{E}_{lm} + \mathrm{i}\mathcal{B}_{lm}, $$
$$ {}_{-2}(\bar{\eth} w)^2_{lm} = \mathcal{E}_{lm} - \mathrm{i}\mathcal{B}_{lm}. $$

$$ \mathrm{cov}\left(C_l^B, \tilde{C}_{l'}^B\right) = \frac{1}{4\pi}\sum_{LM}\Bigg\{[1+(-1)^K]\Bigg|\sqrt{C_l^B C_{l'}^B}(w^2)_{LM}\begin{pmatrix} l & l' & L \\ -2 & 2 & 0\end{pmatrix} + \frac{2\left(\sqrt{C_l^B C_{l'}^B}-\sqrt{C_l^E C_{l'}^E}\right)}{\sqrt{l(l+1)l'(l'+1)}}\left[|\eth w|^2_{LM}\begin{pmatrix} l & l' & L \\ -1 & 1 & 0\end{pmatrix} + \mathcal{E}_{LM}\begin{pmatrix} l & l' & L \\ -1 & -1 & 2\end{pmatrix}\right]\Bigg|^2\Bigg\} + \frac{1}{\pi}\left[\frac{\sqrt{C_l^B C_{l'}^B}-\sqrt{C_l^E C_{l'}^E}}{\sqrt{l(l+1)l'(l'+1)}}\right]^2 \sum_{LM}[1-(-1)^K]|\mathcal{B}_{LM}|^2\begin{pmatrix} l & l' & L \\ -1 & -1 & 2\end{pmatrix}^2, $$

$$ \mathrm{cov}\left(\tilde{C}_l^E, \tilde{C}_{l'}^B\right) \approx \frac{1}{16\pi}\sum_{LM}[1-(-1)^K]\left(\sqrt{C_l^E C_{l'}^E}+\sqrt{C_l^B C_{l'}^B}\right)^2|(w^2)_{LM}|^2\begin{pmatrix} l & l' & L \\ -2 & 2 & 0\end{pmatrix}^2 $$

Illustration: EE

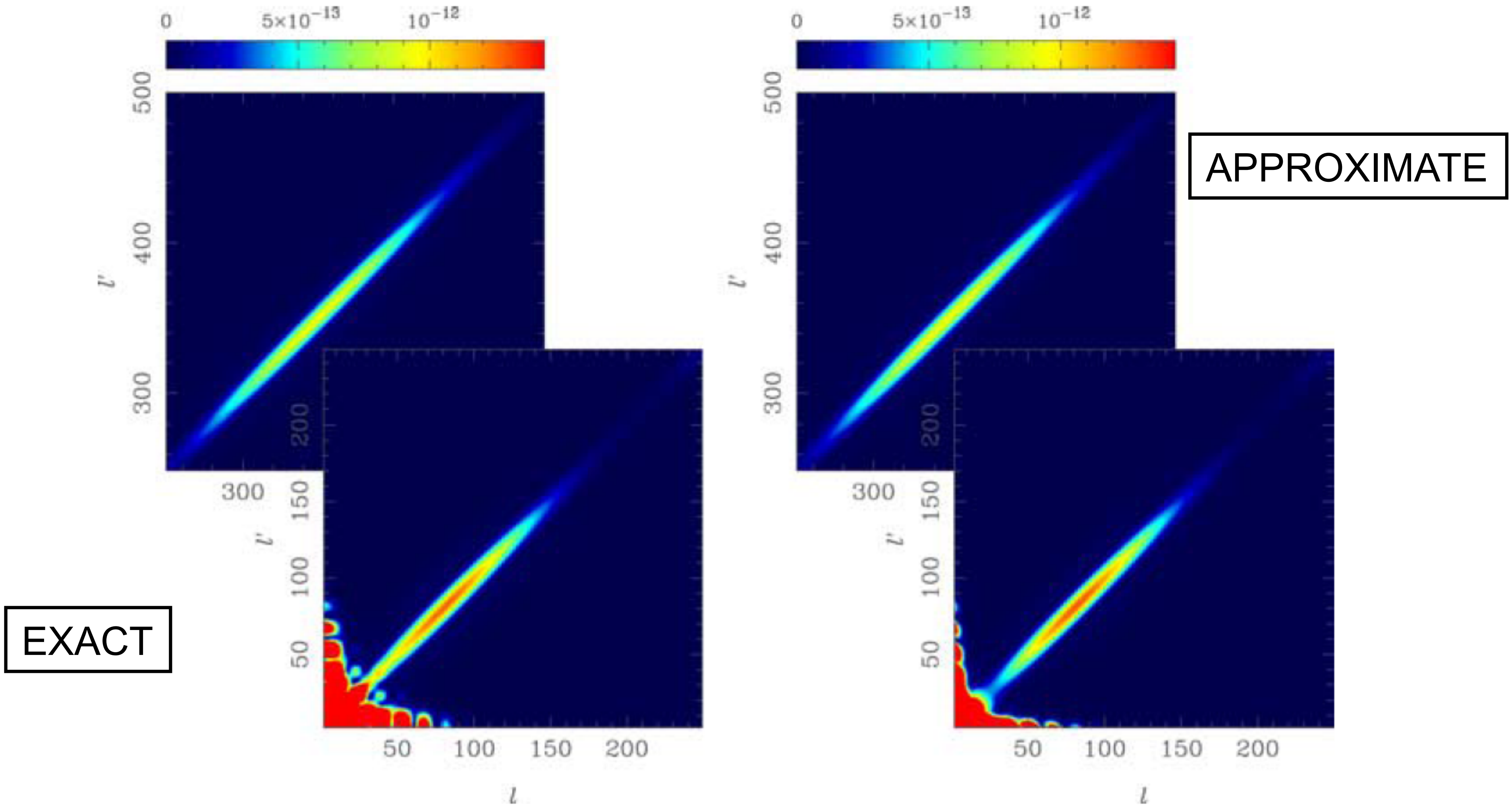


Figure 6. Blocks of the covariance matrix $\mathrm{cov}(\tilde{C}^E_l, \tilde{C}^E_{l'})$ for a 15° radius region with the same weighting as in Fig. 1. The exact covariance matrix is shown on the left-hand side, and its approximation, equation (78), on the right.

Illustration: BB

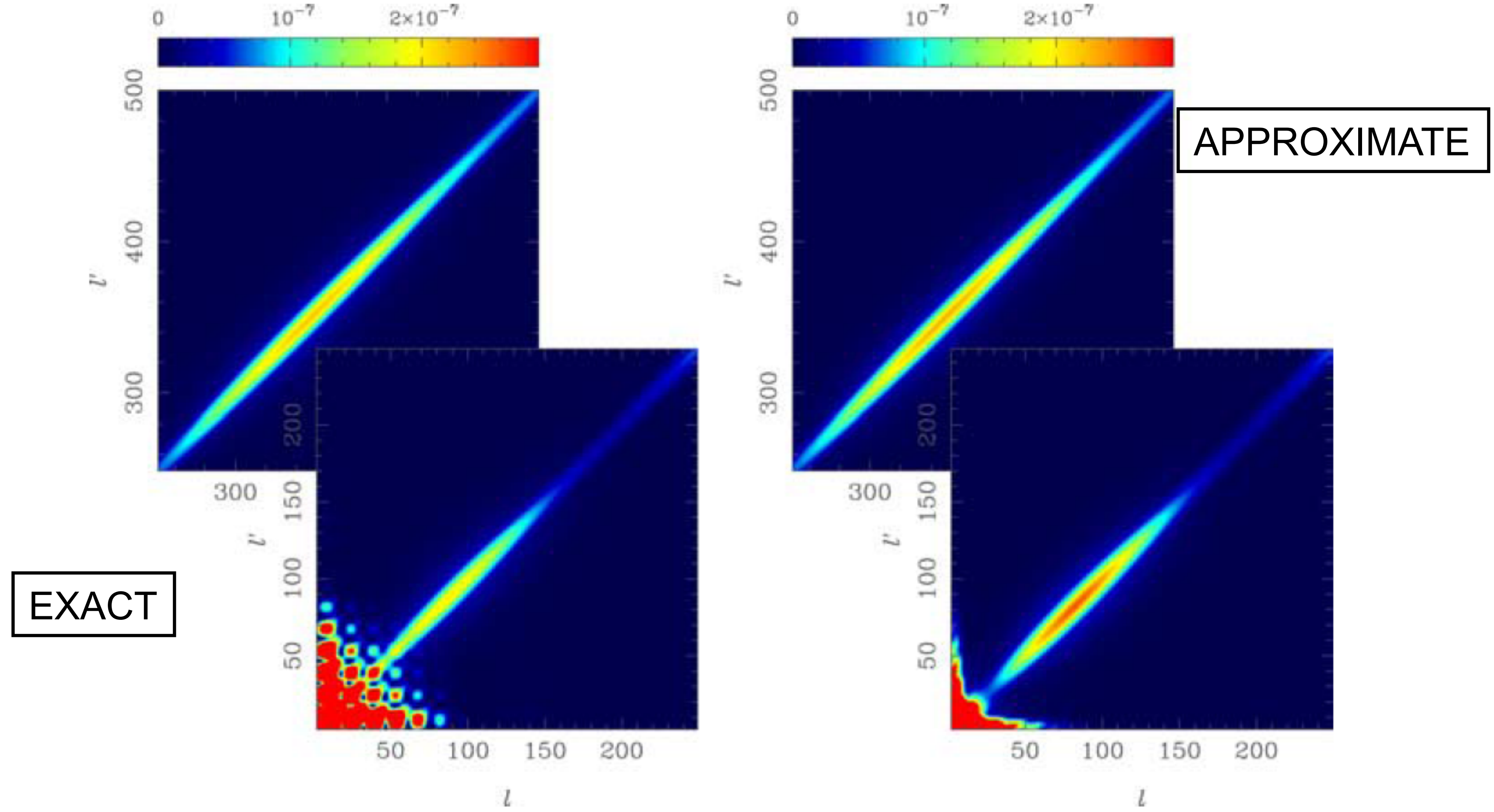


Figure 7. As Fig. 6 but for the covariance matrix $l(l+1)l'(l'+1)\mathrm{cov}(\tilde{C}_l^B, \tilde{C}_{l'}^B)$. (Note the colour scale in this figure differs slightly from that in Fig. 5.)

Illustration: EB

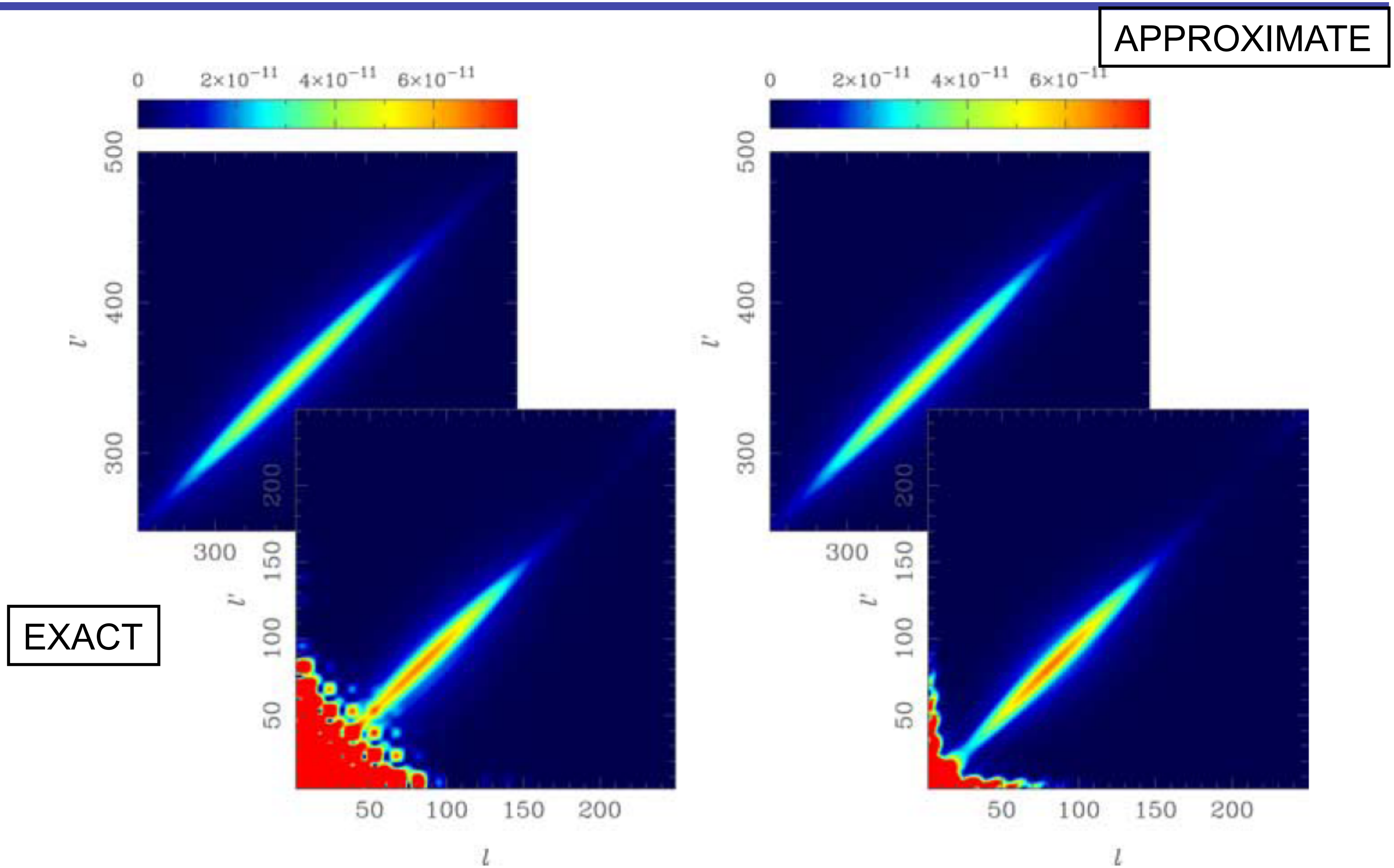


Figure 9. As Fig. 6 but for the covariance matrix $ll' \mathrm{cov}(\tilde{C}_l^E, \tilde{C}_{l'}^B)$.

Approximations: second stage

Approximate covariance of pseudo-spectra band powers:

$$\sum_{l'=\bar{l}-\Delta l/2}^{\bar{l}+\Delta_l/2} \mathrm{cov}\left(\tilde{C}_l^E, \tilde{C}_{l'}^E\right) \approx \frac{2}{2\bar{l}+1}\frac{C_l^E C_{\bar{l}}^E}{4\pi}\sum_{LM}|(w^2)_{LM}|^2 = \frac{2}{2\bar{l}+1} w^{(4)} f_{\rm sky} C_l^E C_{\bar{l}}^E$$

$$\mathrm{var}\left(\tilde{C}_{\bar{l}}^E\right) \approx \frac{2}{(2\bar{l}+1)\Delta l} w^{(4)} f_{\rm sky} C_{\bar{l}}^{E\,2} \qquad \left(C_l^B = 0\right).$$

$$\mathrm{var}(\tilde{C}_{\bar{l}}^B) \approx \frac{3}{(2\bar{l}+1)\bar{l}^2(\bar{l}+1)^2\Delta l}\frac{C_{\bar{l}}^{E\,2}}{\pi}\int (\nabla w)^4 \, \mathrm{d}\hat{\boldsymbol{n}} \qquad (C_l^B = 0).$$

$$\mathrm{cov}\left(\tilde{C}_{\bar{l}}^E, \tilde{C}_{\bar{l}}^B\right) \approx \frac{1}{(2\bar{l}+1)\bar{l}(\bar{l}+1)\Delta l}\frac{C_{\bar{l}}^{E\,2}}{4\pi}\sum_{LM} L(L+1)\left|(w^2)_{LM}\right|^2 = \frac{1}{(2\bar{l}+1)\bar{l}(\bar{l}+1)\Delta l}\frac{C_{\bar{l}}^{E\,2}}{4\pi}\int (\nabla w^2)^2 \, \mathrm{d}\hat{\boldsymbol{n}}.$$

$$\ell, \ell', \Delta\ell \gg \ell_{max}$$

Approximate recovered spectra as weighted sums of pseudo-spectra:

$$\hat{C}_l^E = \frac{1}{N_l}\left(\alpha_l \tilde{C}_l^E - \beta_l \tilde{C}_l^B\right), \qquad \hat{C}_l^B = \frac{1}{N_l}\left(\alpha_l \tilde{C}_l^B - \beta_l \tilde{C}_l^E\right)$$

$$\mathrm{var}\left(\hat{C}_{\bar{l}}^E\right) \approx \frac{2 w^{(4)} C_{\bar{l}}^{E\,2}}{(2\bar{l}+1)\Delta l f_{\rm sky} w^{(2)2}} \qquad \left(C_l^B = 0\right),$$

$$\mathrm{var}\left(\hat{C}_{\bar{l}}^B\right) \approx \frac{2 C_{\bar{l}}^{E\,2}}{(2\bar{l}+1)\Delta l f_{\rm sky}}\frac{6}{\bar{l}^2(\bar{l}+1)^2}\left[\frac{(\nabla w)^{(4)}}{w^{(2)2}} + \frac{2}{3}\frac{w^{(4)}(\nabla w)^{(2)2}}{w^{(2)4}} - \frac{4}{3}\frac{(\nabla w)^{(2)}(w\nabla w)^{(2)}}{w^{(2)3}}\right] \qquad \left(C_l^B = 0\right)$$

$$\alpha_l \equiv \textstyle\sum_{ll'} P_{ll'}$$

$$\beta_l \equiv \textstyle\sum_{l'} M_{ll'}$$

$$N_l \equiv \alpha_l^2 - \beta_l^2$$

$$4\pi w^{(i)} f_{\rm sky} \equiv \int w^i(\hat{\boldsymbol{n}})\, \mathrm{d}\hat{\boldsymbol{n}}$$

Illustration: E-mode leakage in BB covariance

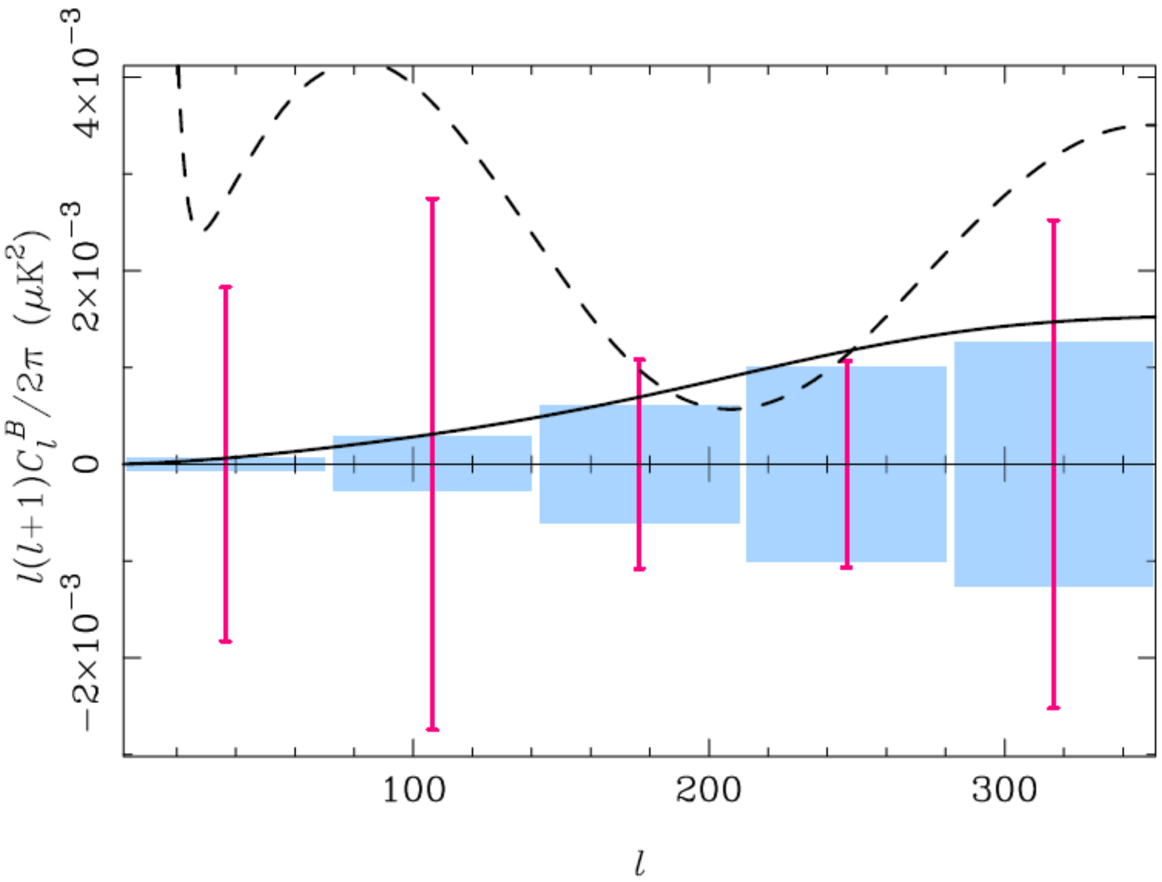
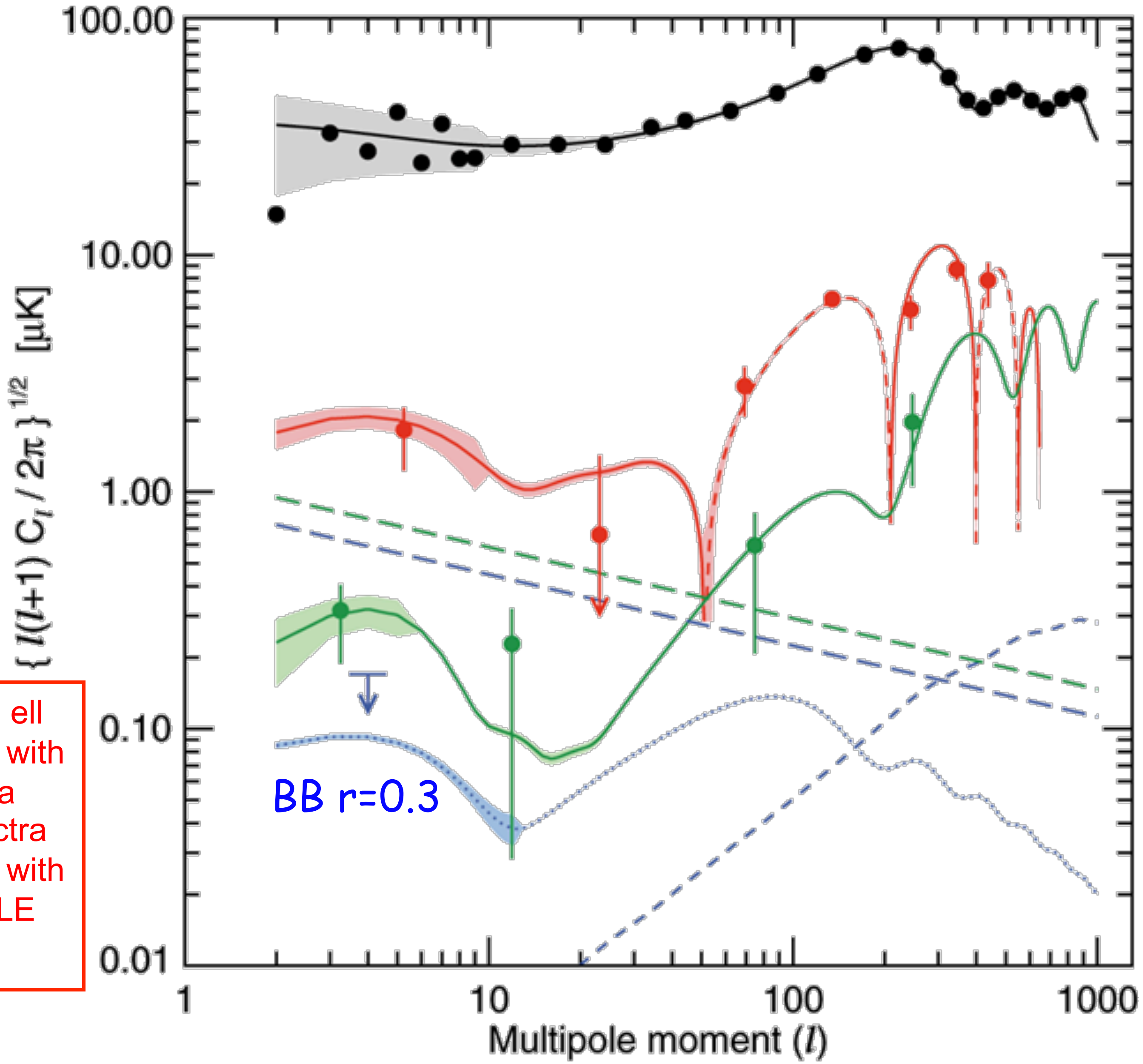


Figure 10. Sample variance errors on the recovered $\hat{C}_l^B$ using the estimator of Chon et al. (2004); see also Section 2.1. The survey area and weight function is the same as in Fig. 1. The error boxes are the contribution purely from C_l^B to the one-sigma errors on flat band-power estimates with $\Delta l = 70$. These are thus representative of the errors that would be obtained if E and B modes were separated (without loss) at the level of the map. They agree well with the simple rule of thumb in equation (98), plotted as the solid line. The error bars are the contribution purely from C_l^E. They agree reasonably with the dashed line, which is the rule of thumb in equation (99). Critically, these dominate the errors due to C_l^B in the angular range relevant to gravity-wave searches with B-mode polarization. All errors are computed in the null hypothesis $r = 0$, so the contribution from C_l^B arises only from sample variance of the lens-induced B modes.

Application: WMAP3 SPECTRA



•TT/TE at high ell were obtained with pseudo spectra
• All other spectra were obtained with pixel-based MLE

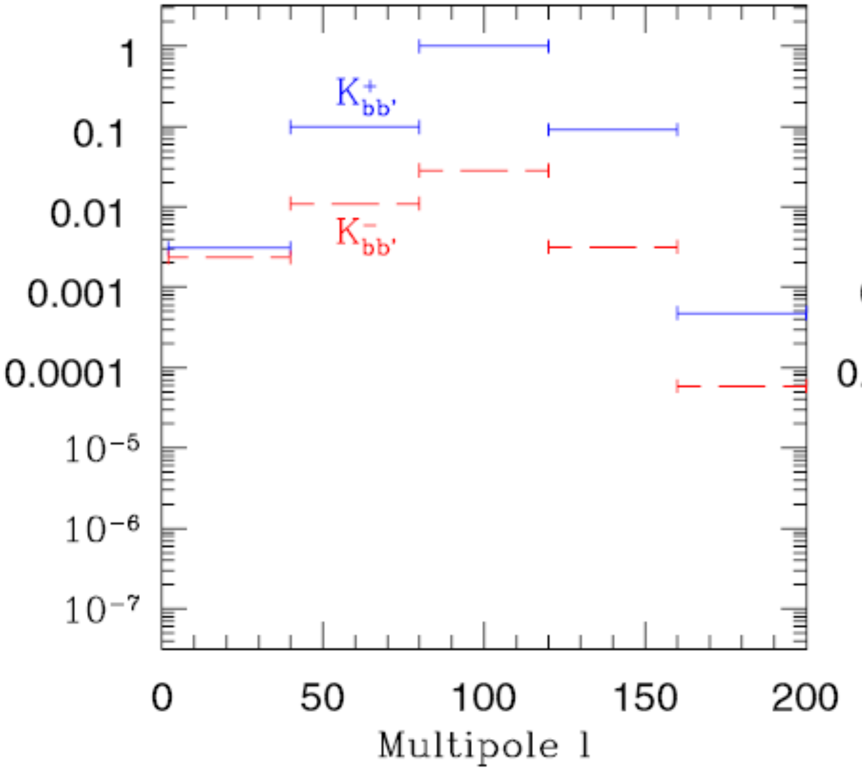
Approx EE/BB foreground

Page et al. 06

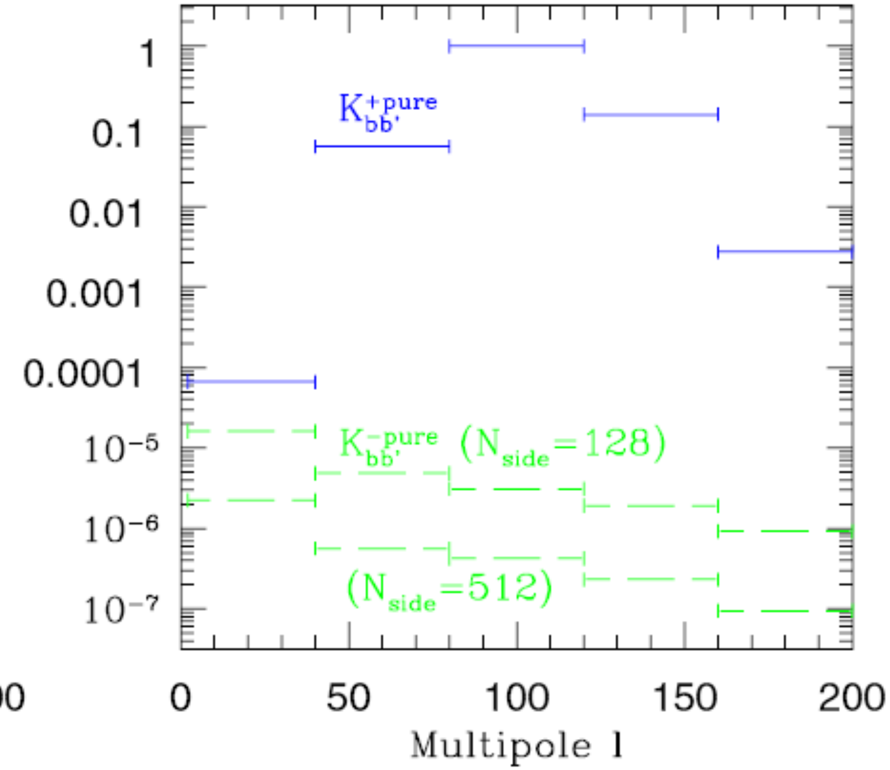
E/B separation “in the map”

- Idea: E/B mode separation in the map, using fast spherical harmonics transforms
- Can be achieved if weight function and its derivatives have specific properties
- BB spectrum covariance is not contaminated by E-mode power

Regular fast estimator



New fast estimator



Bunn et al. 03
Smith 06

E/B separation in the map (cont.)

$$\chi_E = (\eth\eth(Q+iU) + \bar{\eth}\bar{\eth}(Q-iU))/2$$

$$\chi_B = i(\eth\eth(Q+iU) - \bar{\eth}\bar{\eth}(Q-iU))/2$$

$$\tilde{a}^{\chi}_{\ell m} \equiv \frac{1}{\sqrt{(\ell-1)\ell(\ell+1)(\ell+2)}} \int d\mathbf{n}\chi_B(\mathbf{n})W(\mathbf{n})Y^*_{\ell m}(\mathbf{n})$$

Integrate by parts:

$$\tilde{a}^{\chi}_{\ell m} = \frac{i}{2}\int d\mathbf{n}(Q(\mathbf{n}+iU(\mathbf{n}))\eth\eth\frac{W(x)Y^*_{\ell m}(\mathbf{n})}{\sqrt{(\ell-1)\ell(\ell+1)(\ell+2)}} + c.c.$$

$$= \frac{i}{2}\int d\mathbf{n}(Q(\mathbf{n})+iU(\mathbf{n}))[W(\mathbf{n})\,{}_2Y^*_{\ell m}$$

$$+ \frac{2}{\sqrt{(\ell-1)(\ell+2)}}\bar{\eth}W(\mathrm{n})\,{}_1Y^*_{\ell m}(\mathbf{n})$$

$$+ \frac{1}{\sqrt{(\ell-1)\ell(\ell+1)(\ell+2)}}\bar{\eth}\bar{\eth}W(\mathrm{n})Y^*_{\ell m}(\mathbf{n})]$$

Usual pseudo-Cl

Contreterms

Smith & Zaldarriaga 07

W window function of finite support:

$W = 0$ and $\eth W = 0$ on survey boundary

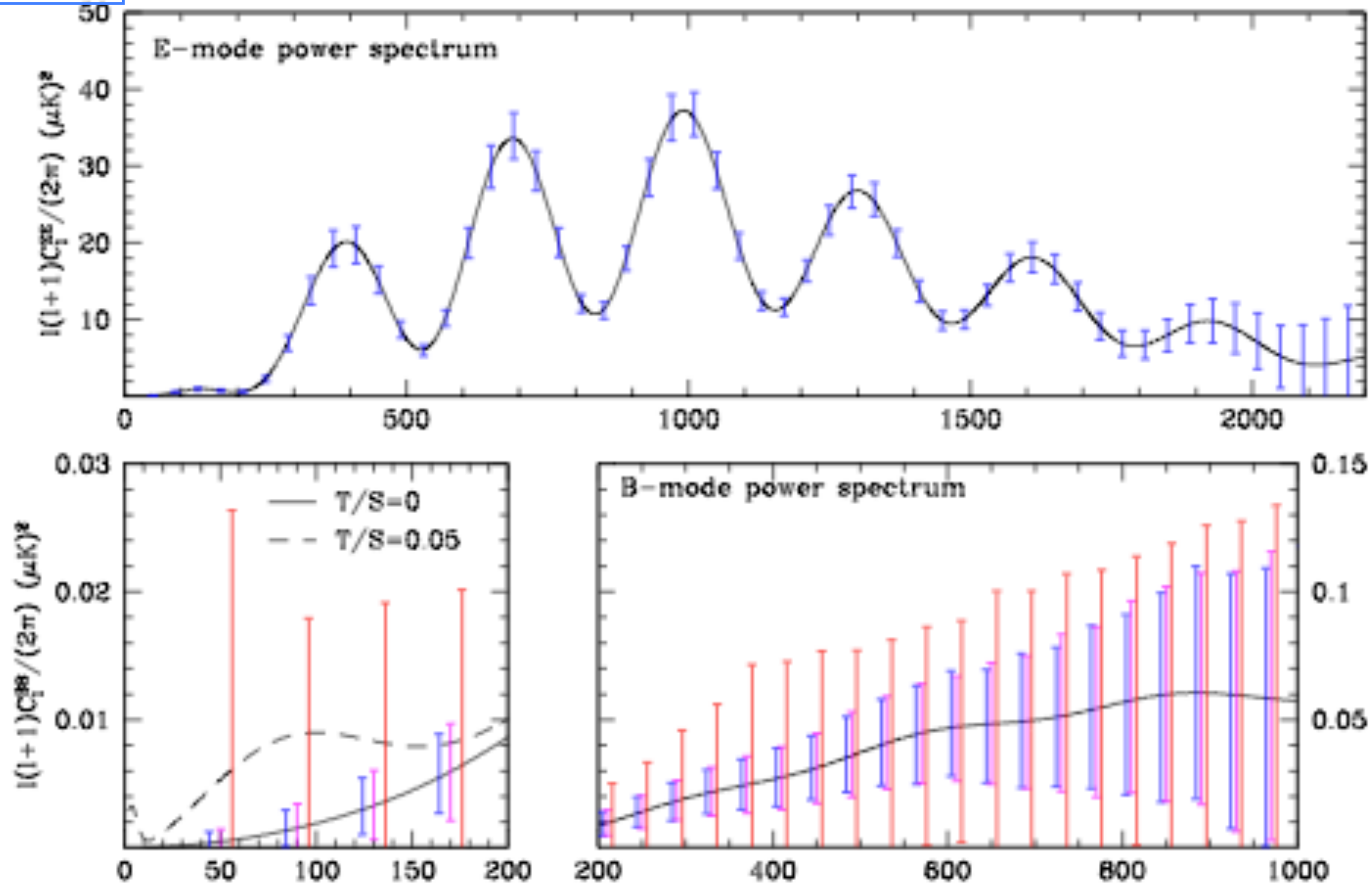
- Needs to be implemented on discretized data: remaining pixelization issues
- Difficulty Is in (numerical) design of the window: make estimator close to QML

Effect on covariance

- 10 deg. Circular patch
- 5.64 μK.arcmin
- 8' fwhm beam

Smith & Zaldarriaga 07

Pseudo-Cl
with contreterms



Difficulty is in (numerically) optimizing windows to get minimum covariance

Spectral estimation in CMB observations (reminders)

Simple data model: gaussian fields on a pixelized sphere
- scalar (temperature fluctuations)
- vector (T and polarization)

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

$$\begin{aligned} \mathbf{C}(p) &= \mathbf{S}(p) + \mathbf{N} \\ \mathbf{S}(p)_{ij} &= \sum_{\ell=0}^{\infty} (2\ell+1) C_\ell P_\ell(\cos\theta_{ij}) \end{aligned}$$

Temperature only: scalar case

Full sky case, no noise:

$$\mathbf{S}_{\ell m,\ell' m'} = \delta_{\ell\ell'}\delta_{mm'}C_\ell$$

$$P(a_{\ell m}|C_\ell) = \frac{1}{\sqrt{2\pi C_\ell}} \exp\left[-\frac{|a_{\ell m}|^2}{2C_\ell}\right]$$

$$\widehat{C}_\ell = \sum_{m=-\ell}^{\ell} |a_{\ell m}|^2$$

Sufficient statistics

Full-sky likelihoods

$$P(\hat{C}_\ell|C_\ell) \propto C_\ell \left(\frac{\hat{C}_\ell}{C_\ell}\right)^{\frac{2\ell-1}{2}} \exp\left[-\frac{(2\ell+1)\hat{C}_\ell}{2C_\ell}\right]$$

Chi2 distribution

Flat prior

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

Inverse Gamma distribution

Generalizes easily to vector-valued, pixelized fields
E.g. Temperature and polarisation anisotropies

$$\mathbf{a} = \begin{pmatrix} a^T_{\ell m} \\ a^E_{\ell m} \\ a^B_{\ell m} \end{pmatrix}, \quad \mathbf{V}_\ell = \begin{pmatrix} C^{TT}_\ell & C^{TE}_\ell & 0 \\ C^{TE}_\ell & C^{EE}_\ell & 0 \\ 0 & 0 & C^{BB}_\ell \end{pmatrix}$$

$$\hat{\mathbf{C}}_\ell = \frac{1}{2\ell+1}\sum_m \mathbf{a}\mathbf{a}^T = \begin{pmatrix} \hat{C}^{TT}_\ell & \hat{C}^{TE}_\ell & \hat{C}^{TB}_\ell \\ \hat{C}^{TE}_\ell & \hat{C}^{EE}_\ell & \hat{C}^{EB}_\ell \\ \hat{C}^{TB}_\ell & \hat{C}^{EB}_\ell & \hat{C}^{BB}_\ell \end{pmatrix}$$

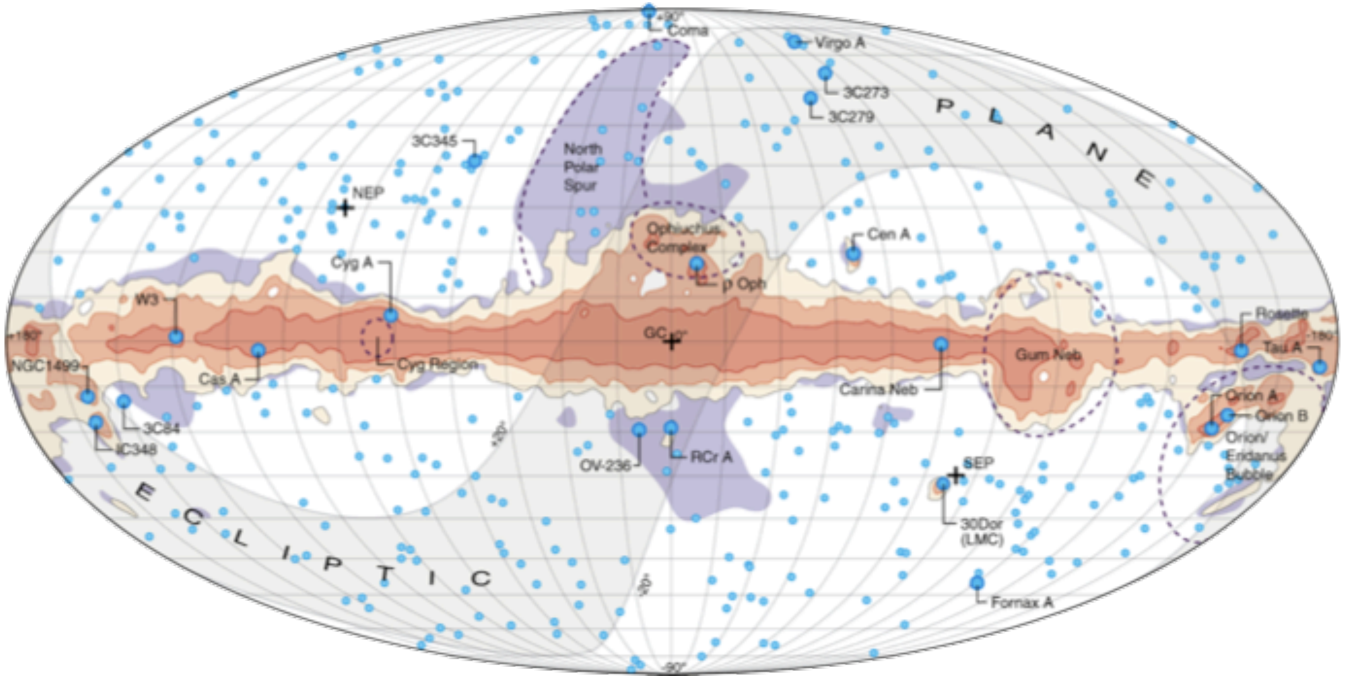
Wishart distribution

$$\begin{aligned} P(\hat{\mathbf{C}}_\ell|\mathbf{W}_\ell) &\propto |\hat{\mathbf{C}}_\ell|^{(2\ell-3)/2}|\mathbf{W}_\ell|^{-(2\ell+1)/2}\exp\left[-\frac{1}{2}Tr(\mathbf{W}_\ell^{-1}\mathbf{S}_\ell)\right] \\ \mathbf{W}_\ell &= \mathbf{V}_\ell/(2\ell+1) \end{aligned}$$

Generalizes easily to isotropic noise + beam:

$$-2\ln P(C_\ell|\hat{C}_\ell) = (2\ell+1)\left(\frac{\hat{C}_\ell}{C_\ell B_\ell^2 + N_\ell} + \ln(C_\ell B_\ell^2 + N_\ell)\right) + \text{cst}$$

Realistic case: anisotropic noise, partial coverage…



In case of anisotropic noise / partial coverage, the factorisation of the posterior (based on spherical harmonics orthogonality) is lost…
Empirical power spectrum is not a sufficient statistics anymore…

Possible path:
- different approach for large and small scales (resp. low and high multipoles)
- find parametrized, more or less sophisticated approximations of the (joint) posterior at small scales
- sample the exact posterior at large scales where approximations fail

Beware: this is still a simplified data model, reality involves asymmetric beams, noise filtering in time domain, etc.

Strategy at high multipoles

- Partial sky coverage: no product form for the posterior
- Huge number of pixels: cannot efficiently explore the posterior
- but gaussian asymptotics should help to construct approximations
- Partial sky coverage: no exactly sufficient statistics in terms of empirical power spectra, but very good approximations, especially in noise dominated regime
- data compression in terms of heuristically weighted/corrected empirical power spectrum
- construction of approximate, analytical posterior functions inspired from the full-sky case, and (empirical) covariance of the (empirical) PS

Parametrized approximations using pseudo spectra

Start from the full-sky distribution of the empirical power spectrum:

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

Assume we have an unbiased empirical power spectrum estimate, with known covariance, we look for a function of this estimate which is approximately gaussian

$$\ln P(\hat{C}_\ell|\theta) \approx \ln A - \frac{1}{2}\sum_{\ell\ell'} M^{-1}_{\ell\ell'}(\hat{x}_\ell(\hat{C}_\ell)-\mu_\ell)(\hat{x}_{\ell'}(\hat{C}_{\ell'})-\mu_{\ell'})$$

Choice: (approximately) put the third derivative of lnP to zero in the full-sky case:

$$2(\mathrm{d}\hat{C}_\ell/\mathrm{d}\hat{x}_\ell)^2 = 3\hat{C}_\ell \mathrm{d}^2\hat{C}_\ell/\mathrm{d}\hat{x}_\ell^2 \Rightarrow \hat{x}_\ell = \hat{C}_\ell^{1/3}$$

Further inspection of the full-sky distribution (peak and curvature at peak) leads to this choice:

$$\mu_\ell = \left(\frac{2\ell-1}{2\ell+1}C_\ell\right)^{1/3}$$

$$M^{-1}_{\ell\ell'} = \sqrt{\frac{2\ell+1}{2\ell-1}}\frac{\mathrm{d}\hat{C}_\ell}{\mathrm{d}\hat{x}_\ell} S^{-1}_{\ell\ell'} \sqrt{\frac{2\ell'+1}{2\ell'-1}}\frac{\mathrm{d}\hat{C}_{\ell'}}{\mathrm{d}\hat{x}_{\ell'}}$$

and

$$A^{-1} \propto \sqrt{|M_{\ell\ell'}|}\prod_\ell \mu_\ell^2$$

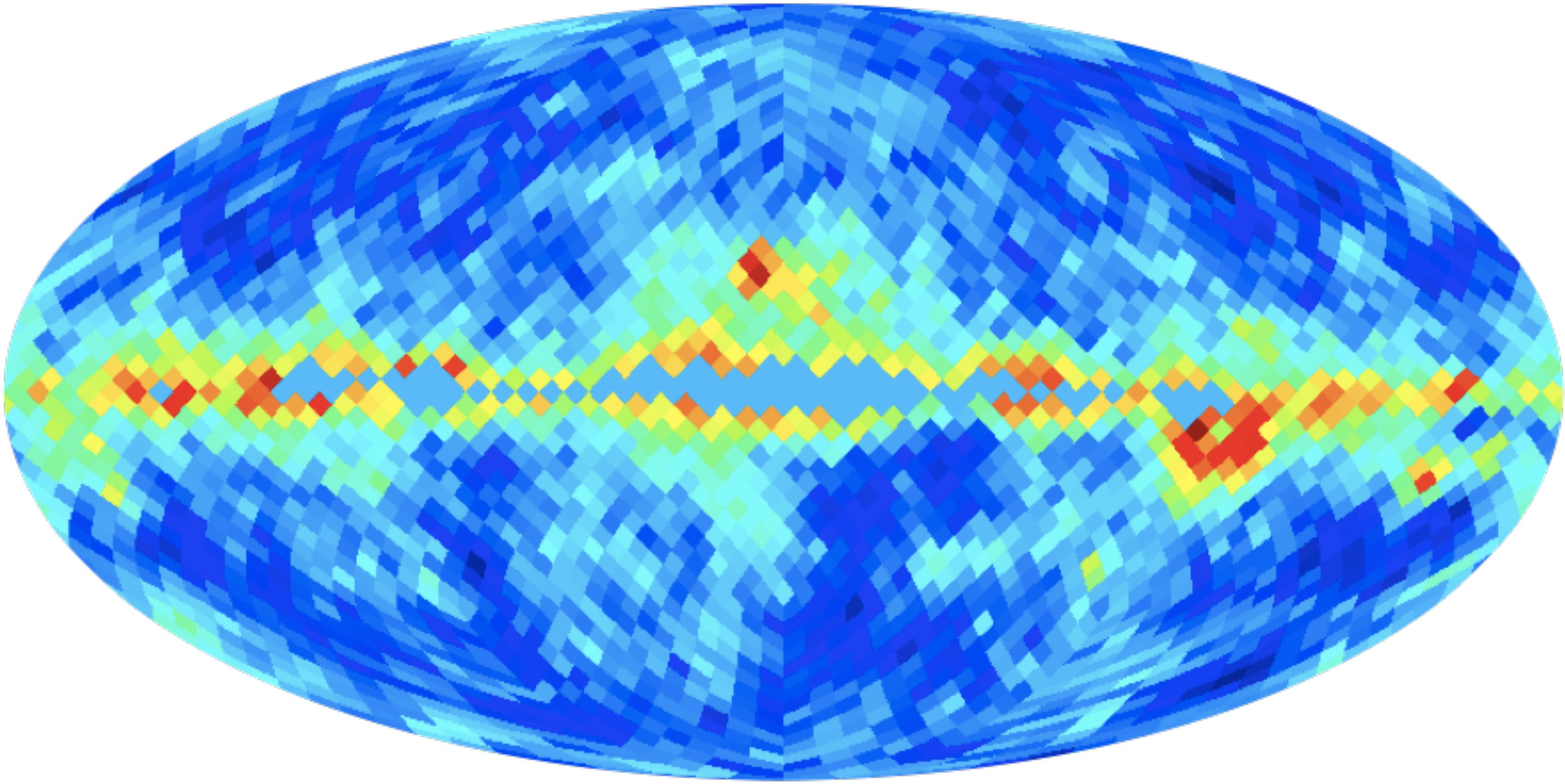
Other possible choices: offset log-normal approximation, (direct) gaussian approximation, inverse-gamma approximation, etc.

- Beware: in general pseudo-spectra are not sufficient statistics
- This is still an open problem, especially for polarization

Smith et al. 06
Percival & Brown 06
Hamimeche & Lewis 08

Strategy at low multipoles

- Small number of observable modes/pixels
- Partial sky coverage: simple approximations to the posterior fail
- Need to sample from the “exact” posterior (or build clever approximations)



Exploring the full likelihood: Gibbs samplers

IDEA: factorize (complicated) likelihood into (simpler) conditional probabilities

Goal: sampling of the posterior:

$$P(C_\ell|m) \propto G[m, S(C_\ell) + N]P(C_\ell)$$

Can be achieved by marginalizing:

$$P(C_\ell, s, m) = P(m|s)P(s|C_\ell)P(C_\ell)$$

How to do that ? Iterative sampling from the conditional probabilities:

$$s^{i+1} \hookleftarrow P(s|C_\ell^i, m)$$

$$C_\ell^{i+1} \hookleftarrow P(C_\ell|s^{i+1}).$$

Wiener filter

$$P(s|C_\ell^i, m) \propto G[s - S^i(S^i + N)^{-1}m, ((S^i)^{-1} + N^{-1})],$$

$$P(C_\ell|s^i) = P(C_\ell)\prod_l \frac{\sigma_\ell^{2\ell-1}}{2^{2\ell-1}\Gamma[2\ell-1]} \frac{e^{-\sigma_\ell/(2C_\ell)}}{\sqrt{C_\ell^{2\ell+1}}}$$

$$\sigma_\ell = \sum_{m=-\ell}^{+\ell} |s_{\ell m}^i|^2$$

Wandelt et al. 2004

Gibbs sampling (cont.)

Drawing from $P(s|C_\ell, m)$

$$[1 + (S^i)^{1/2} N^{-1} (S^i)^{1/2}](S^i)^{-1/2} x = (S^i)^{1/2} N^{-1} m.$$

Wiener filter map, computed Using CG iterations

$$(1 + S^{1/2} N^{-1} S^{1/2}) S^{-1/2} y = \xi + S^{1/2} N^{-1/2} \chi$$

Fluctuation part, ξ and χ are normal random variables

Drawing from $P(C_\ell|s, m) = P(C_\ell|s)$

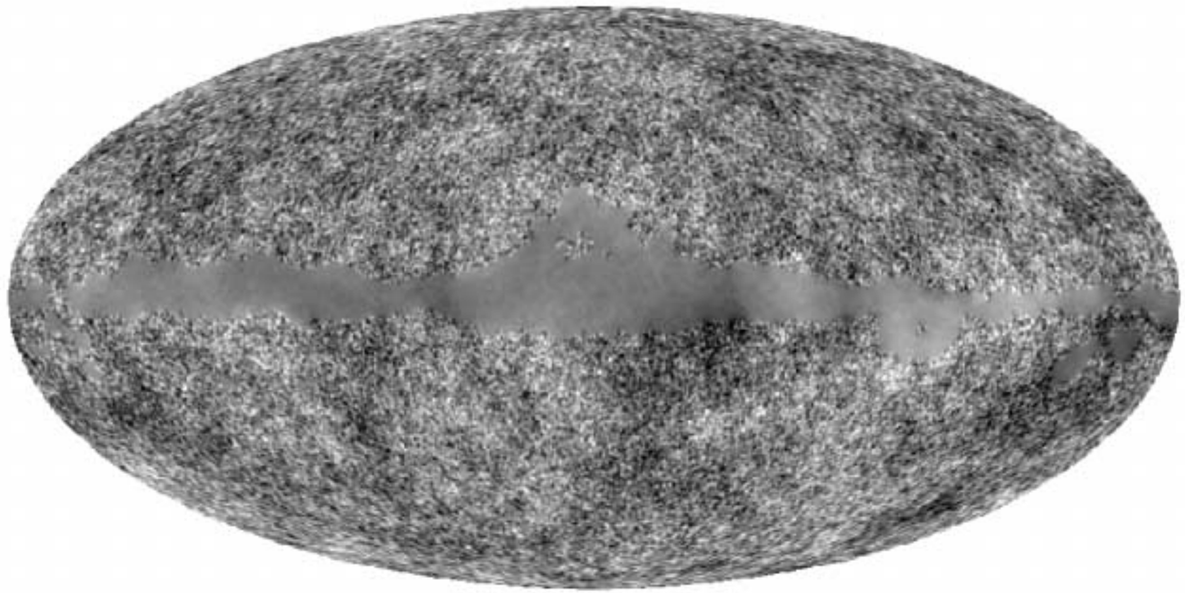
Compute $\sigma_\ell = \sum_{m=-\ell}^{\ell} s_{\ell m}^2$

ρ_ℓ Vector of normal random variables, of size $(2\ell - 1)$

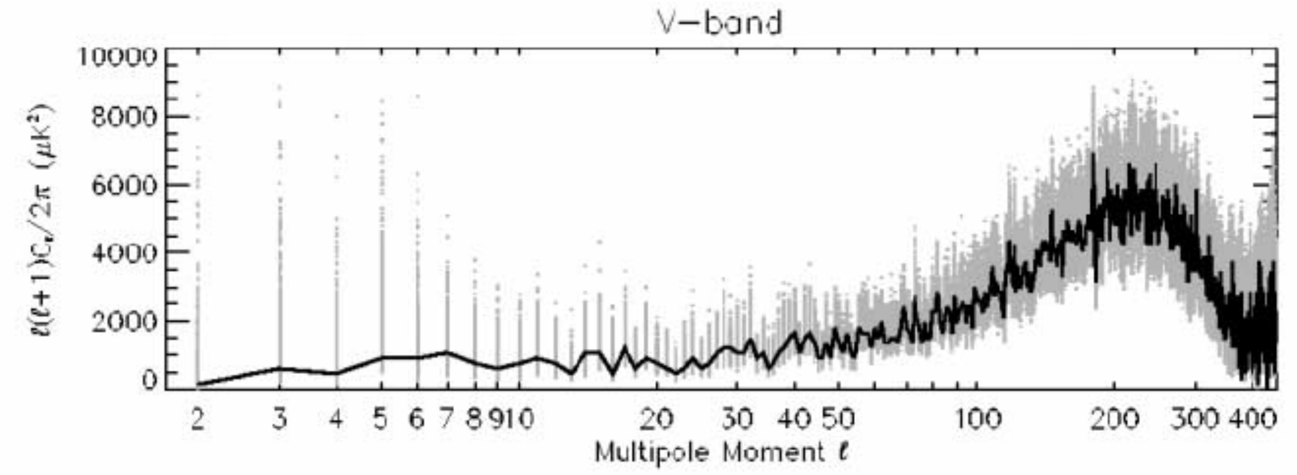
$$C_\ell = \frac{\sigma_\ell}{|\rho_\ell|^2}$$

Or replace this sampling step with analytical knowledge of this conditional distribution ➔ Blackwell-Rao estimator (leads to reduced variance on integrals of the posterior, see later)

Results on WMAP1 data



Map average over posterior: Generalized wiener filter



V-band spectrum: Samples and mean posterior

Eriksen et al. (2004)

Main issue: slow convergence of sampler if S/N << 1

Gibbs sampling and Blackwell-Rao

We don't know want to sample directly from $P(C_\ell|\mathrm{d})$

But we know how to sample from the conditionals:

Product of inverse Gamma (Wishart) distribution, analytic $P(\mathrm{C}|\mathrm{s},\mathrm{d})$

Constrained gaussian realisation, iterative solvers… $P(\mathrm{s}|\mathrm{C},\mathrm{d})$

Draw from each conditional iteratively: Gibbs sampling

$$P(\mathrm{C}|\mathrm{s},\mathrm{d}) = P(\mathrm{C}|\mathrm{s}) = \prod_\ell P(C_\ell|\widehat{C}_l(\mathrm{s}))$$

But we know this analytically → Blackwell-Rao estimator:

$$\begin{aligned} P(\mathrm{C}|\mathrm{d}) &= \int \left(\prod_\ell P(C_\ell|\mathrm{s})\right) P(\mathrm{s}|\mathrm{d})\mathrm{ds} \\ &= \int \left(\prod_\ell P(C_\ell|\widehat{C}_\ell(\mathrm{s}))\right) P(\hat{\mathrm{C}}|\mathrm{d})\mathrm{d}\hat{\mathrm{C}} \\ &\approx \langle \prod_\ell P(C_\ell|\widehat{C}_\ell^{(i)})\rangle_{Gibbs} \end{aligned}$$

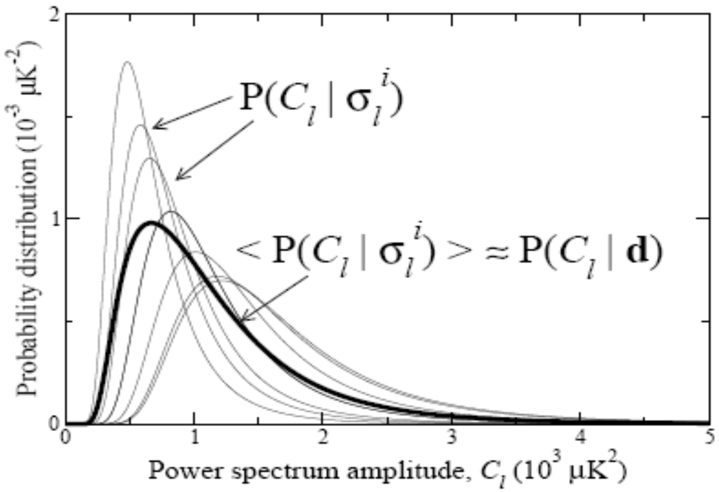


FIG. 1: A one-dimensional illustration of the BR estimator. The thin lines indicate the $P(C_\ell|\sigma_\ell^i)$ distributions, and the thick line shows their average. This average converges toward the true density $P(C_\ell|\mathrm{d})$ as the number of samples increases.

Chu et al. 05

Comparison of Gibbs results to (one choice of) analytical, parameterized posteriors

Comparison of Gibbs results to analytical (WMAP) posterior

Validation of the Gibbs sampler for a 2 parameters case

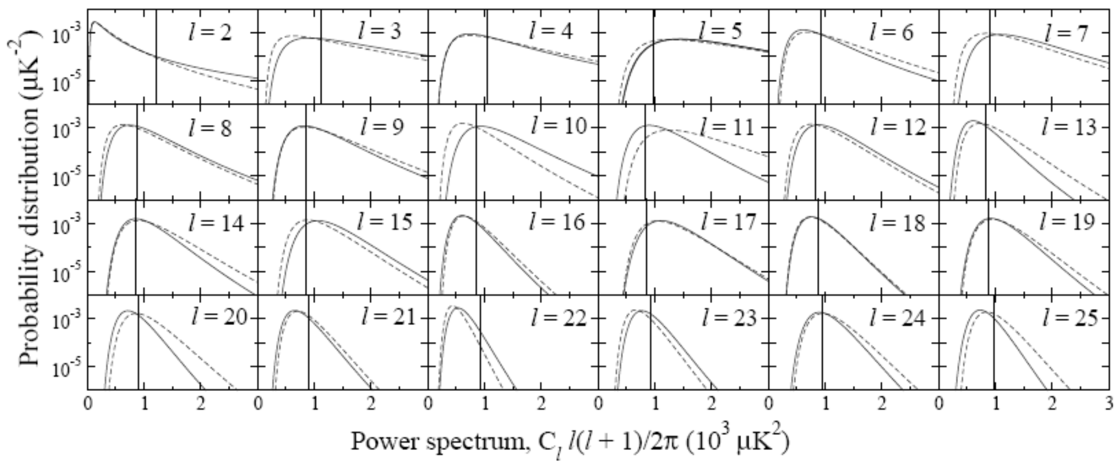


FIG. 6: Comparison of the BR (solid curve) and the analytic *WMAP* (dashed curve) univariate likelihood functions for each multipole up to $\ell = 25$. The vertical lines indicates the value of the best-fit *WMAP* power-law model (not including a running spectral index). The univariate likelihood functions are computed by slicing through the multivariate likelihood, fixing all other multipoles at the corresponding best-fit value. Notice that all distributions shown here are strongly non-Gaussian.

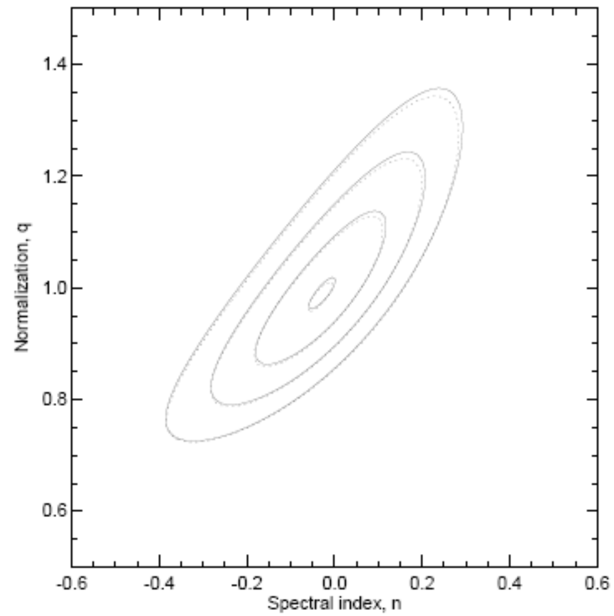


FIG. 2: Contours in (q, n) space of constant probability given the simulated data described in text, for both the BR estimator (solid) and brute-force evaluation of the likelihood (dashed). Contours are where $-2 \ln P(C_\ell|\mathbf{d})$ rises by 0.1, 2.3, 6.17, and 11.8 from its minimum value, corresponding (for Gaussian distributions) to the peak, and the 1, 2 and 3σ confidence regions.

Chu et al. 05

Plus and minuses…

- As any MC sampling, can easily refine the posterior (e.g. foreground templates)
- It can "deal" with intermediate scales, where "exact" computation of the posterior is impossible (too many pixels/modes)
- What about convergence ???
  - Very bad behavior at low Signal-to-Noise ratio
    - need to rebin at small scales
  - Fortunately, at large scales Signal is really dominant (at least for temperature)
  - Lack of parallelism
  - Recent developments to overcome convergence problems: Jewell et al. 08

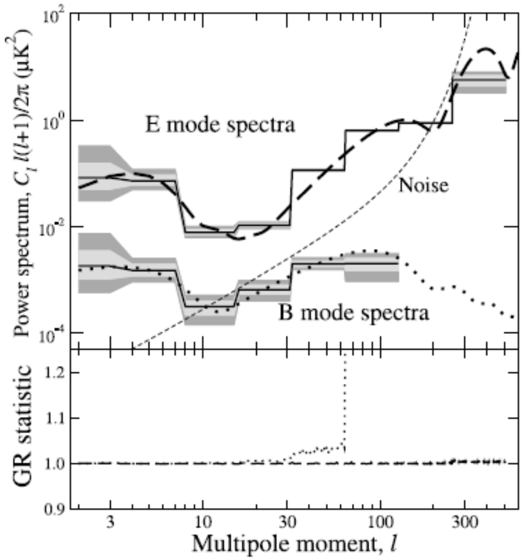


FIG. 3.—Reconstructed E- and B-mode power spectra from the low-resolution analysis. Input spectra are shown as dashed and dotted lines, respectively, while the reconstructed posterior distributions are indicated by solid lines (posterior maximum) and gray regions (1 and 2 σ confidence regions). The corresponding noise spectrum is given by a thin dashed line. The Gelman-Rubin convergence statistic as a function of multipole is shown in the bottom frame.

Low resolution, high S/N prospective experiment

Need to bin (logarithmically) to get reasonable convergence (5 chains of 1000 samples, 10 min/sample)

Larson et al. 07

Same for high resolution, higher noise experiment (e.g. Planck)

- Heavily binned (20 bins only)
- Very small number of samples (8 chains of 100 samples)
- Each constrained realisation takes ~16 CPU hours
- Even with this binning, convergence indicators are not very good (not surprising given the number of samples, and the parameter space of the chain)

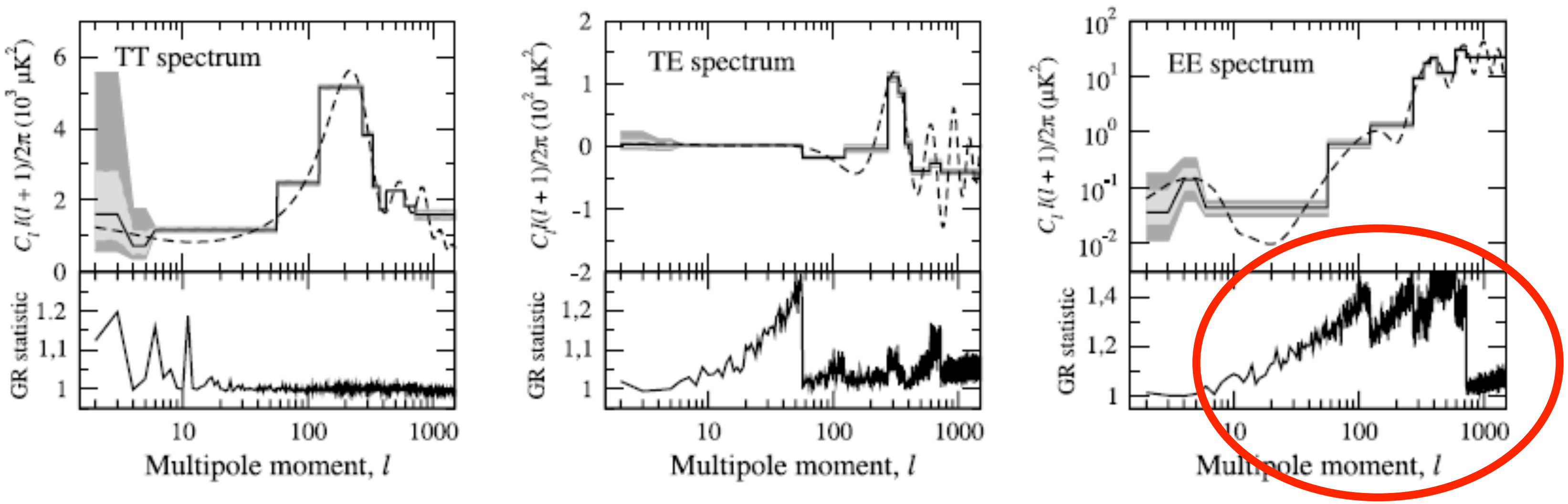


FIG. 5.—Reconstructed power spectra from the high-resolution *Planck* 100 GHz simulation. The true spectra are shown as dashed lines, and the reconstructed posterior distributions are given by a maximum posterior value (*solid lines*) and 68% and 95% confidence regions. The Gelman-Rubin convergence statistics are shown in the bottom frames.

Larson et al. 07

A proposed alternative: Hybrid/Hamiltonian MC

Data model

$$d = \mathbf{RY}a + n$$

$$C_{\ell m \ell' m'} = \langle a_{\ell m} a^*_{\ell' m'} \rangle = C_\ell \delta_{\ell\ell'} \delta_{mm'}.$$

Decomposition of joint posterior decomposition (as in Gibbs)

$$\Pr(C_\ell, a|d) \propto \Pr(d|a)\Pr(a|C_\ell)\Pr(C_\ell).$$

$$\Pr(d|a) \propto \exp\left[-\frac{1}{2}(d - \mathbf{RY}a)^{\mathrm{T}} \mathbf{N}^{-1} (d - \mathbf{RY}a)\right],$$

where $\mathbf{N} = \langle nn^{\mathrm{T}} \rangle$, and

$$\Pr(a|C_\ell) \propto \frac{1}{\sqrt{|\mathbf{C}|}} \exp\left(-\frac{1}{2} a^{\mathrm{T}} \mathbf{C}^{-1} a\right)$$

$$\Pr(a|C_\ell) \propto \prod_l \left(\frac{1}{C_\ell}\right)^{\frac{2\ell+1}{2}} \exp\left(-\frac{2\ell+1}{2}\frac{\sigma_\ell}{C_\ell}\right)$$

HMC Implementation

Augmented log-posterior, with kinetic term

$$H = \sum_i \frac{p_i^2}{2m_i} + \psi(x)$$

$$\psi(x) = -\log \Pr(x)$$

Draw moments p_i from Gaussian law
Integrate Hamilton's equation (e.g. leap-frog)
Estimate acceptance probability

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = \frac{\partial H}{\partial p_i}$$

$$\frac{\mathrm{d}p_i}{\mathrm{d}t} = -\frac{\partial H}{\partial x_i} = -\frac{\partial \psi(x)}{\partial x_i}$$

$$p_A = \min(1, \exp(-\delta H))$$

Useful non-trivial gradients

$$\psi(a, C_\ell) = \frac{1}{2}(d - \mathbf{YB}a)^{\mathrm{T}} \mathbf{N}^{-1} (d - \mathbf{YB}a) + \sum_\ell \left(\ell + \frac{1}{2}\right)\left(\ln C_\ell + \frac{\sigma_\ell}{C_\ell}\right) + \text{const}$$

and the gradient of the potential can be computed exactly by

$$\frac{\partial \psi(a, C_\ell)}{\partial a} = -\mathbf{BY}^{\mathrm{T}}\mathbf{N}^{-1}(d - \mathbf{YB}a) + \mathbf{C}^{-1}a$$

$$\frac{\partial \psi(a, C_\ell)}{\partial C_\ell} = \left(l + \frac{1}{2}\right)\frac{1}{C_\ell}\left(1 - \frac{\sigma_\ell}{C_\ell}\right).$$

Fast because of fast Spherical harmonics transforms

Fast because scalar

Results on WMAP simulation

Starting point: Constrained realization

$$\left(\mathsf{C}^{-1} + \mathsf{B}\mathsf{Y}^{\mathrm{T}}\mathsf{N}^{-1}\mathsf{Y}\mathsf{B}\right) x = \mathsf{B}\mathsf{Y}^{\mathrm{T}}\mathsf{N}^{-1} d$$

and a fluctuation term y that corrects for the bias in x

$$\left(\mathsf{C}^{-1} + \mathsf{B}\mathsf{Y}^{\mathrm{T}}\mathsf{N}^{-1}\mathsf{Y}\mathsf{B}\right) y = \mathsf{C}^{-1/2}\omega_0 + \mathsf{B}\mathsf{Y}^{\mathrm{T}}\mathsf{N}^{-1/2}\omega_1$$

Diagonal mass matrix, inversely proportional to approx. variances of parameters

$$\mathrm{var}\left(C_\ell\right) = \frac{2\ell+1}{2}\left(C_\ell + N_\ell/B_\ell^2\right)^2$$

$$\mathrm{var}\left(a_{\ell m}\right) = \left(C_\ell^{-1} + B_\ell N^{-1}_{\ell m,\ell' m'}\delta_{\ell\ell'}\delta_{mm'}B_\ell\right)^{-1}$$

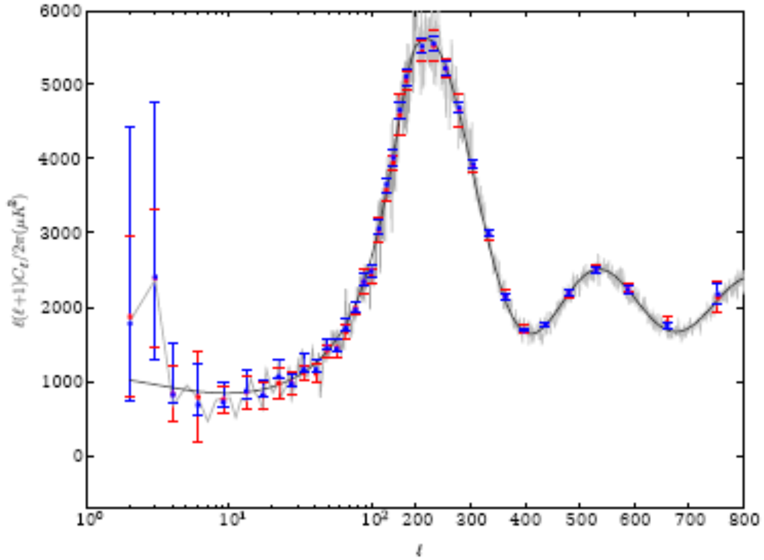


Figure 1. Binned power spectrum and 68 percent confidence intervals as compared to the results of an application of the MASTER method to the same simulated *WMAP* data. The black solid line shows the power spectrum from which the simulation was generated while the grey shows the power spectrum of the realisation. The red points and error bars show the MASTER results. The blue circles and error bars show the mean and 68 per cent confidence intervals found from 4050 samples generated with the HMC sampler.

Hanson's convergence criterion

$$R_i = \frac{\sigma_i^2 = \int_{-\infty}^{\infty} (x_i - \bar{x}_i)^2 \Pr(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \approx \frac{1}{M}\sum_k \left(x_i^k - \bar{x}_i\right)^2}{\sigma_i^2 = \frac{1}{3}\int_{-\infty}^{\infty} (x_i - \bar{x}_i)^3 \frac{\partial \psi(\boldsymbol{x})}{\partial x_i} \Pr(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.}$$

We compute (21) from the samples in our chain by

$$\sigma_i^2 \approx \frac{1}{M}\frac{1}{3}\sum_k \left(x_i^k - \bar{x}\right)^3 \left.\frac{\partial \psi}{\partial x_i}\right|_{x_i^k}.$$

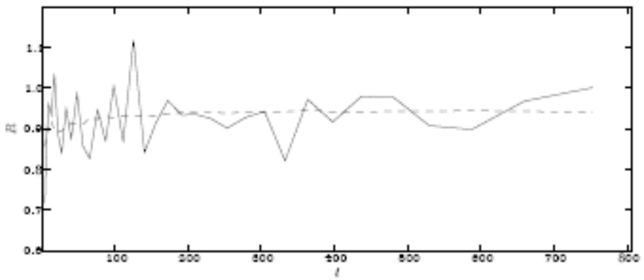


Figure 2. A summary of the convergence statistics of the 4050 samples used to produce the power spectrum in Fig. 1. Although convergence is judged from the R for every parameter we show here only the average R for in each bin for the C_ℓ (solid line) and $\boldsymbol{\alpha}$ (dashed line). R values between 0.8 and 1.2 represent excellent convergence while any that lie between 0.6 and 1.4 are acceptable.

3000 "burn-in" discarded samples
4000 used samples

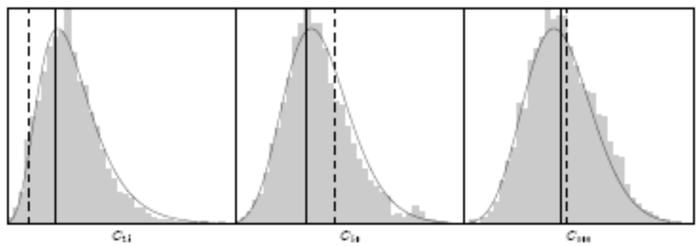


Figure 3. Posterior distributions for C_{15}, C_{50} and C_{200}. The histograms are generated from the C_ℓ samples and the solid line shows the results of the Rao-Blackwell estimator. The solid vertical line marks the value of the input theoretical spectrum whilst the dashed marks the value for the realisation used in the simulations.

TEASING : analytic approximations at large scales

Pixel-based likelihood (=> target)

$$p(\boldsymbol{x}|\mathsf{R}) = |2\pi\mathsf{R}|^{-1/2} \mathrm{e}^{-\frac{1}{2}\boldsymbol{x}^T \mathsf{R}^{-1} \boldsymbol{x}}$$

$$\mathbf{R}_{ij}^{\mathrm{var}}(\boldsymbol{D}) = \sum_{\ell=\ell_{\min}}^{\ell=\ell_{\max}} \frac{2\ell+1}{4\pi} (D_\ell - N_\ell) P_\ell(\cos\theta_{ij}),$$

$$\mathbf{R}_{ij}^{\mathrm{cst}} = \sum_{\ell \mathrm{fixed}} \frac{2\ell+1}{4\pi} W_\ell C_\ell P_\ell(\cos\theta_{ij}) + \sigma^2 \delta_{ij}.$$

$$D_\ell = W_\ell C_\ell + N_\ell$$

Sampling proposal

$$g(\boldsymbol{D};\theta) = \prod_{\ell=\ell_{\min}}^{\ell=\ell_{\max}} i\Gamma(D_\ell; \alpha_\ell, \beta_\ell)$$

$$\theta = \{\alpha_\ell, \beta_\ell\}_{\ell=\ell_{\min}}^{\ell=\ell_{\max}}$$

$$\alpha_\ell = \frac{(2\ell+1)}{2} f_{\mathrm{sky}} - 1$$

$$\beta_\ell = \frac{(2\ell+1)}{2} f_{\mathrm{sky}} D_\ell^{\mathrm{ML}}$$

Parameters starting points

Iteratively minimize divergence between the distributions using adaptive importance sampling (Population MC)

Benabed, Cardoso, Prunet, Hivon 2009

Population Monte Carlo basics

Also known as Adaptive Importance Sampling

$$\mathbb{E}_{\pi}^{X}[f(X)] = \int f(x)\pi(x)dx \approx \frac{1}{N}\sum_{X_i \sim \pi} f(X_i)$$

- MC basic sampling
- Needs to know how to sample under π

$$\begin{aligned}
\mathbb{E}_{\pi}^{X}[f(X)] &= \int f(x)\pi(x)dx \\
&= \int f(x)\frac{\pi(x)}{q(x)}q(x)dx \\
&= \int f(x)\omega(x)q(x)dx \\
&\approx \frac{1}{N}\sum_{X_i \sim q} \omega(X_i)f(X_i)
\end{aligned}$$

- Importance sampling
- No need to sample under π
- Need to know how to compute π
- Only need to sample under q
- Beware: q cannot be too far from π
- q and π need to have same support

$$\begin{aligned}
\mathbb{E}_{\pi}^{X} &= \frac{\int f(x)\omega(x)q(x)dx}{\int \omega(x)q(x)dx} \\
&\approx \frac{\sum_{X_i \sim q} \omega(X_i)f(X_i)}{\sum_{X_i \sim q} \omega(X_i)}
\end{aligned}$$

- Normalized importance sampling
- Useful when π is not normalized

Importance sampling: need for adaptation

$$F_n(X) \equiv \frac{1}{N} \sum_{X_i \sim q} \omega(X_i) f(X_i)$$

IS estimate

$$\begin{aligned} \mathbb{E}_q^X[F_n(X)] &= \frac{1}{N} \sum_{i=1}^{N} \int \frac{\pi(x)}{q(x)} f(x) q(x) dx \\ &= \mathbb{E}_\pi^X[f(X)] \end{aligned}$$

No bias

$$\begin{aligned} Var[F_n(X)] &= \mathbb{E}_q^X[(F_n(X) - \mathbb{E}_q^X[F_n(X)])^2] \\ &= \frac{1}{N} Var[\omega(X) f(X)] \\ &= \frac{1}{N} \int \left(\frac{\pi(x) f(x)}{q(x)} - \mathbb{E}_\pi^X[f(X)] \right)^2 q(x) dx \end{aligned}$$

minimized when $q(x) \propto \pi(x) f(x)$

PMC base algorithm

IS under parametrized mixture proposal:

$$q_{(\alpha,\theta)}(x) = \sum_{d=1}^{D} \alpha_d q_d(x;\theta_d)$$

Iterative minimization of KL divergence:

$$D(\pi \| q_{(\alpha,\theta)}) = \int \log\left(\frac{\pi(x)}{\sum_{d=1}^{D} \alpha_d q_d(x;\theta_d)}\right)\pi(x)\mathrm{d}x$$

1. Generate a sample $(X_{i,t})$ from the current mixture IS proposal (3) parameterised by and compute the normalised importance weights

$$\bar{\omega}_{i,t} = \frac{\pi(X_{i,t})}{\sum_{d=1}^{D} \alpha_d^{t,N} q_d(X_{i,t};\theta_d^{t,N})} \Bigg/ \sum_{j=1}^{N} \frac{\pi(X_{j,t})}{\sum_{d=1}^{D} \alpha_d^{t,N} q_d(X_{j,t};\theta_d^{t,N})}$$

   and the mixture posterior probabilities

$$\rho_d(X_{i,t};\alpha^{t,N},\theta^{t,N}) = \alpha_d^{t,N} q_d(X_{i,t};\theta_d^{t,N}) \Bigg/ \sum_{\ell=1}^{D} \alpha_\ell^{t,N} q_\ell(X_{i,t};\theta_\ell^{t,N}),$$

   for $i = 1,\ldots,N$ and $d = 1,\ldots,D$.

2. Update the parameters α and θ as

$$\alpha_d^{t+1,N} = \sum_{i=1}^{N} \bar{\omega}_{i,t}\rho_d\left(X_{i,t};\alpha^{t,N},\theta^{t,N}\right),$$

$$\theta_d^{t+1,N} = \arg\max_{\theta_d}\left[\sum_{i=1}^{N} \bar{\omega}_{i,t}\rho_d\left(X_{i,t};\alpha^{t,N},\theta^{t,N}\right)\log\left\{q_d\left(X_{i,t};\theta_d^{t,N}\right)\right\}\right]$$

   for $d = 1,\ldots,D$.

Monitor convergence with Perplexity

$$0 \le P = \exp(H^{t,N})/N \le 1$$

$$H^{t,N} = -\sum_{i=1}^{N} \bar{\omega}_{i,t} \log \bar{\omega}_{i,t}$$

Cappe et al. 07

PMC algorithm: derivation

Goal: maximize $\int \log\left(\sum_{d=1}^{D} \alpha_d q_d(x;\theta_d)\right) \pi(x)\,\mathrm{d}x$

As in EM algorithm, use mixture index as latent variable: $f(z) = \alpha_z \quad \text{and} \quad f(x|z) = q_z(x;\theta_z)$

Update as in EM, with additional expectation on X $(\alpha^{t+1}, \theta^{t+1}) = \arg\max_{(\alpha,\theta)} \mathbb{E}_\pi^X \left[\mathbb{E}^Z_{(\alpha^t,\theta^t)} \{\log(\alpha_Z q_Z(X;\theta_Z))|X\}\right]$

Using: $f(z|x) = \alpha_z^t q_z(x;\theta_z^t) \Big/ \sum_{d=1}^{D} \alpha_d^t q_d(x;\theta_d^t)$

And defining: $\rho_d(X;\alpha,\theta) = \alpha_d q_d(X;\theta_d) \Big/ \sum_{\ell=1}^{D} \alpha_\ell q_\ell(X;\theta_\ell)$

$$\alpha^{t+1} = \arg\max_{\alpha} \mathbb{E}_\pi^X \left[\sum_{d=1}^{D} \rho_d(X;\alpha^t,\theta^t)\log(\alpha_d)\right],$$

$$\theta^{t+1} = \arg\max_{\theta} \mathbb{E}_\pi^X \left[\sum_{d=1}^{D} \rho_d(X;\alpha^t,\theta^t)\log(q_d(X;\theta_d))\right]$$

$\sum_{d=1}^{D} \alpha_d^{t+1} = 1$

⇒

$$\alpha_d^{t+1} = \mathbb{E}_\pi^X \left[\rho_d(X;\alpha^t,\theta^t)\right],$$

$$\theta_d^{t+1} = \arg\max_{\theta_d} \mathbb{E}_\pi^X \left[\rho_d(X;\alpha^t,\theta^t)\log(q_d(X;\theta_d))\right]$$

PMC: case of Gaussian mixtures

$$\theta_d^{t+1,N} = \arg\min_\theta \mathbb{E}_\pi^X \left[\rho_d(X;\alpha^t,\theta^t)\left(\log|\Sigma_d| + (X-\mu_d)^{\mathrm{T}}\Sigma_d^{-1}(X-\mu_d)\right)\right]$$

Solutions:

$$\mu_d^{t+1} = \frac{\mathbb{E}_\pi^X\left[\rho_d(X;\alpha^t,\theta^t)X\right]}{\mathbb{E}_\pi^X\left[\rho_d(X;\alpha^t,\theta^t)\right]},$$

$$\Sigma_d^{t+1} = \frac{\mathbb{E}_\pi^X\left[\rho_d(X;\alpha^t,\theta^t)(X-\mu_d^{t+1})(X-\mu_d^{t+1})^{\mathrm{T}}\right]}{\mathbb{E}_\pi^X\left[\rho_d(X;\alpha^t,\theta^t)\right]}.$$

$$\alpha_d^{t+1,N} = \sum_{i=1}^N \bar\omega_{i,t}\xi_{i,t},$$

$$\mu_d^{t+1,N} = \frac{\sum_{i=1}^N \bar\omega_{i,t}\xi_{i,t}X_{i,t}}{\sum_{i=1}^N \bar\omega_{i,t}\xi_{i,t}} = \sum_{i=1}^N \bar\omega_{i,t}\xi_{i,t}X_{i,t} \Big/ \alpha_d^{t+1,N},$$

$$\Sigma_d^{t+1,N} = \sum_{i=1}^N \bar\omega_{i,t}\xi_{i,t}(X_{i,t}-\mu_d^{t+1,N})(X_{i,t}-\mu_d^{t+1,N})^{\mathrm{T}} \Big/ \alpha_d^{t+1,N}$$

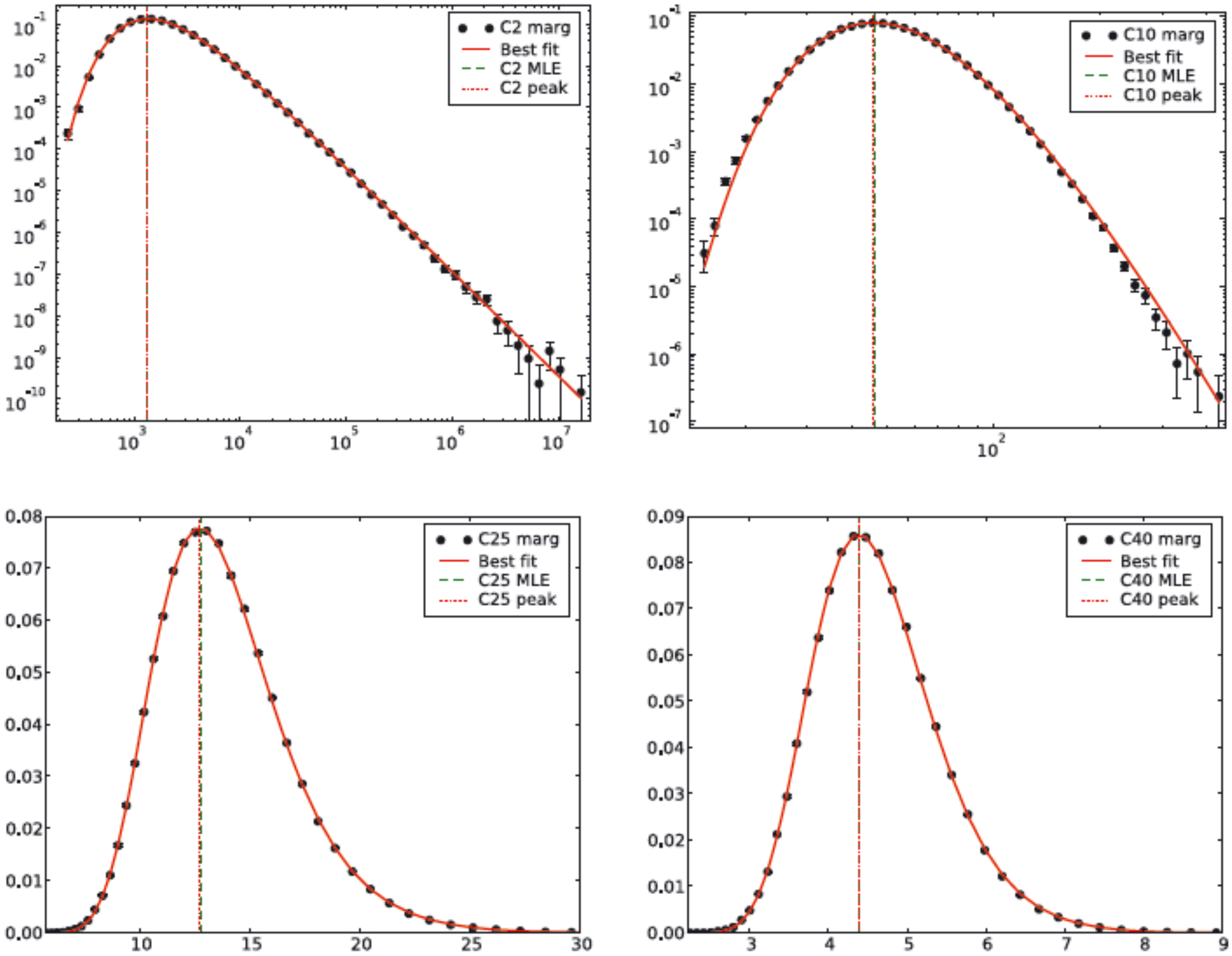
where

$$\xi_{i,t} = \rho_d(X_{i,t};\alpha^{t,N},\theta^{t,N})$$

Can be easily generalized to log-normals, but also multivariate Student mixtures

Results on synthetic (low-l) data

Benabed, Cardoso, Prunet, Hivon 2009



Taking into account the covariance: Copula approximation

$$\mathcal{N}(G_\ell;0,1)dG_\ell = i\Gamma(D_\ell;\alpha_\ell,\beta_\ell)dD_\ell$$

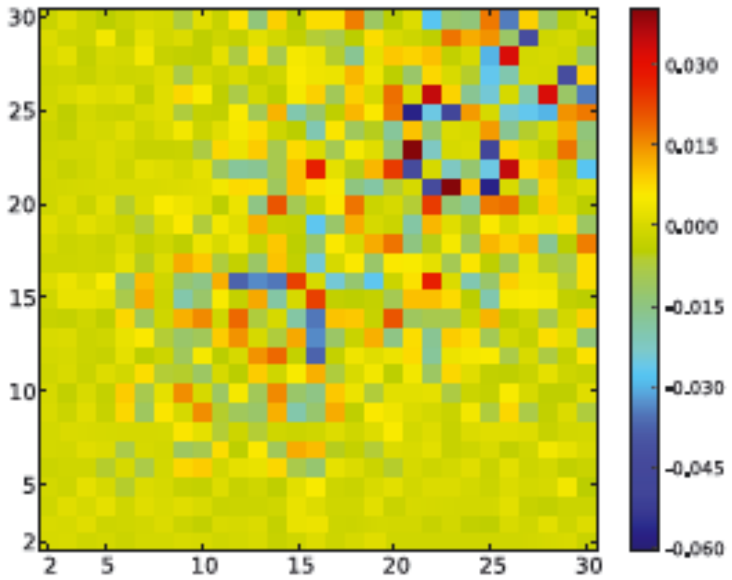
$$\tilde{\pi}(D_\ell) \equiv \prod_k i\Gamma(D_k;\alpha_k,\beta_k)\,\frac{\mathcal{N}^{(d)}(G_\ell;0,\mathrm{M_G})}{\prod_k \mathcal{N}^{(1)}(G_k;0,1)}.$$

Gaussianization of marginals, defines G_l as functions of D_l

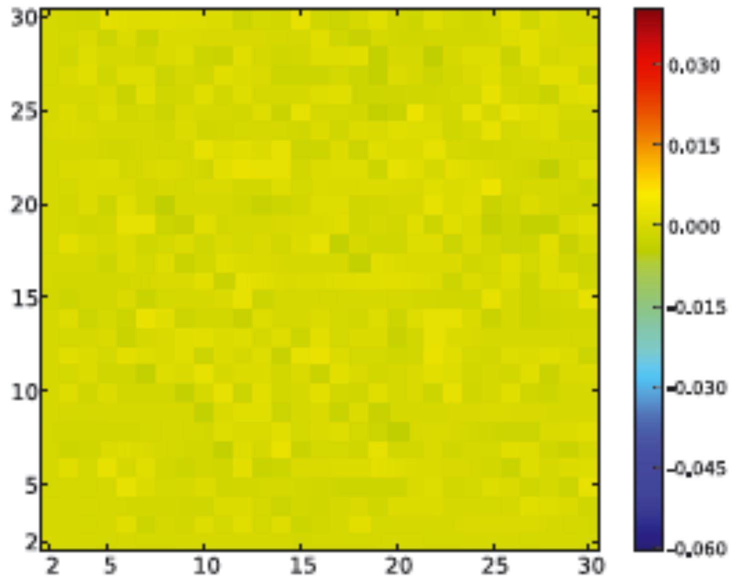
Gaussian copula approximation

M_G is measured (for now) on the (Gaussianized) samples

WMAP5 low-res



Measured D_l correlation



Correlation predicted from copula

Benabed, Cardoso, Prunet, Hivon 2009

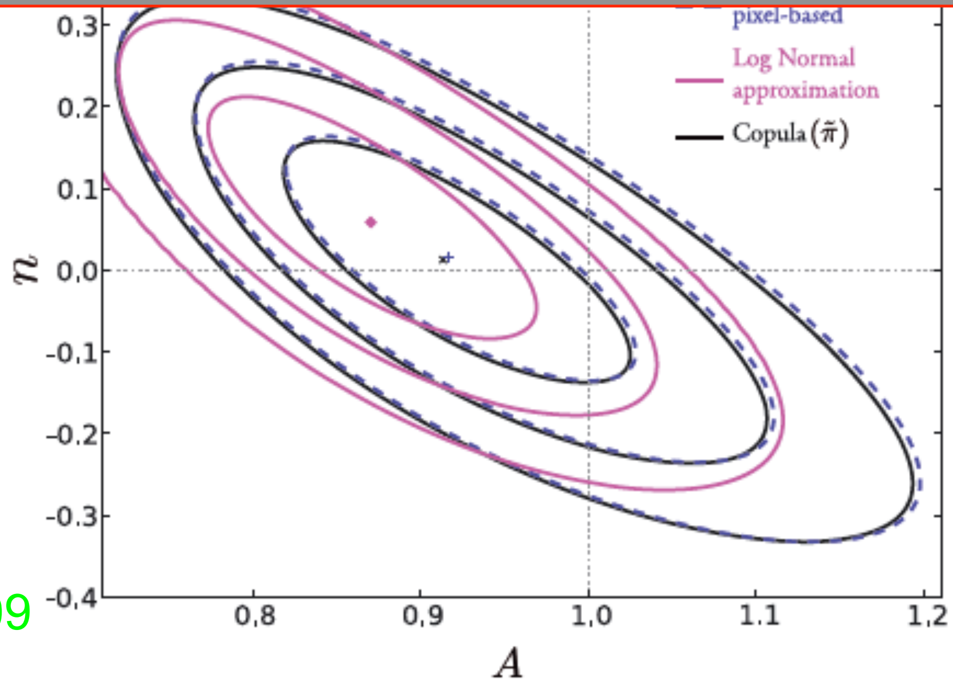
Quality of approximation: parameter posteriors

Approximation	Perplexity	Kullback ($\times 10^{-3}$)
Copula $\tilde{\pi}$	0.991	8.6
Uncorrelated copula $\tilde{\pi}_0$	0.965	35.2
Uncorrelated last run	0.956	45.0
Naive $\tilde{\pi}_{\text{naive}}$	0.779	249.6
Lognormal	0.191	1655.3

TO COME SOON: APPROXIMATIONS WITH NO SAMPLING REQUIRED !

- WMAP5 data
- Simple 2-parameters model

$$\tilde{C}_\ell \equiv C_\ell^{\text{ref}} \times A \left(\frac{\ell}{\ell_0}\right)^n$$



Benabed, Cardoso, Prunet, Hivon 2009

Summary

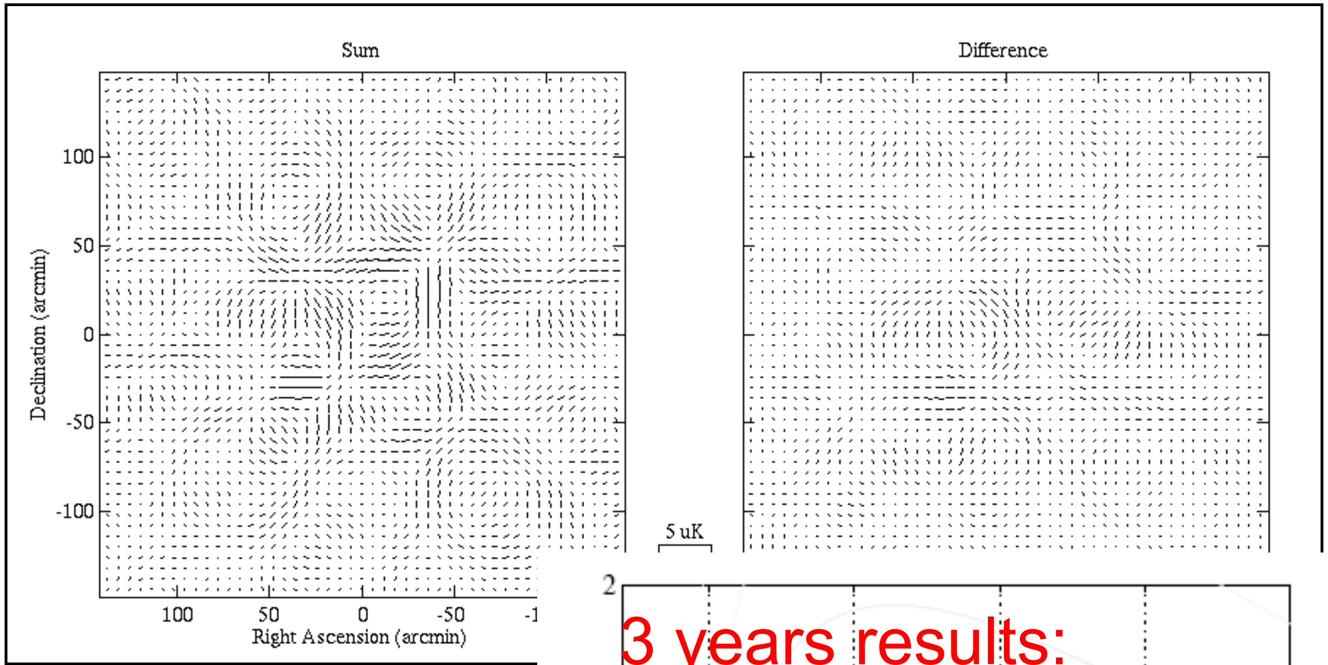
- Provided with a linear instrument and Gaussian noise, (optimal) maps are
  - available (through iterative solvers)
  - sufficient statistics
- Given a map, we know how to find the peak of the spectrum likelihood
  - exactly at large scales
  - approximately at small scales
- Given a peak (and curvature at peak) of the PS likelihood, we can have:
  - (proven) good approximation of the likelihood at large scales (polar ?)
  - approximations at small scales ? Use asymptotics ? (open pb)
  - how to stitch large and small scales ?? (open pb)
- Multi-channel, multi-components likelihoods (see JFC's lectures)
- Boltzmann codes + PS likelihoods ➔ cosmological inference
  - via MCMC codes (e.g. COSMOMC)
  - via PMC codes

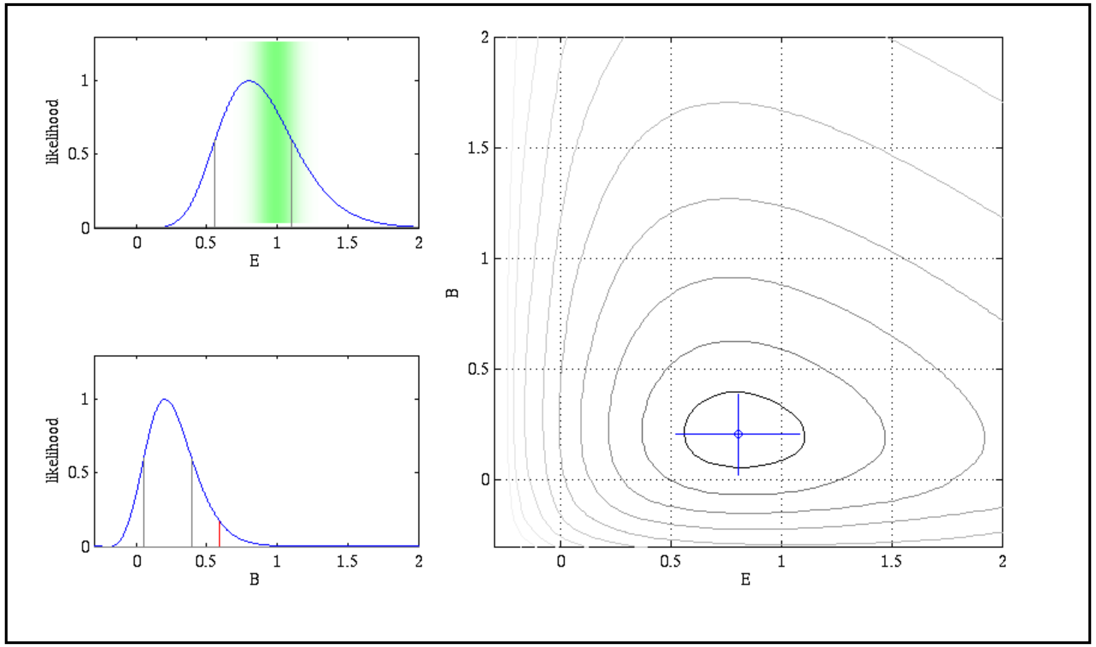
Polarisation measurements
Status and perspectives

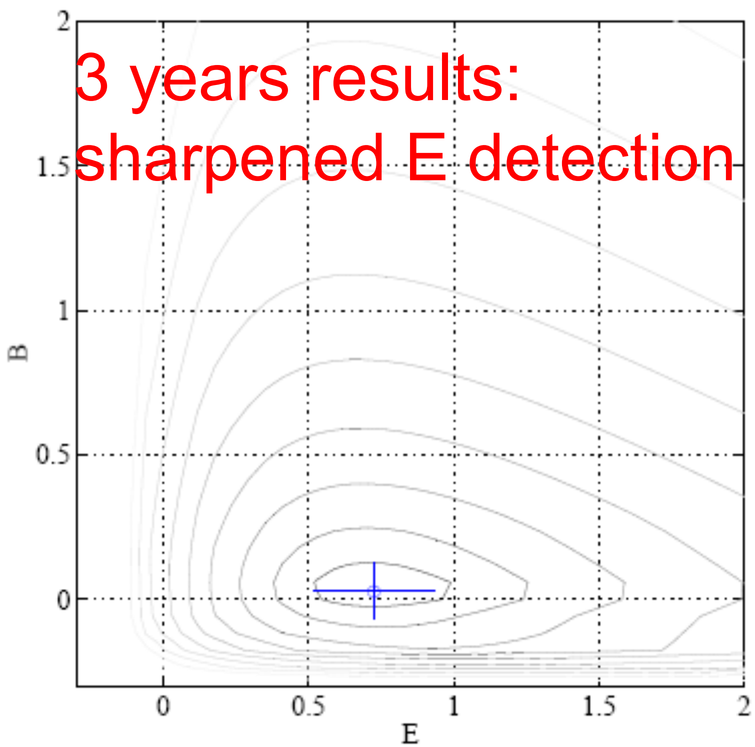
Polarisation: first measurement

First detection of E-mode polarisation by DASI (2002)

Kovac et al. 2002
Leitch et al. 2005





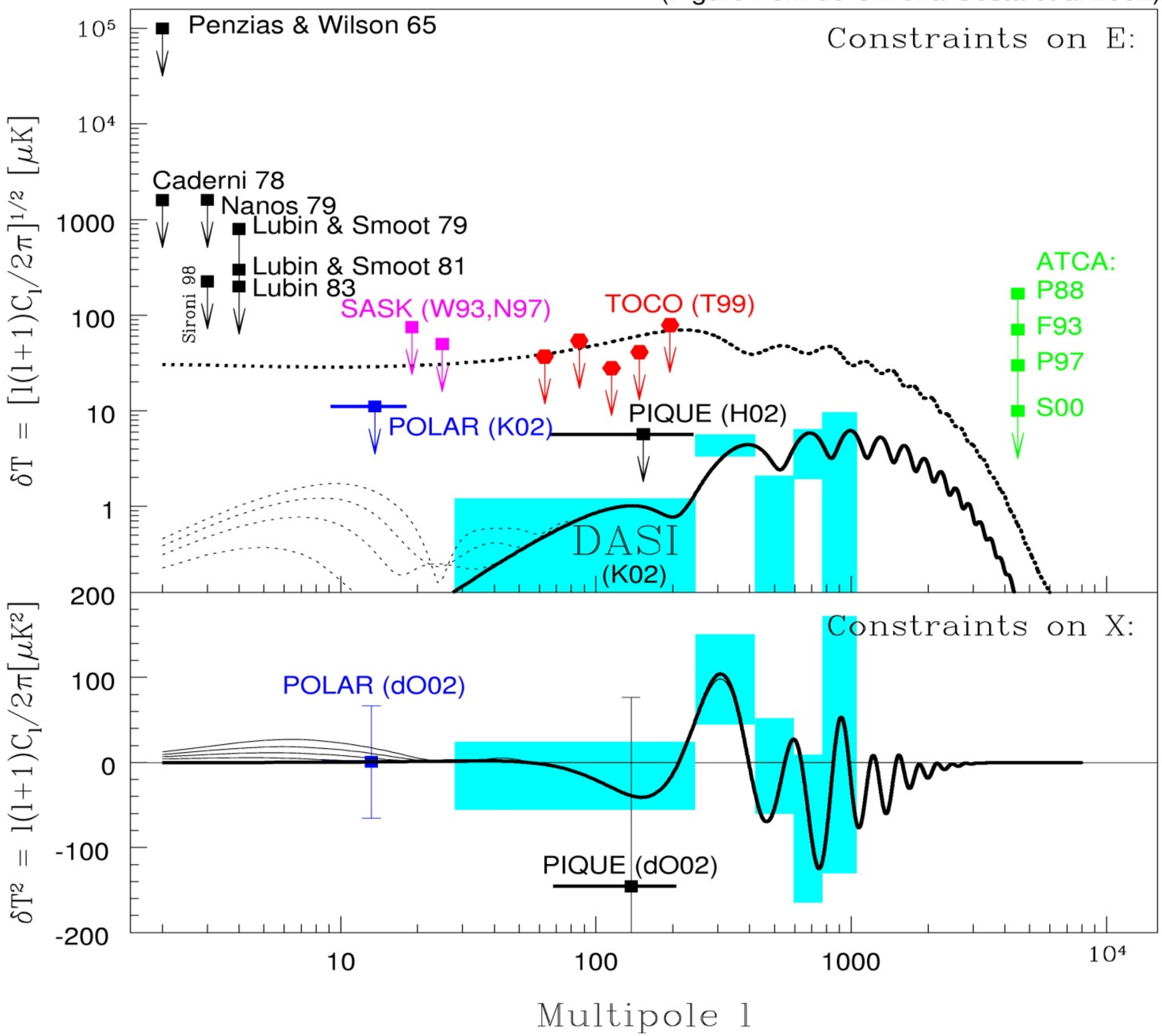


3 years results: sharpened E detection

Polarisation

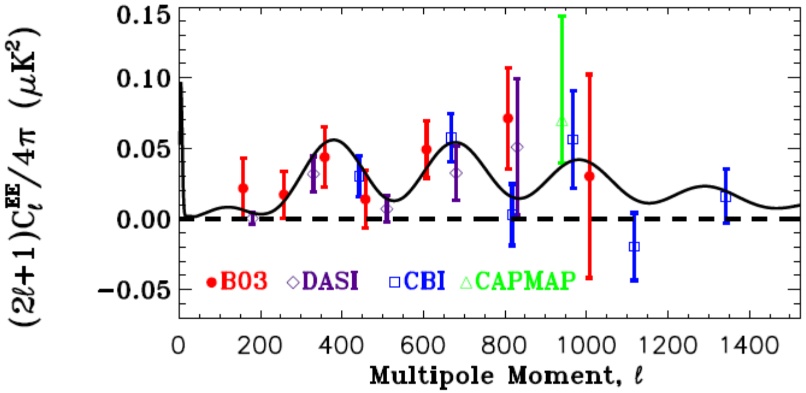
Upper limits and first measurement by DASI

(Figure from de Oliveira-Costa et al 2002)



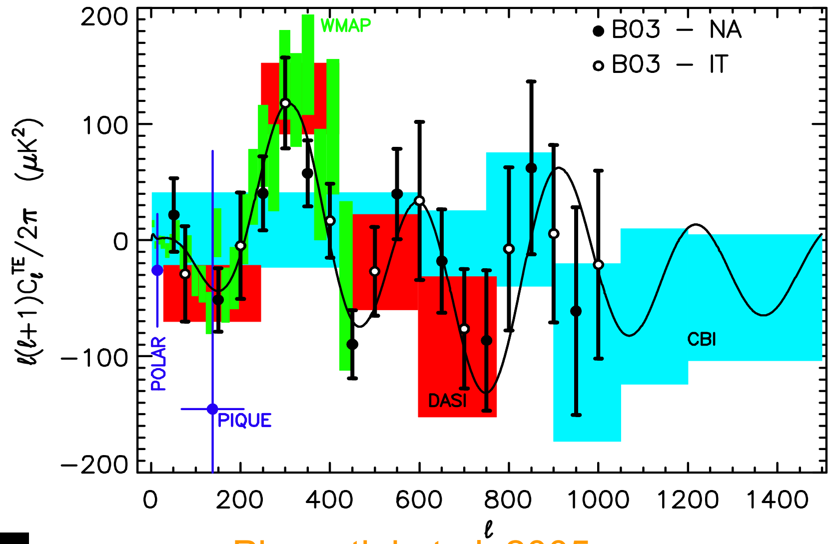
Polarisation spectra: 2005 observational status

EE power spectrum

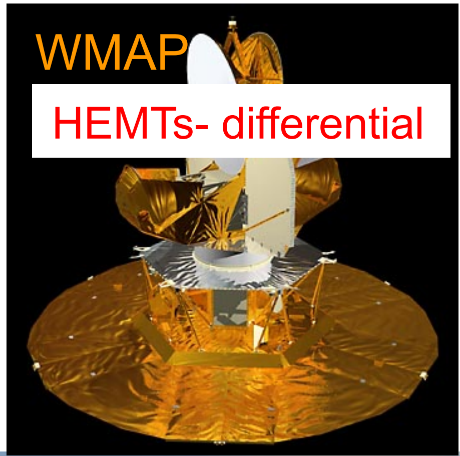


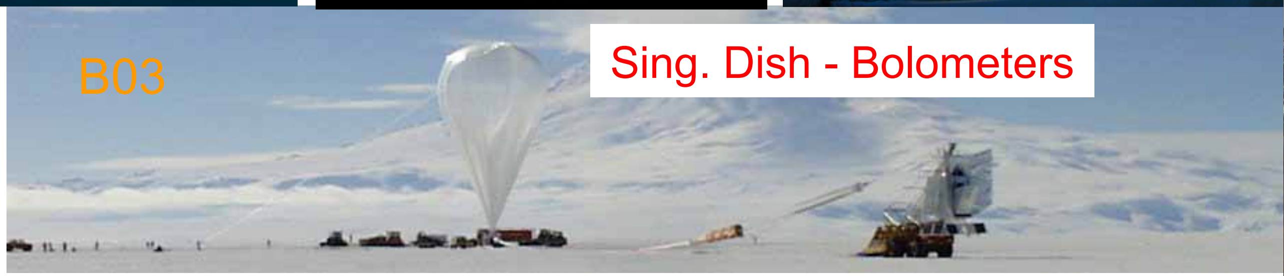
Montroy et al. 2005

TE cross-spectrum

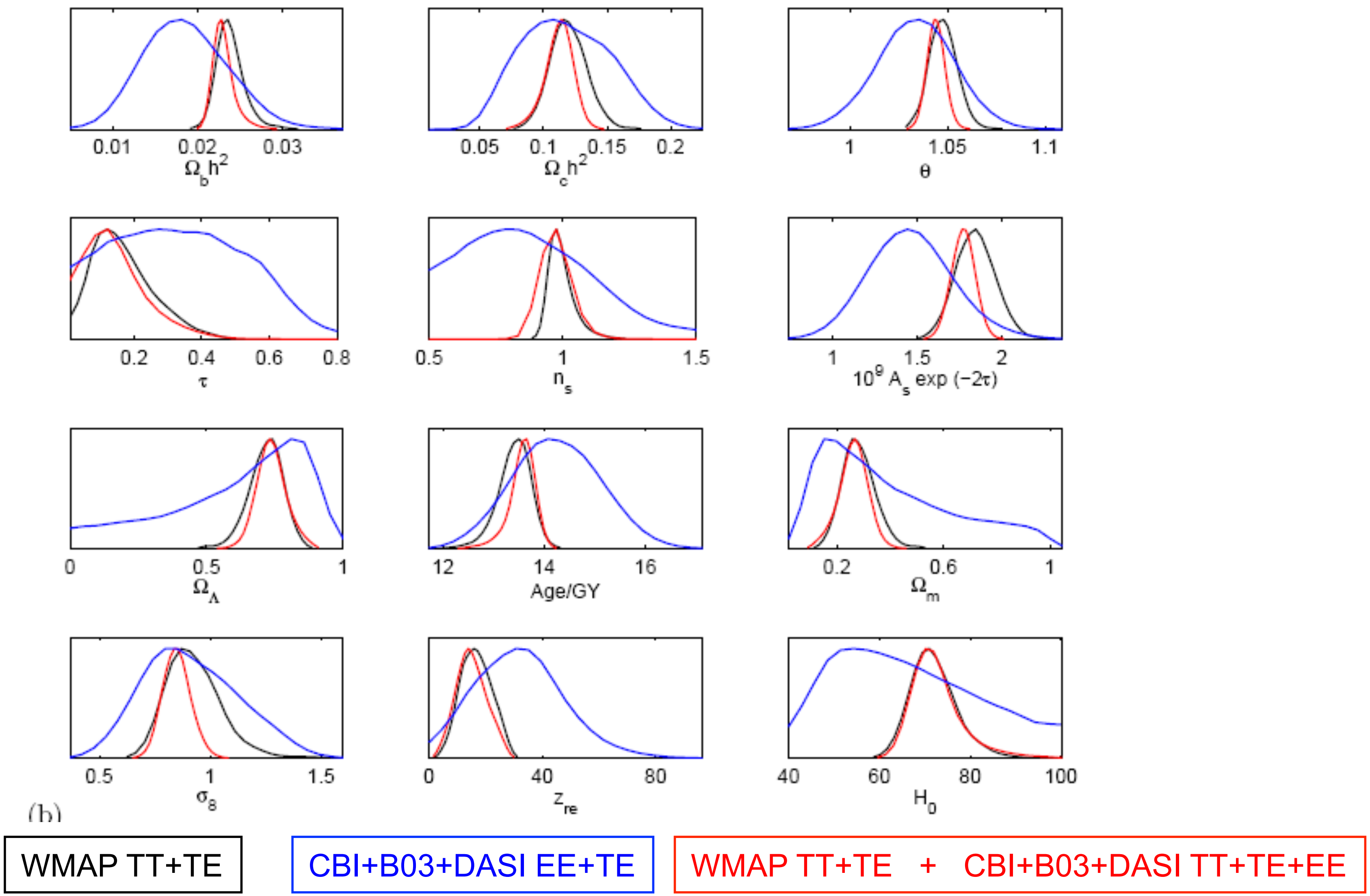


Piacentini et al. 2005



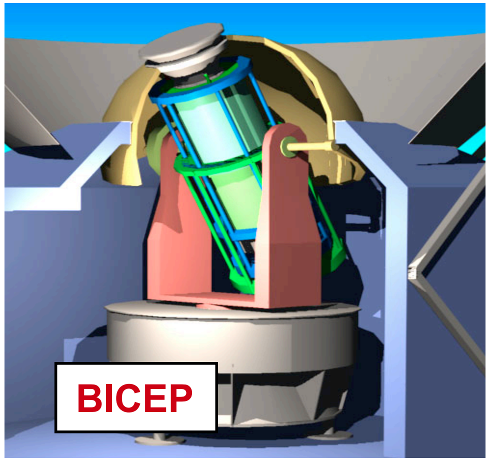


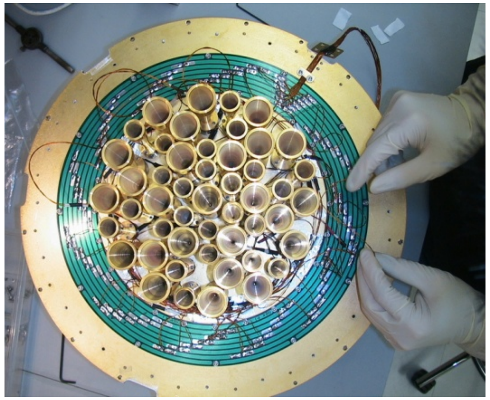
Cosmological consistency

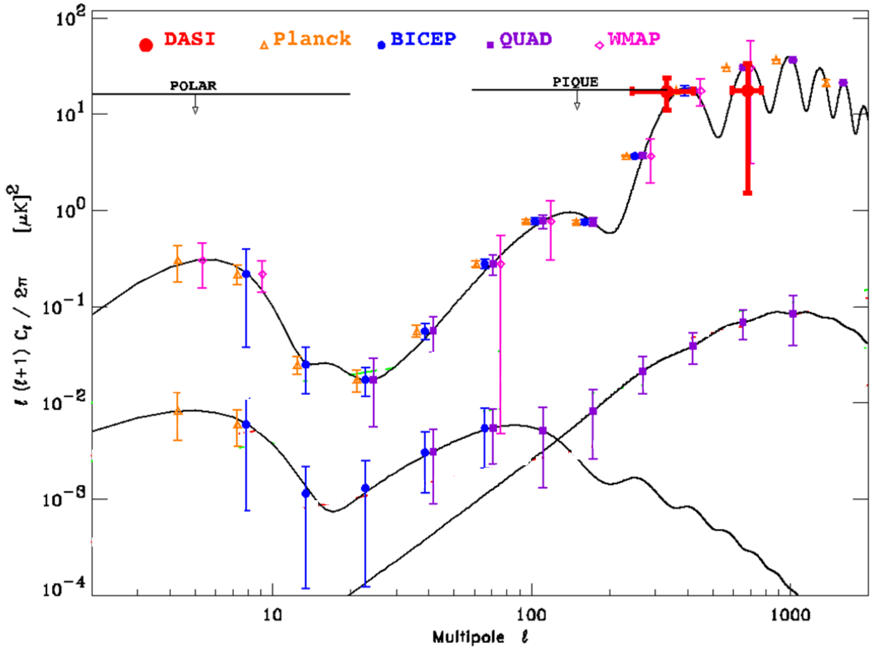


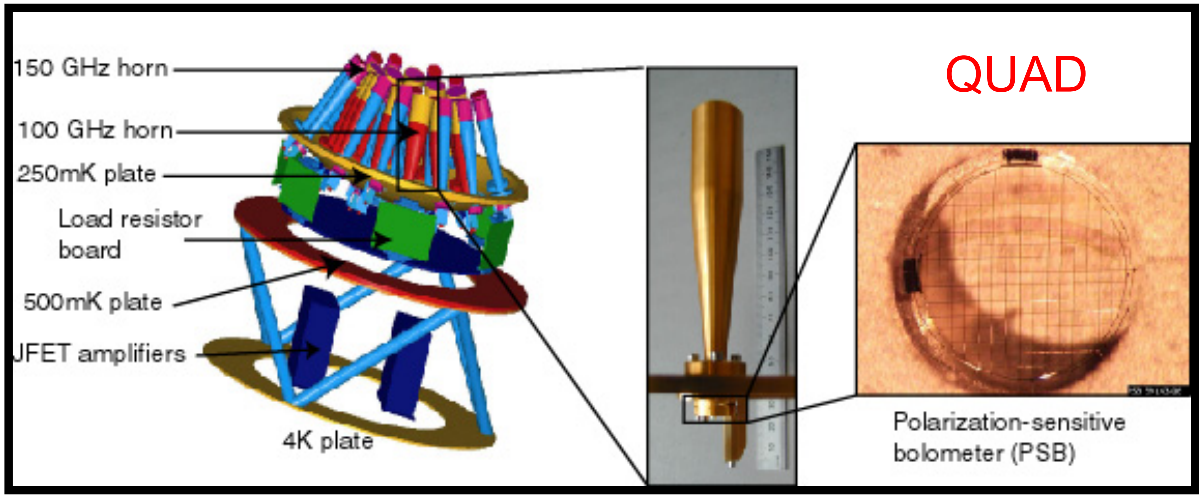
Sievers et al. 2005

Polarisation: on-going









WMAP5: first large-scale polarization maps

Q Stokes

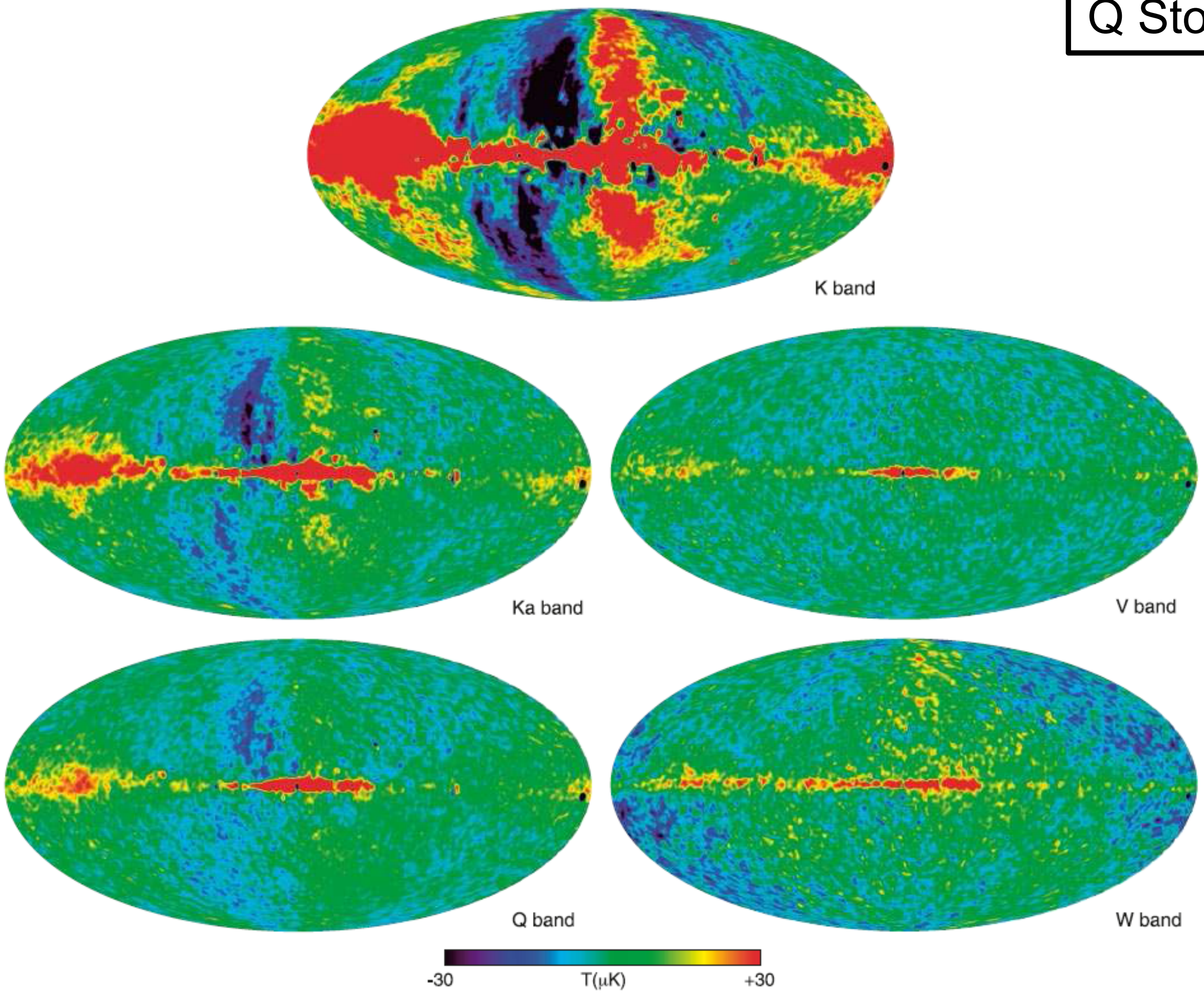


Fig. 2.— Five-year Stokes Q polarization sky maps in Galactic coordinates smoothed to an effective Gaussian beam of 2.0°, shown in Mollweide projection. *top*: K band (23 GHz), *middle-left*: Ka band (33 GHz), *bottom-left*: Q band (41 GHz), *middle-right*: V band (61 GHz), *bottom-right*: W band (94 GHz).

Hinshaw et al. 08

WMAP5: first large-scale polarization maps

U Stokes

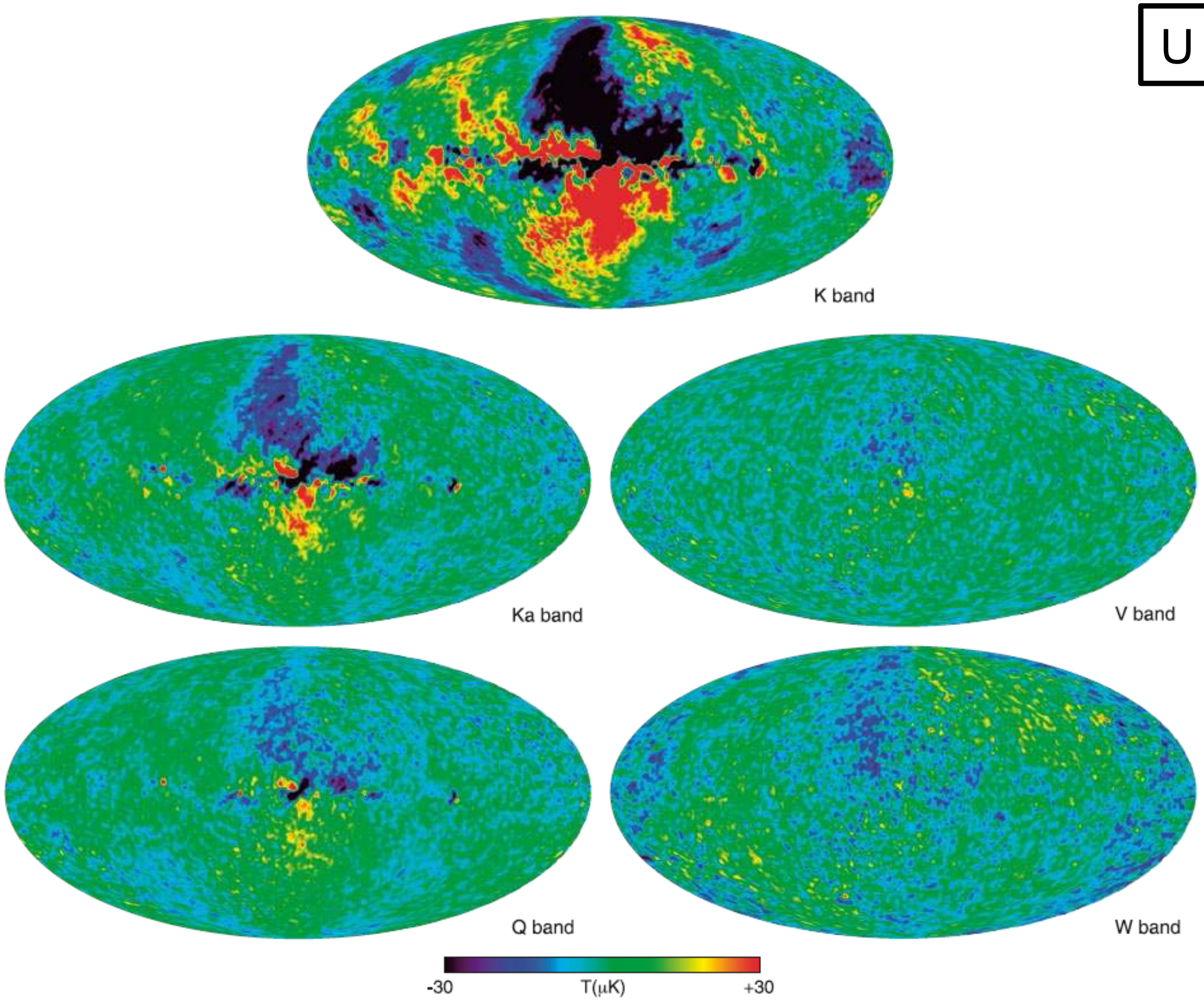
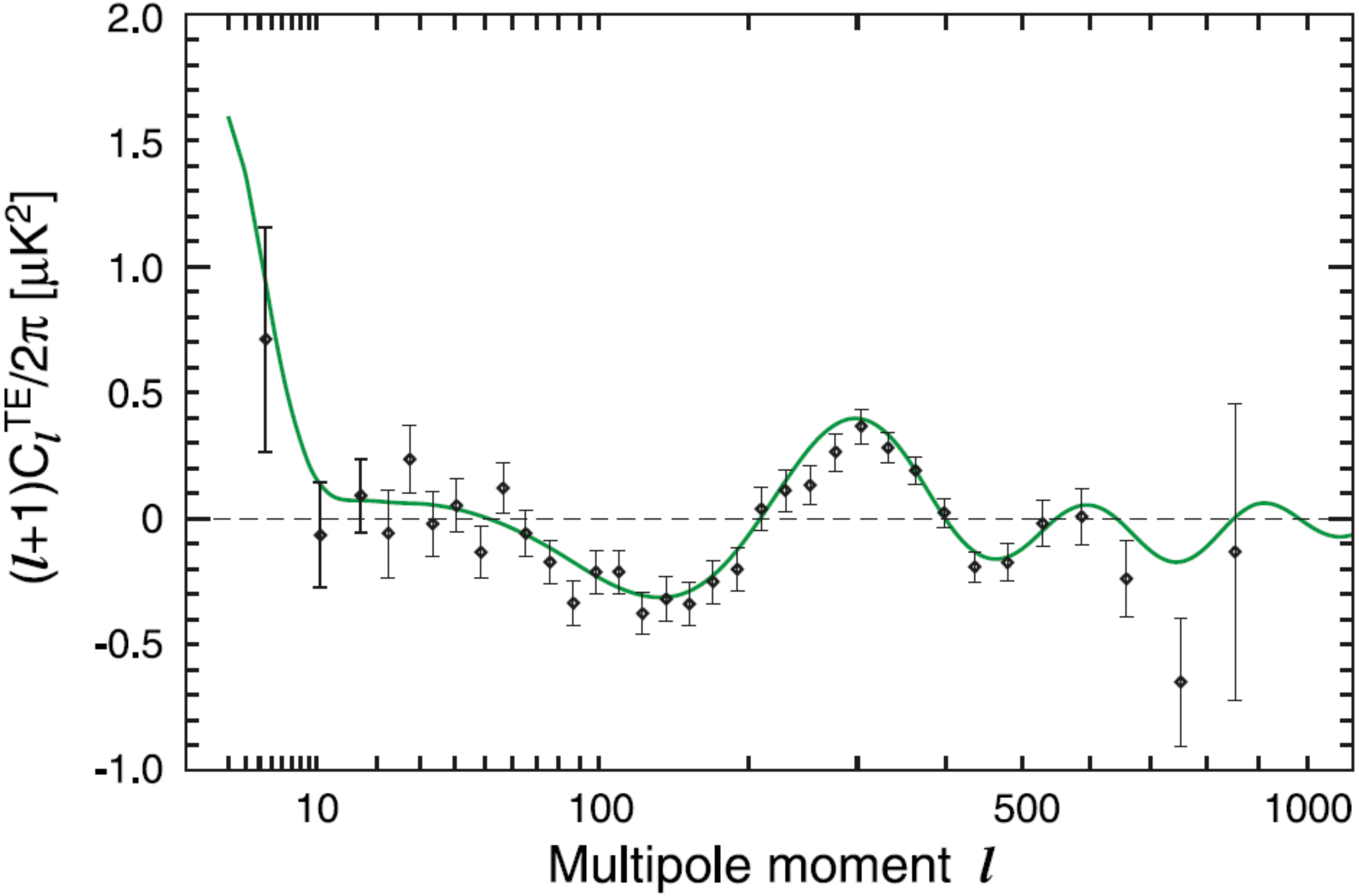


Fig. 3.— Five-year Stokes U polarization sky maps in Galactic coordinates smoothed to an effective Gaussian beam of 2.0°, shown in Mollweide projection. *top*: K band (23 GHz), *middle-left*: Ka band (33 GHz), *bottom-left*: Q band (41 GHz), *middle-right*: V band (61 GHz), *bottom-right*: W band (94 GHz).

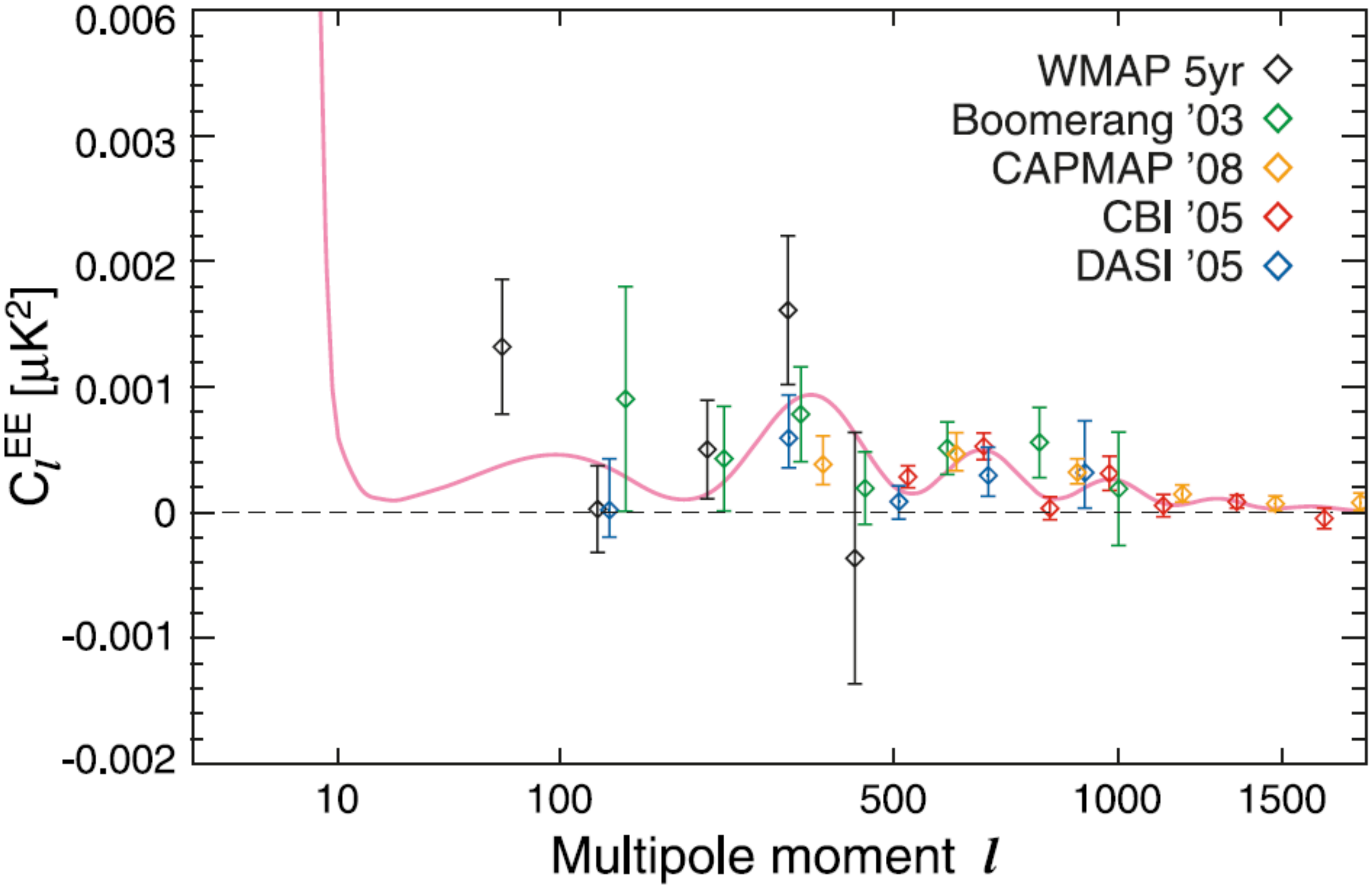
Hinshaw et al. 08

WMAP5: TE cross-spectrum



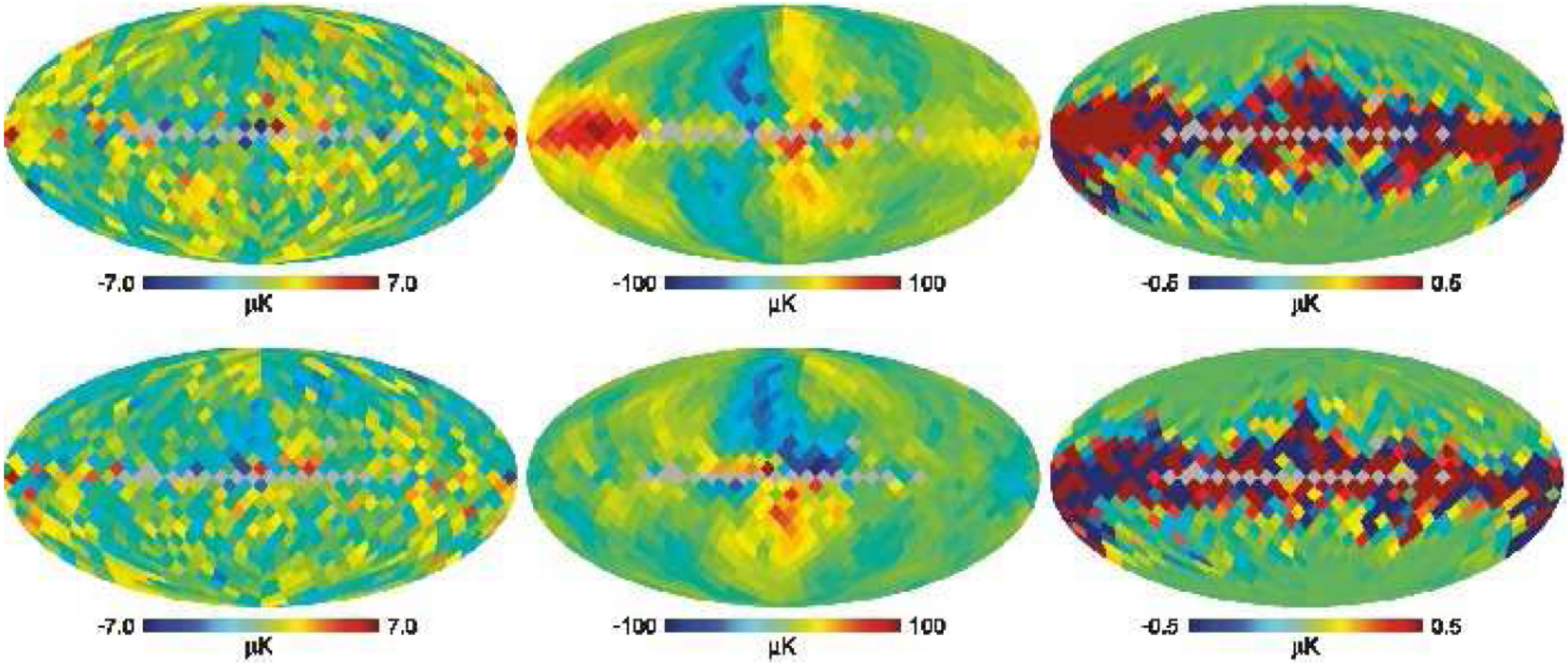
Nolta et al. 08

WMAP5: EE spectrum



Nolta et al. 08

Polarized maps: CMB, sync, dust (Q,U) Stokes parameters

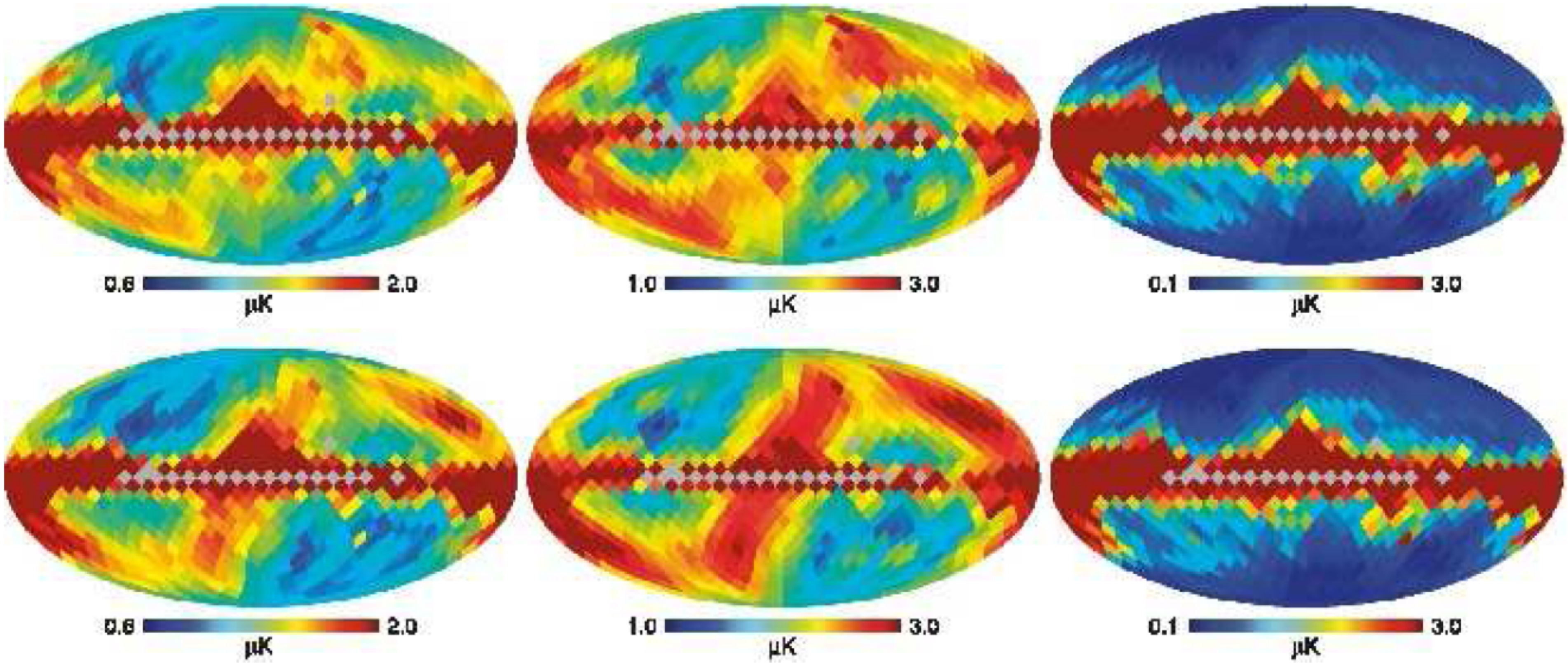


Component separation with Gibbs sampling, see later

Dunkley et al. 08

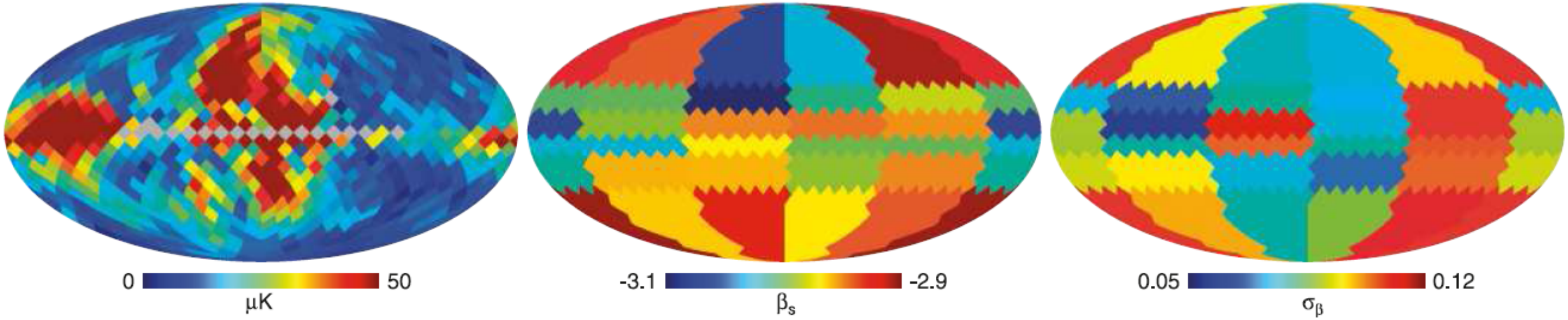
1-sigma pixel errors, CMB, sync, dust

– 8 –



Dunkley et al. 08

Synchrotron amplitude, spectral index and error on latter



- Correlated patterns between synchrotron amplitude and index
- Polarized likelihood on foreground template cleaned map

Dunkley et al. 08

QUAD second release (2008)

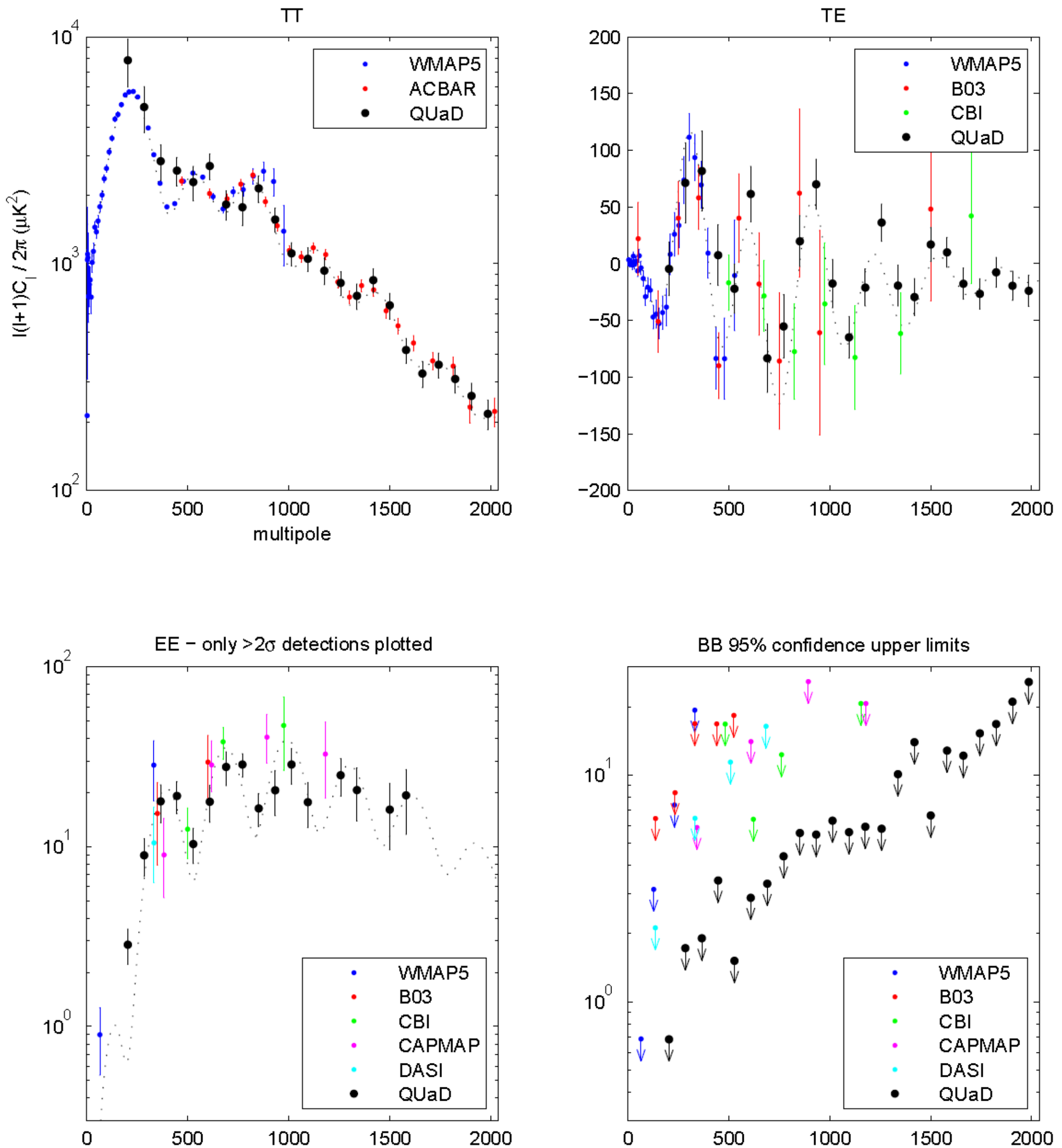
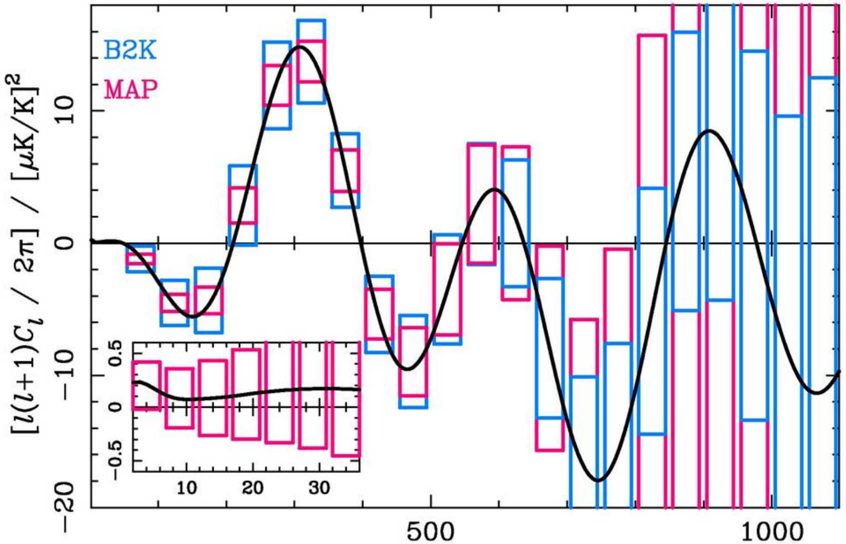


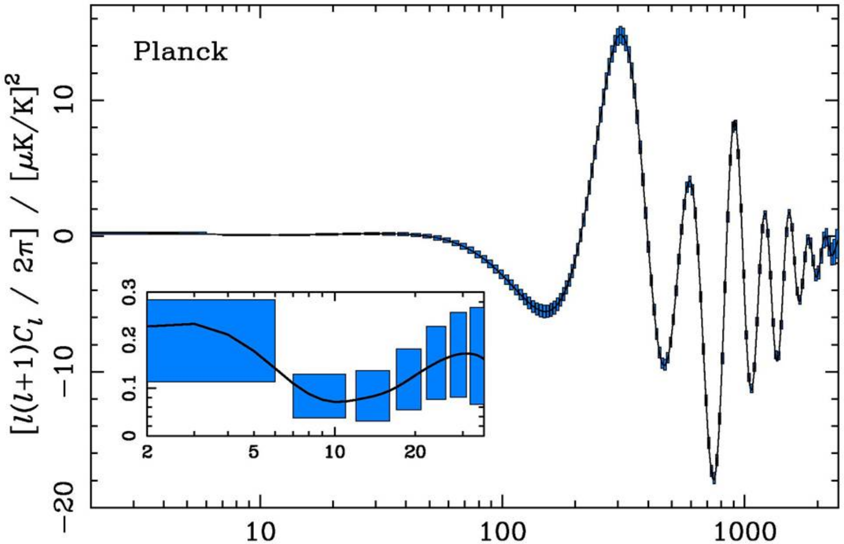
FIG. 22.— QUaD power spectra compared to results from WMAP (Nolta et al. 2008), ACBAR (Reichardt et al. 2008), B03 (Piacentini et al. 2006; Montroy et al. 2006), CBI (Sievers et al. 2007), CAPMAP (Bischoff et al. 2008) and DASI (Leitch et al. 2005). The BB upper limits are stated values where provided, and otherwise the 95% point of the positive part of the bandpower pdf.

Pryke et al. 08

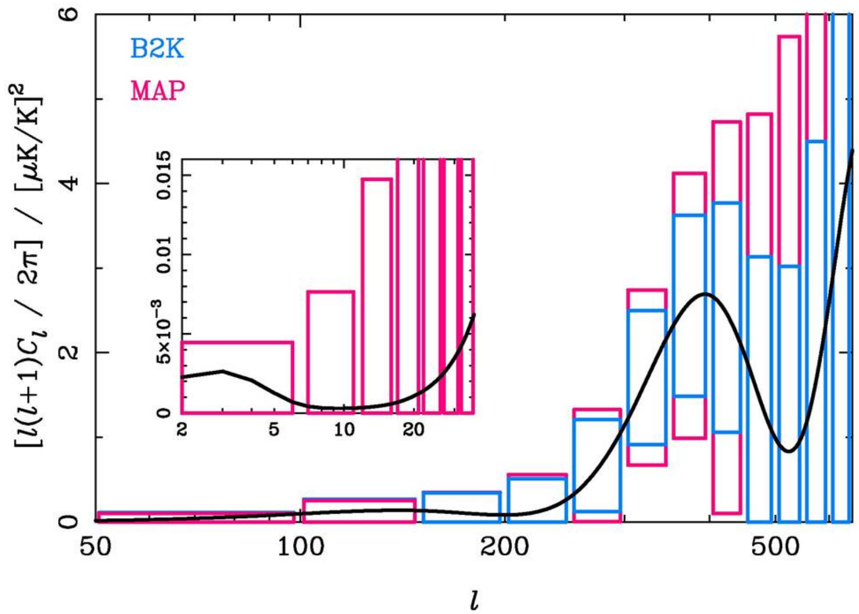
Polarisation: Planck

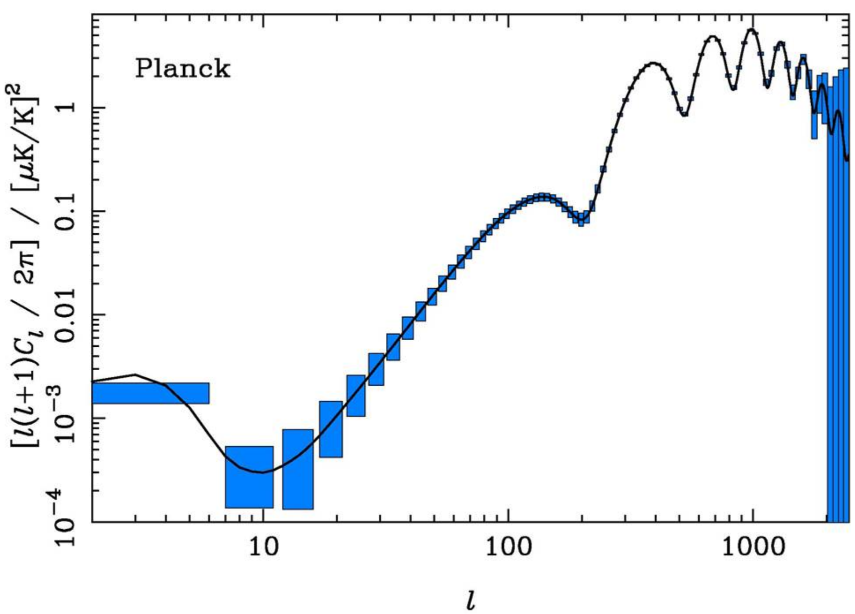
http://background.uchicago.edu/~whu/



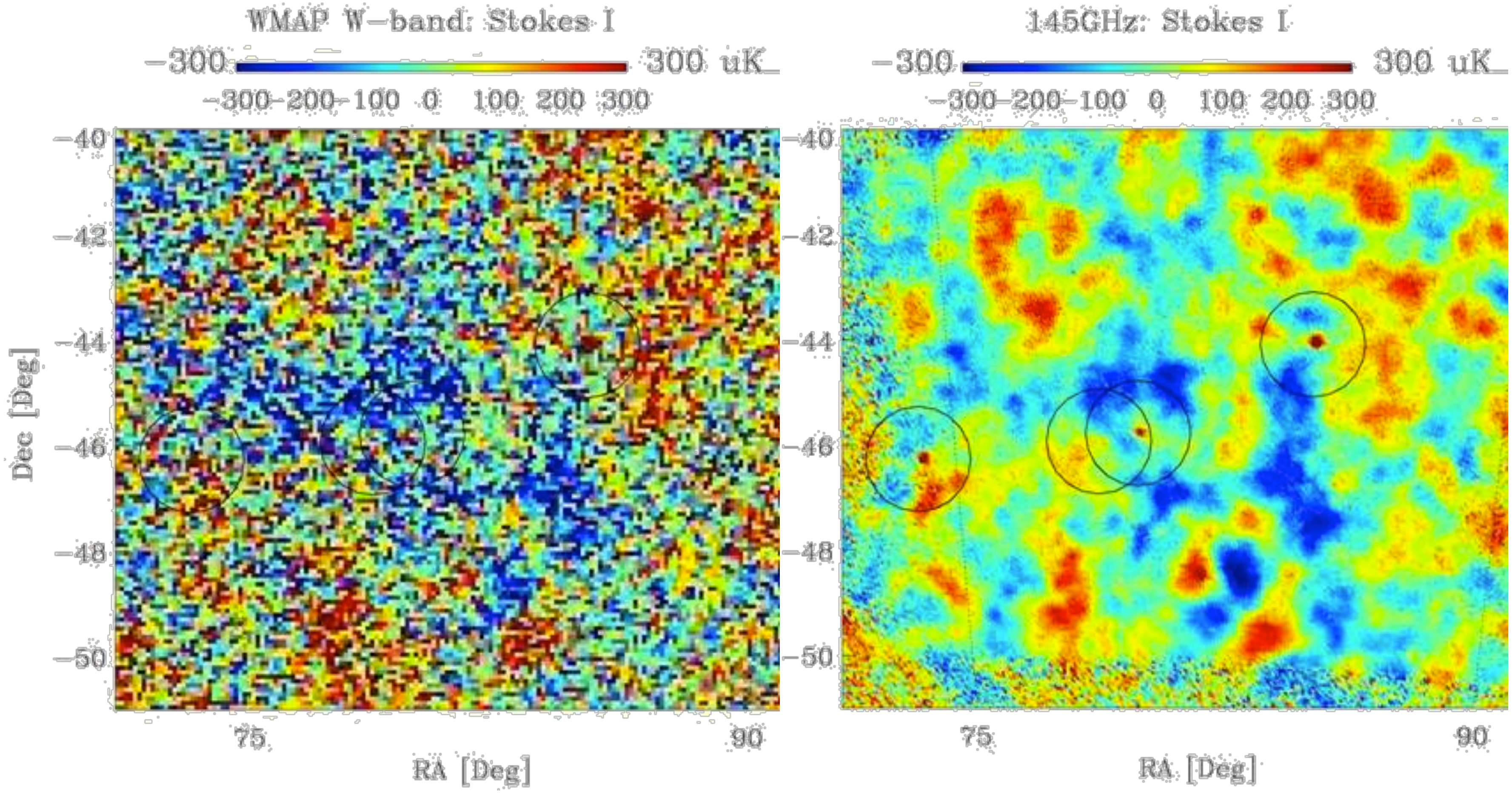


Nice propaganda ! But does not include systematics …





B03 Deep survey

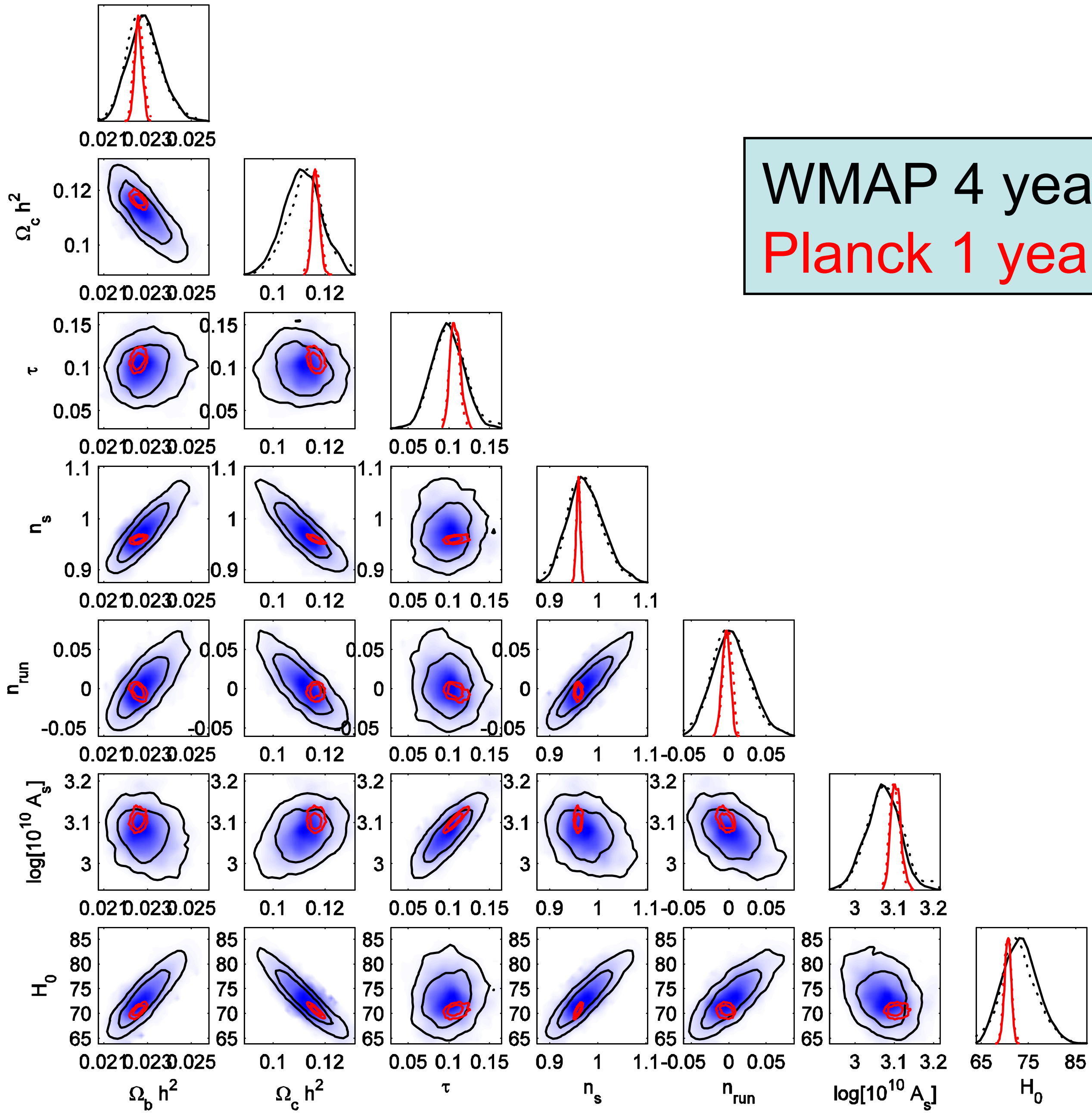


Masi et al. astroph/0507509

A foretaste of Planck-HFI @ 145 GHz but:

- w_I =82μK.arcmin, while HFI goal is w_I =42μK.arcmin @ 143GHz (OK FM bolos delivered~36)
- Planck has matching sensitivities in 9 frequency bands, e.g. ~60 μK.arcmin @ 100 & 217 GHz
- 90 deg^2, i.e. 0.2% of the sky covered, instead of 100% (and deep surveys in Planck too)

Why bother: parameters posteriors



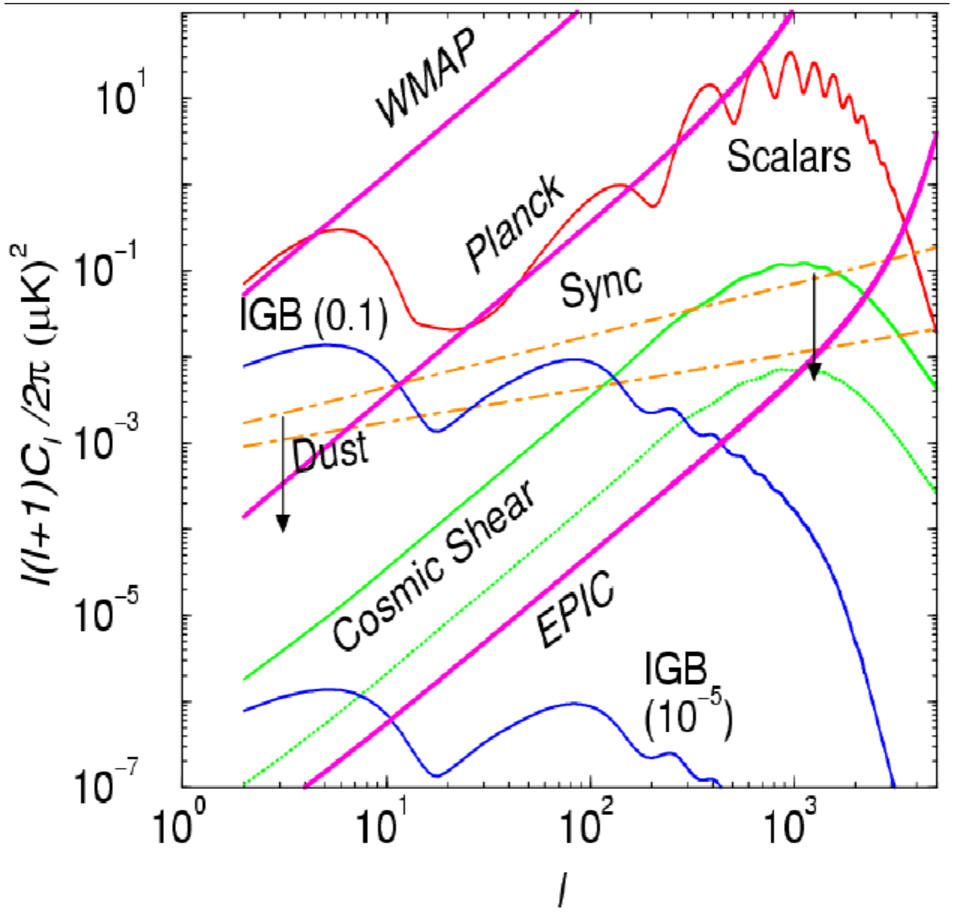
Bond et al. 04

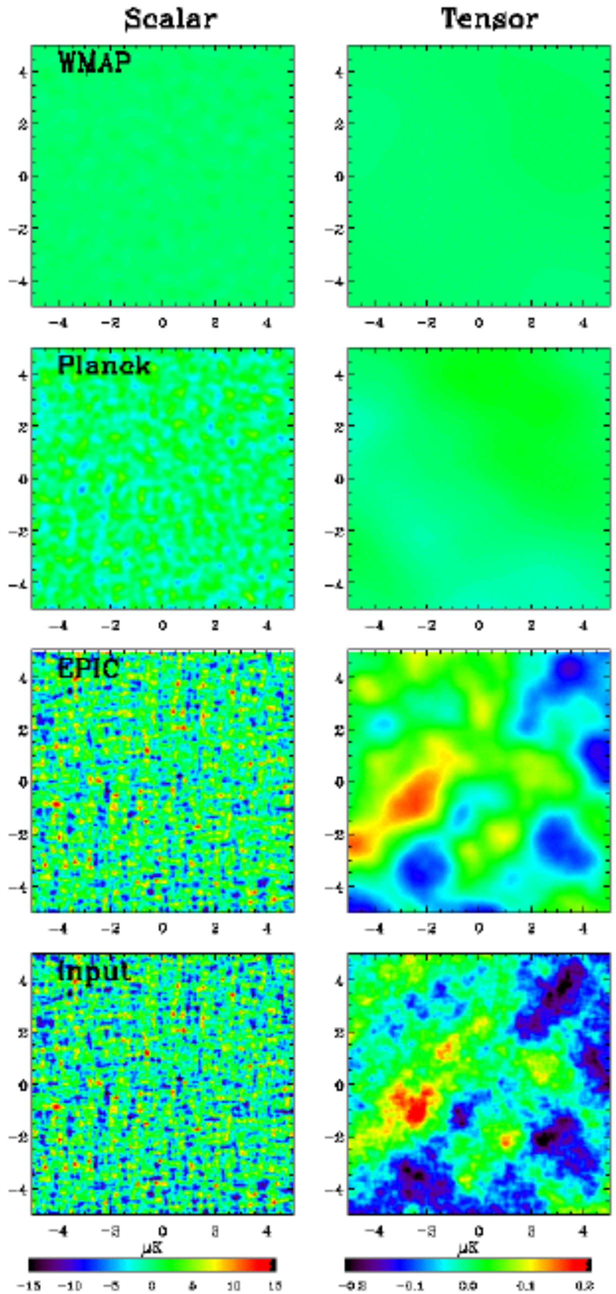
Space-borne polarimeter

Specific needs

- Specific design to control instrumental systematics
  - Thermal stability (tiny signals !)
  - Instrumental polarisation control
  - Optimized scanning strategy
- Detectors are ~background limited
  - Need a lot of them !!
  - Detector arrays, no horns, big focal planes

Polarisation: the future challenge





Courtesy EPIC consortium

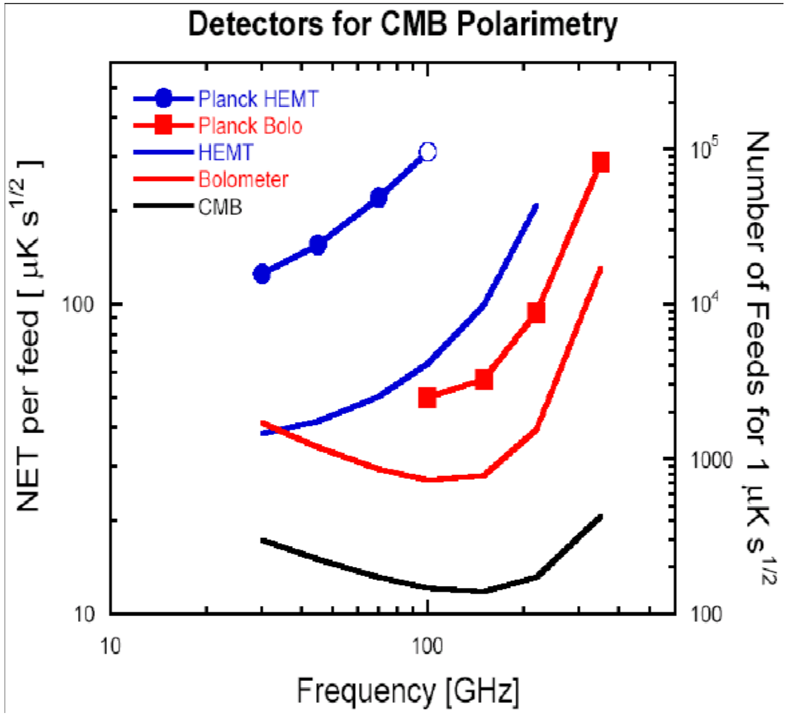
- Primordial GW background: no theoretical prior on amplitude…
- One-field inflationary models: Tensor amplitude varies as E_{inf}^4
- Lensing-induced B-modes: dominant at least on small scales
- Polarized foreground emissions are nearly unknown ...

Polarisation from space: requirements

- Large scales: space required
- Stable environment: space …
- Detectors are background limited
    - need lots of them !
    - detector arrays
    - large telemetry …
- Stringent systematics control

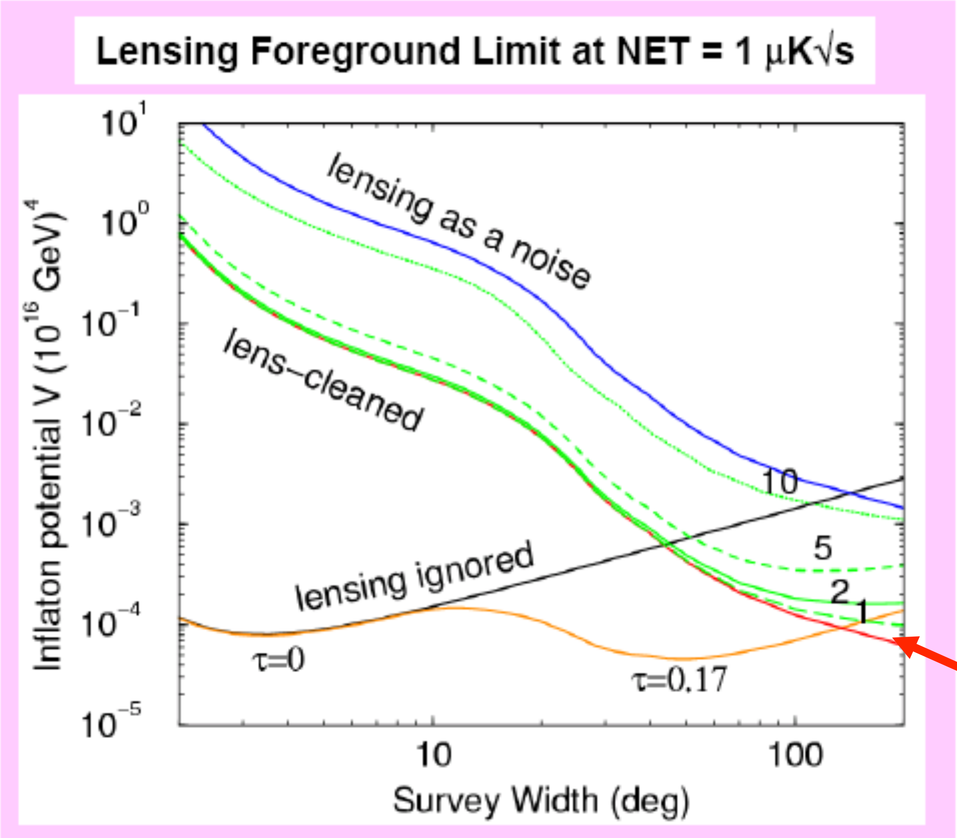
Current and future focal planes

	Future		Planck	
Freq	NET (calc)	# feeds for 1 uK√ s	NET (goal)	# feeds
30	38	1500	125	2
45	42	1750	155	3
70	25	750	220	6
100	25	750	55	4
150	25	750	57	4
220	38	1500	95	4
350			290	4



Courtesy EPIC consortium

Lensing-induced B-mode cleaning



Substract lensing-induced BB by reconstruction of deflection angle using 4-point minimum variance estimators (Hu & Okamoto 2002)

Exponantial cut-off of CMB anisotropies at small scales limits lensing reconstruction

Kesden, Cooray, Kamionkowski (2002)

Lensing "cleaning": improvement ?

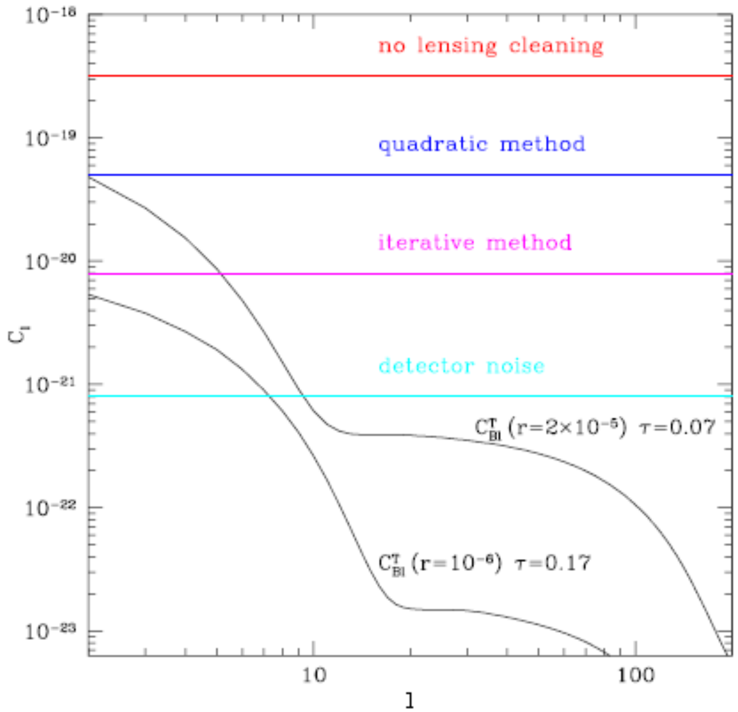


FIG. 2: Power spectra of noise for 2′, 0.25μK arcmin instrument with no lensing cleaning, cleaning with quadratic method and cleaning with iterative maximum-likelihood method. Also shown are two theoretical power spectra for $r = 2 \times 10^{-5}$ and $r = 10^{-6}$. Assuming this instrument specifications and iterative method the former can be detected (at 2-σ) both in reionization peak ($l < 20$) and in recombination peak $l > 20$), while the latter is detectable for $l < 20$ only. The noise power spectra have been averaged over the $l < 150$ range.

Hirata & Seljak 2003

- Iterative ML method
- Gains in the low-noise limit

by reducing the CV of the residual