Challenges in Cosmic Microwave Background Data Analysis

Steven T. Myers

National Radio Astronomy Observatory

Socorro, NM

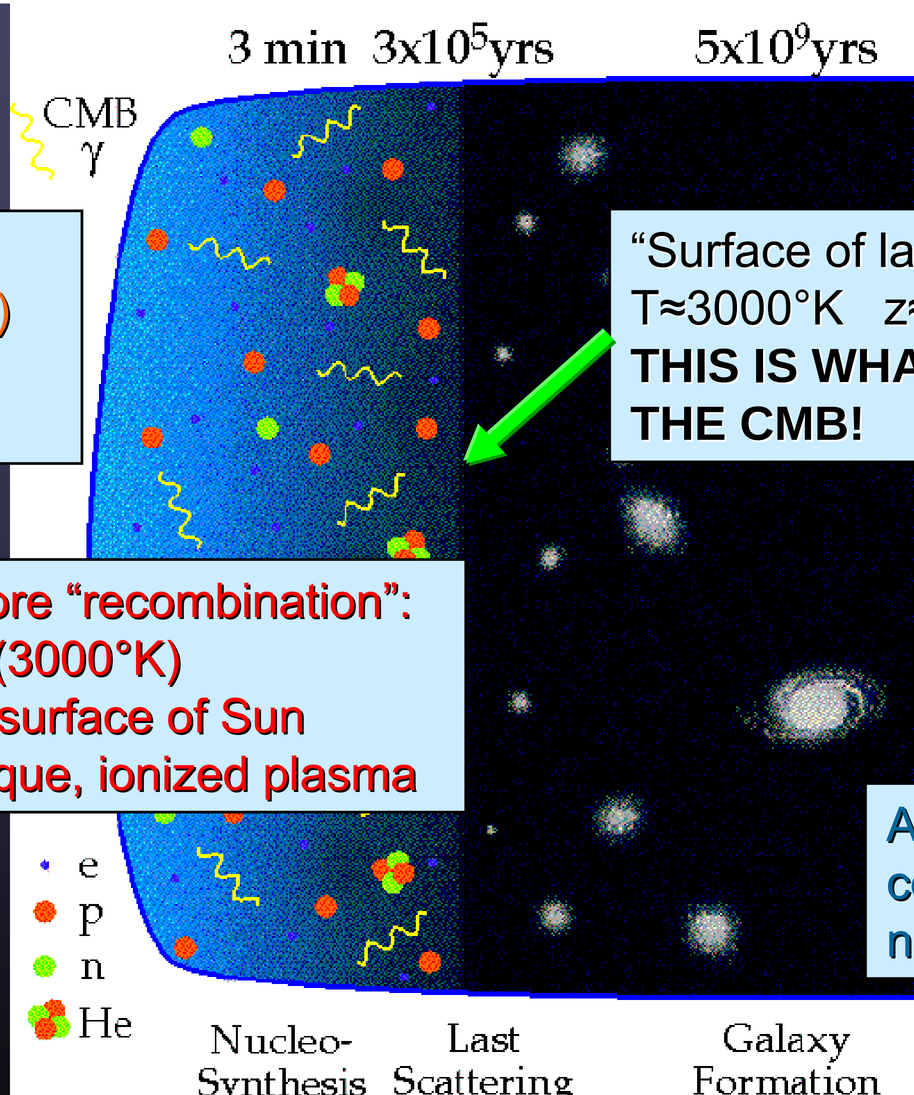
The problem…

- Use observations of the intensity and polarization of the sky at centimeter and millimeter wavelengths to infer the values (or constraints on the values) of fundamental cosmological parameters

The CMB problem

Thermal History of the Universe

“First 3 minutes”:
very hot (10 million °K)
like interior of Sun
nucleosynthesis!



“Surface of last scattering”
T≈3000°K   z≈1000
THIS IS WHAT WE SEE AS THE CMB!

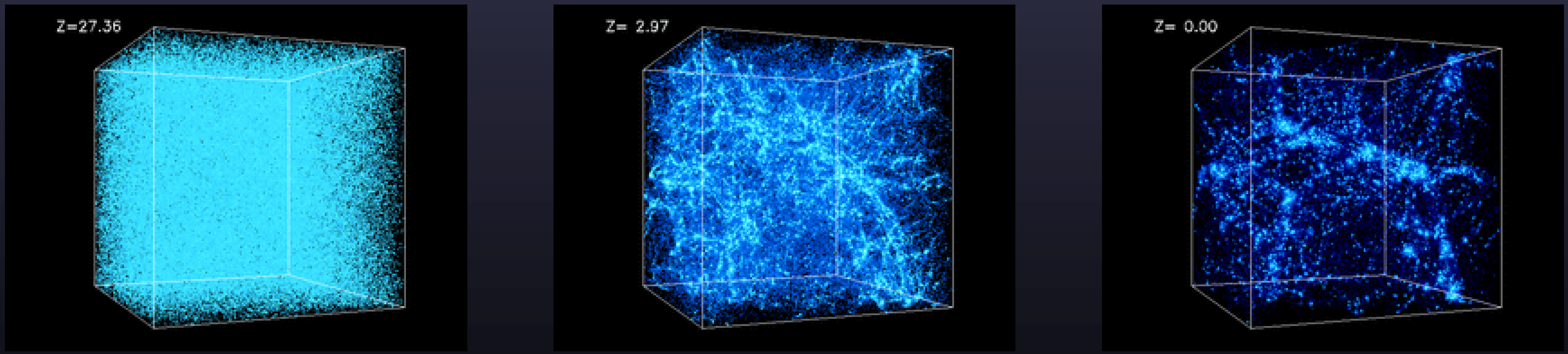
Before “recombination”:
hot (3000°K)
like surface of Sun
opaque, ionized plasma

After “recombination”:
cooler, transparent,
neutral hydrogen gas

Courtesy Wayne Hu – *http://background.uchicago.edu*

Matter History of the Universe

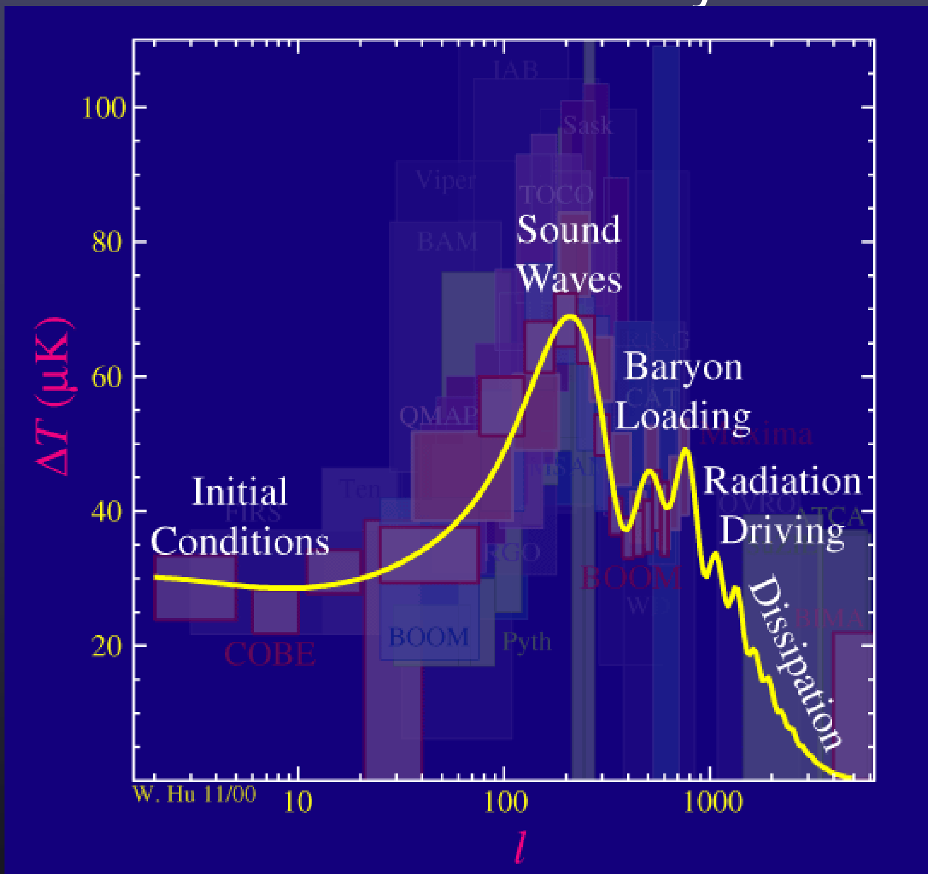
- we see “structure” in Universe now
  - density fluctuations ~1 on 10 h^{-1} Mpc scales
  - scale of clusters of galaxies
- must have been smaller in past (fluctuations grow)
  - in expanding Universe growth is approximately linear
  - CMB @ *a* = 0.001 ➔ density fluctuations ~ 0.001
    - NOTE: density higher in past, but density fluctuations smaller!

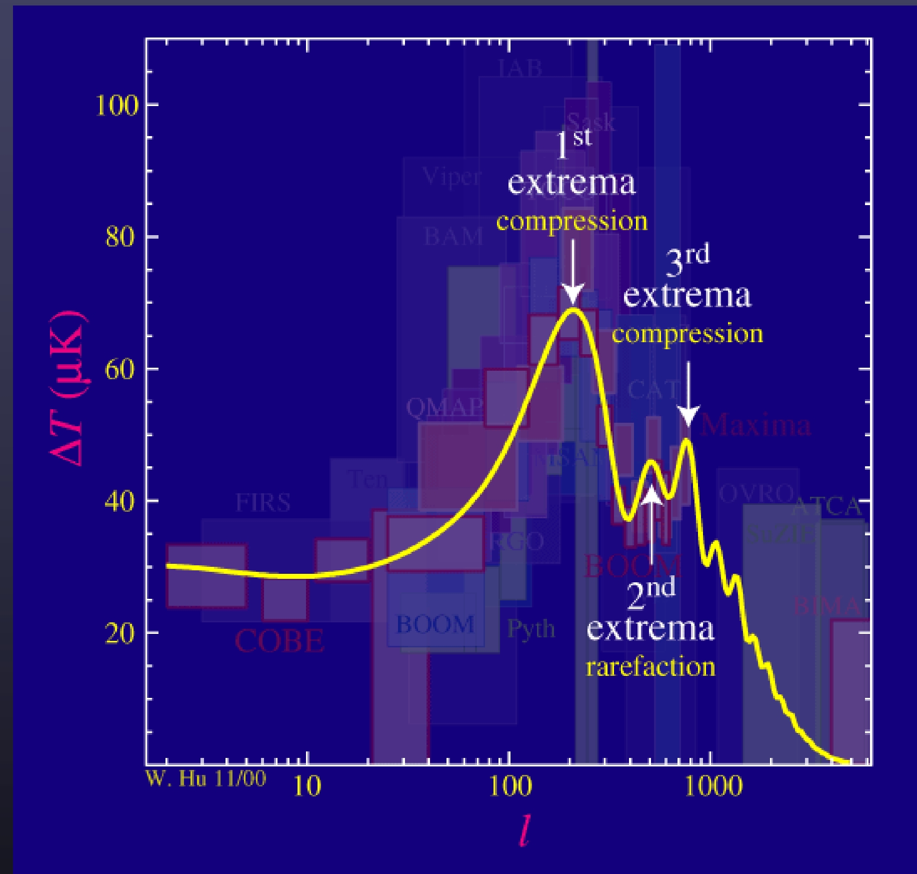


Courtesy Andrei Kravtsov – *http://cosmicweb.uchicago.edu*

CMB Power Spectrum Features

- dark matter + baryons ➔ acoustic oscillations



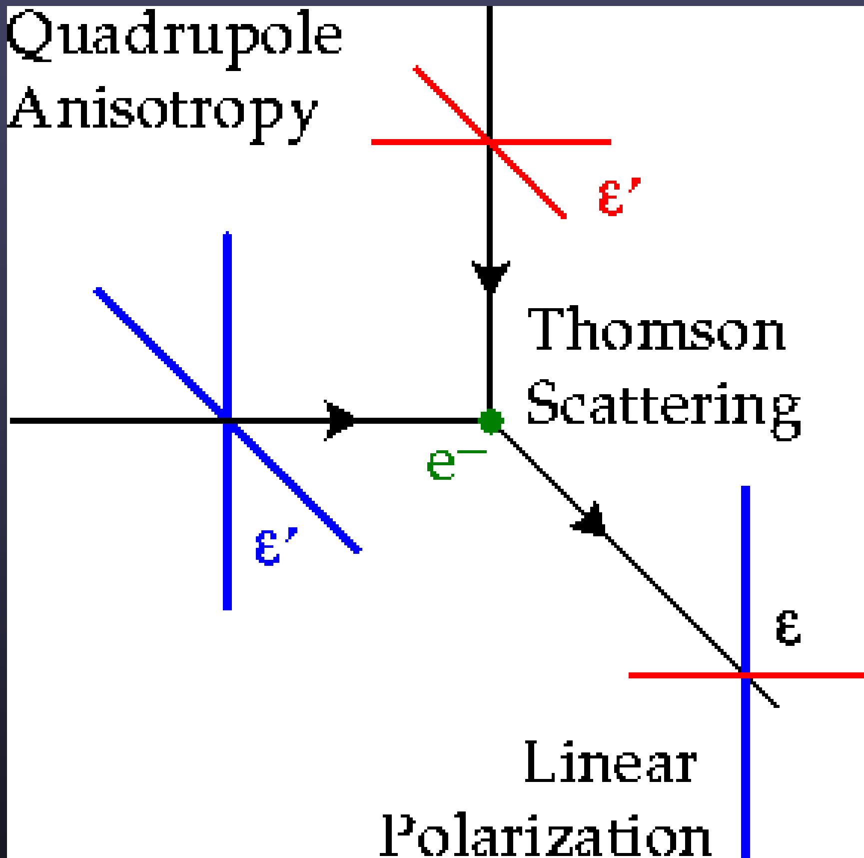


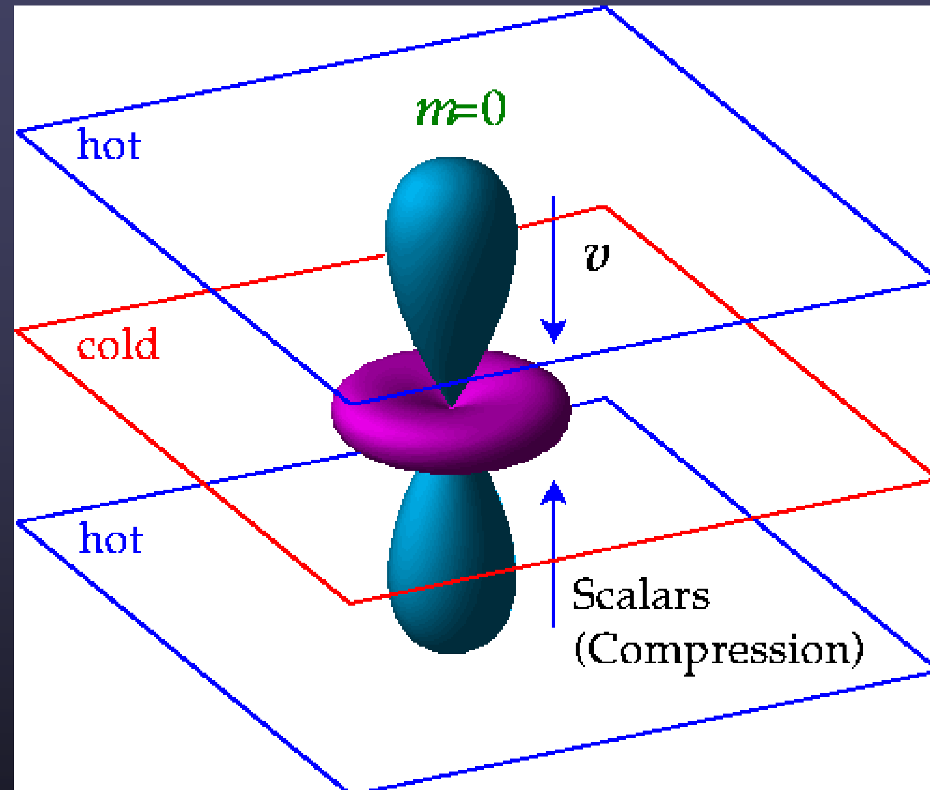
  – position of peaks measures angular scale of sound crossing at last scattering

Courtesy Wayne Hu – *http://background.uchicago.edu*

CMB Polarization

- Due to quadrupolar intensity field at scattering



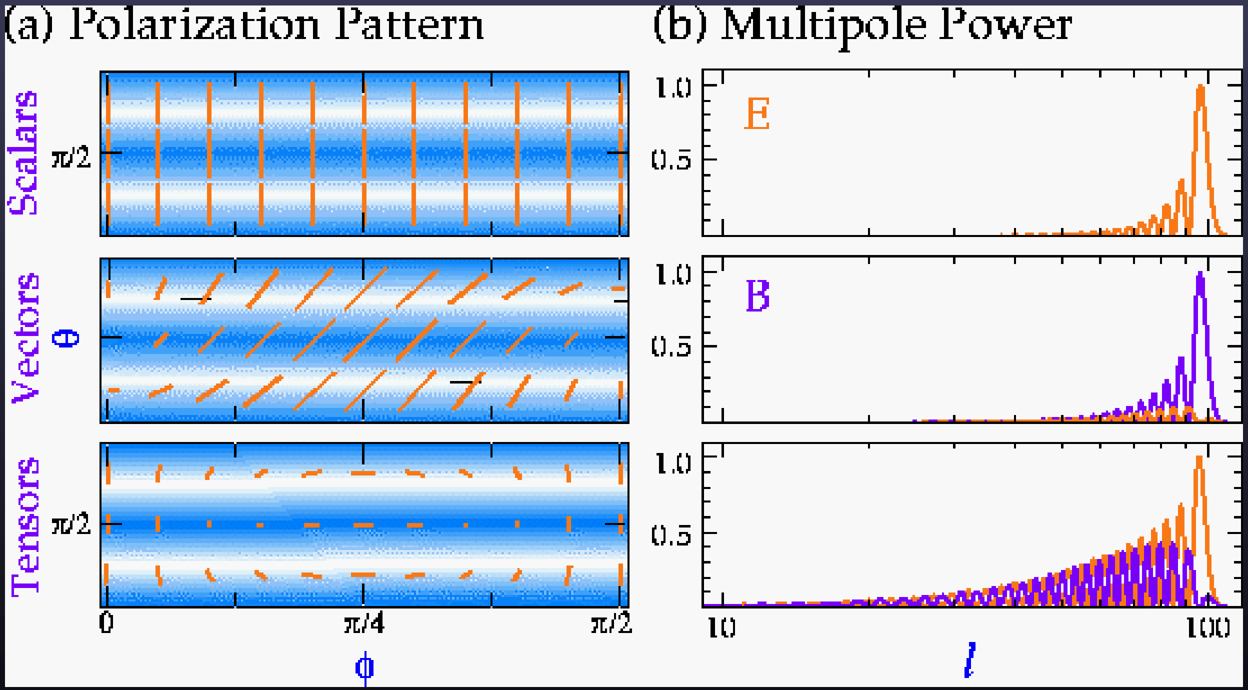


NOTE: polarization maximum when velocity is maximum (out of phase with compression maxima)

Courtesy Wayne Hu – *http://background.uchicago.edu*

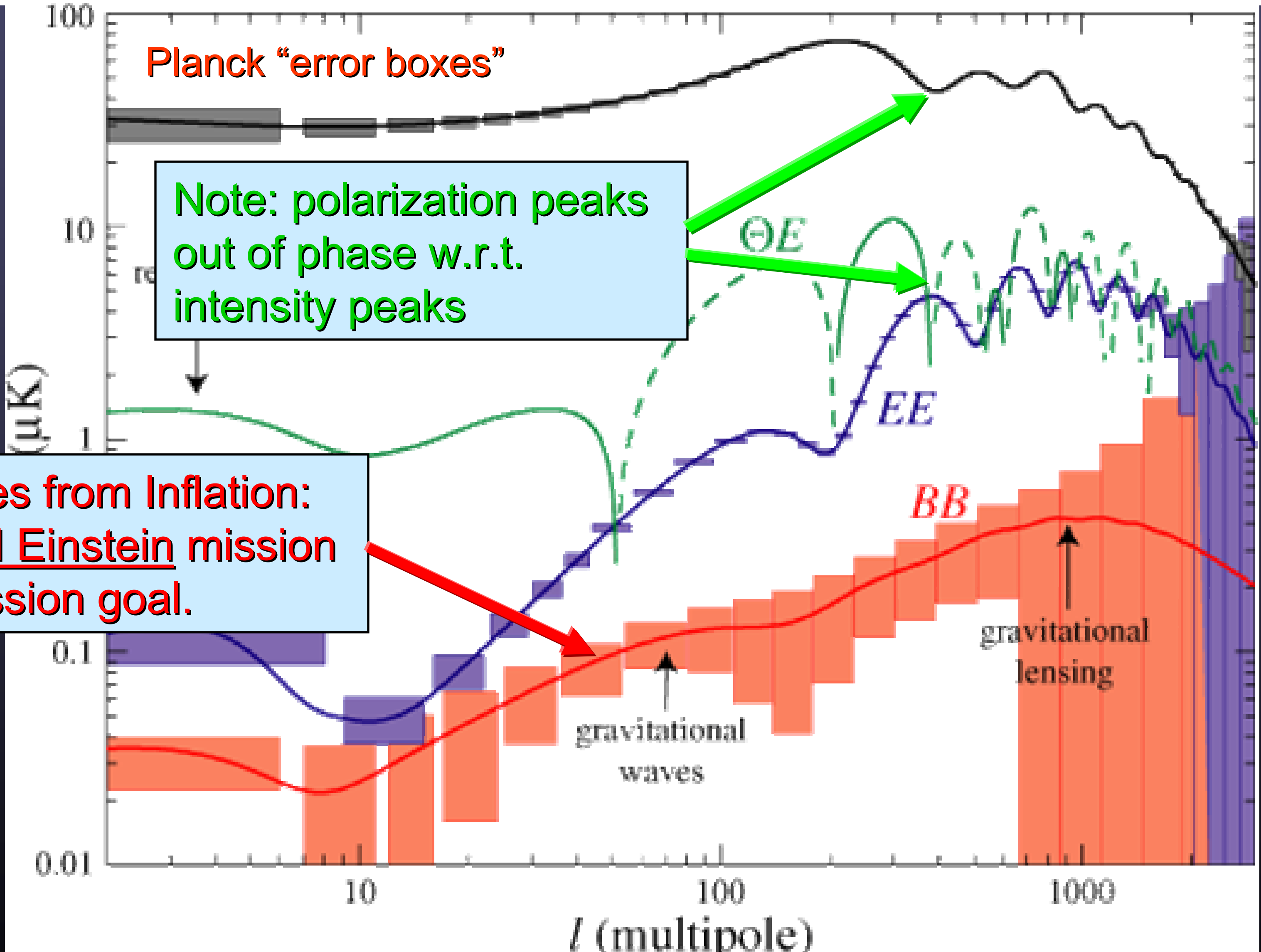
CMB Polarization

- E & B modes: even & odd parity modes on k-vector
  - E (even parity, “gradient”)
    - from scalar density fluctuations ➔ predominant!
  - B (odd parity, “curl”)
    - from gravity wave tensor modes, or secondaries



Courtesy Wayne Hu – *http://background.uchicago.edu*

Polarization Power Spectrum

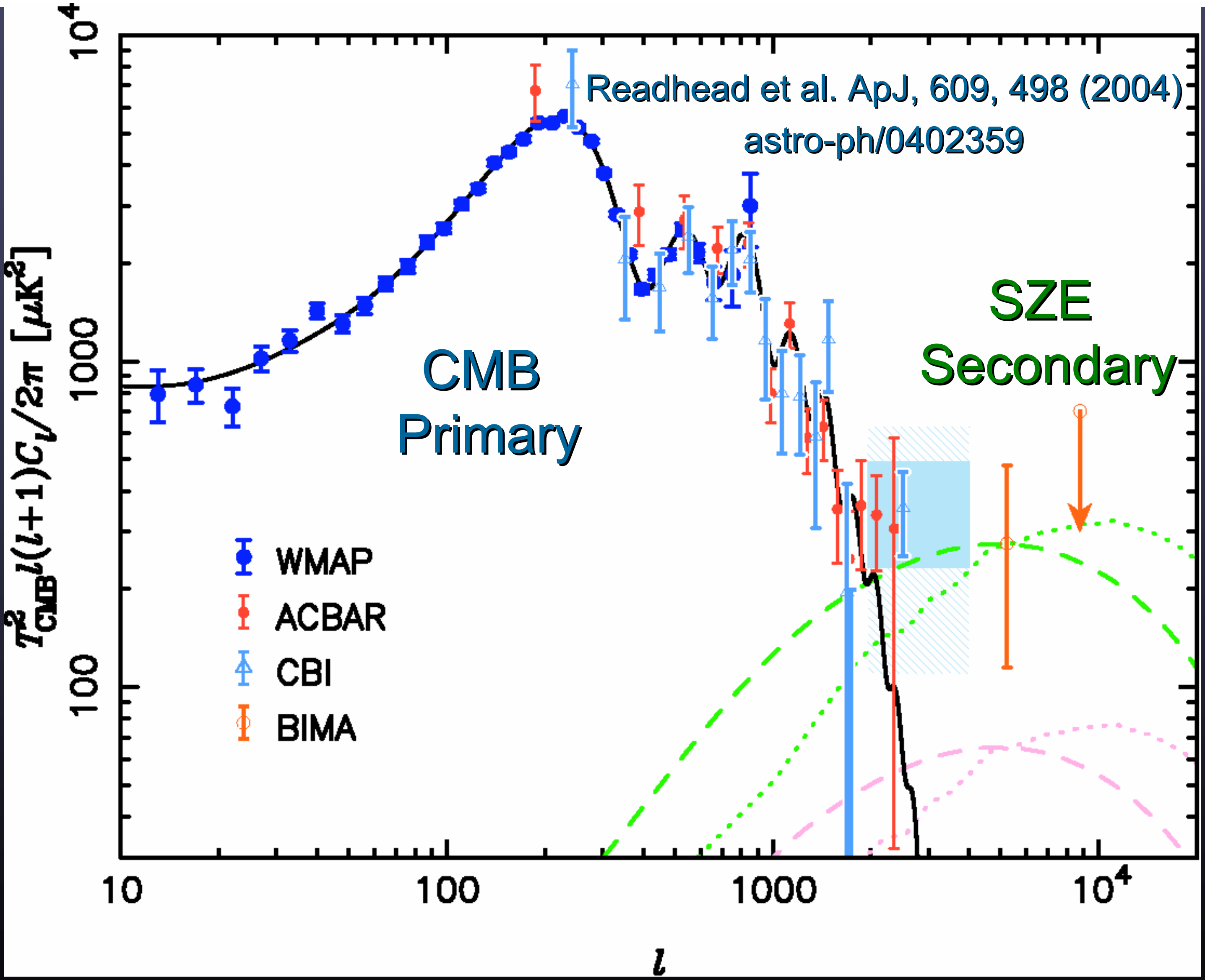


Note: polarization peaks out of phase w.r.t. intensity peaks

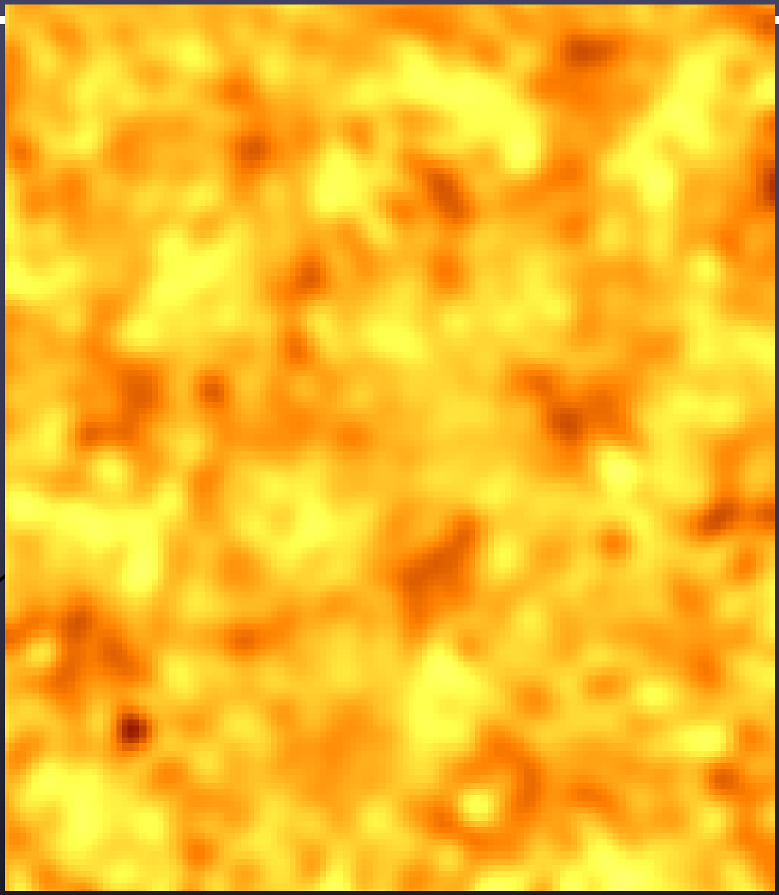
B-modes from Inflation: Beyond Einstein mission key mission goal.

Hu & Dodelson *ARAA 2002*

CBI 2000+2001, WMAP, ACBAR, BIMA



CBI Mosaic Observation



THE PILLARS OF INFLATION

1) super-horizon (>2°) anisotropies
2) acoustic peaks and harmonic pattern (~1°)
3) damping tail (<10')
4) Gaussianity
5) secondary anisotropies
6) polarization
7) gravity waves

But … to do this we need to measure a signal which is 3×10^7 times weaker than the typical noise!

The CMB measures these fundamental constants of cosmology:

Ω_k	Ω_b	Ω_{cdm}	n_s	Ω_Λ	Ω_m	h	τ
geometry of the universe	baryonic fraction protons, neutrons	cold dark matter not protons and neutrons	primordial fluctuation spectrum	dark energy negative pressure of space	matter fraction	Hubble Constant size & age of the universe	optical depth to last scattering of cmb

The analysis problem

The problem (rephrased)…

- Observe the cosmic microwave background (CMB) over a subset of the celestial sphere
- Take into account the observational process including instrumental response (beam, distortions, noise)
- Determine statistical properties (correlation function, angular power spectrum)
- Statistically infer constraints on fundamental cosmological parameters

The problem (for statisticians)…

- From Ben Wandelt (astro-ph/0401622):
  - CMB is an isotropic Gaussian random field *s* on the sphere
  - cosmological parameters $\Theta = \{ \theta_i \;\; i=1,..,n \}$ are related in a non-linear way to the spatial covariance structure $S = \langle s\, s^T \rangle$ of the field
  - observed as a sampled, noisy, filtered and censored/polluted measurement
  - analysis task is two-fold:
    - infer the covariance structure *S* of the field *s*
    - infer the parameters Θ

The problem (mathematical)…

- Time-ordered data (TOD):

$$d = A(s + f) + n$$

- Intrinsic signal *s* over (pixellated spherical) sky
  - signal $s = T$ (CMB temperature field) or polarization
  - spherical harmonic transform $$a_{\ell m} = \int d^2\hat{\mathbf{n}}\, Y_{\ell m}(\hat{\mathbf{n}})\, T(\hat{\mathbf{n}})$$
  - power spectrum $$\langle a_{\ell m} a^*_{\ell m} \rangle = C_\ell$$
- Foregrounds *f* (non-thermal spectral signature)
- Noise n (in time domain) with covariance

$$< n\, n^T > = N$$

Bayesian inference

- set up inverse problem using Bayes' theorem

$$P(f,s,C_l,\Theta|d)\, P(d) = P(d|f,s,C_l,\Theta)\, P(f,s,C_l,\Theta)$$

- given "prior probabilities" and dependencies

$$P(f,s,C_l,\Theta) = P(f)\, P(s|C_l)\, P(C_l|\Theta)\, P(\Theta)$$

- and likelihood of data given model

$$\mathcal{L}(d) = P(d|f,s,C_l,\Theta) \approx P(d|f,C_l)$$

Maximum Likelihood Estimate (MLE)

- maximize the likelihood

$$\mathcal{L}(C_\ell \mid \mathbf{d}) = \frac{\exp\left[-\frac{1}{2}\mathbf{d}^T\left(\mathbf{S}(C_\ell)+\mathbf{N}\right)^{-1}\mathbf{d}\right]}{\left[(2\pi)^{N_d}\det\left(\mathbf{S}(C_\ell)+\mathbf{N}\right)\right]^{1/2}}$$

  - note: the exponential term is $\chi^2/2$ (quadratic = easy!)
  - but: the determinant is expensive!

- covariance matrices
  - S and N are covariances between TOD or map pixels
  - may not be sparse (size N_d^2)

The problem (breakdown)…

- Mapmaking
  - evaluate P(s|d), derive map $m=s+f$
- Power Spectrum
  - evaluate P(C_l|d), often via P(C_l|m)
- Parameter Estimation
  - evaluate P(Θ|d), usually via P(Θ|C_l)
- basis for serial pipeline (2 or 3 steps)

Mapmaking

- often helps to turn TOD into maps (reduce size)
- MLE → optimal beam-deconvolved map for $m=s+f$

$$m = (A^T N^{-1} A)^{-1} A^T N^{-1} d$$

- noise in map

$$N_m = (A^T N^{-1} A)^{-1}$$

- problem: A has non-trivial structure and is often ill-conditioned (or singular!)
- solution: regularization (e.g. factor A=BG)

Power spectrum estimation

- for perfect data (all sky, no noise), estimator is trivial:

$$a_{\ell m} = \int d^2\hat{\mathbf{n}}\, Y_{\ell m}(\hat{\mathbf{n}})\, T(\hat{\mathbf{n}})$$

$$\hat{C}_\ell = \frac{\sum_m |a_{\ell m}|^2}{2\ell + 1}$$

  - real data requires MLE or equivalent process
- MLE: the determinant is costly!
  - almost all real methods use some “lossy” procedure
- Issue: cosmic variance
  - only one sky available to observe!
  - only $2l$+1 “m” values at each l, limits low l precision
  - WMAP limited for l < 354, will not improve!

The problem (size)…

- matrix operations are $\mathcal{O}(N^3)$ – what is N?
  - TOD: N_d is number of data samples (very large!)
  - maps: N_m is number of pixels (can be large)
  - note: ideally should be number of *independent* samples/pixels
    - compression (even lossy) is desirable!
- example: WMAP vs. Planck
  - WMAP
    - TOD: ~10^8 samples? (1-yr = 17 GB)
    - map: 3145728 HEALpix pixels (0.23°-0.93°) for 5 bands
  - Planck
    - TOD: ~10^{10} samples? ~1 TB?
    - map: 10× WMAP! ~3×10^7 pixels, for 100 detectors, 10chan!
  - N~10^7, so 10^{21} ops for one likelihood evaluation
  - At 10^{10} ops/s and $\pi\times10^7$s/yr: **1000s of CPU years**

Parameter estimation

- Multi-dimensional parameter space
  - 10-20 parameters
  - coupled & degenerate (e.g. Ωh^2)
- brute-force MLE prohibitive
  - but do need to know likelihood surface
  - or frequentist approach?
- sampling
  - do have some idea of probability distributions
  - Markov-Chain Monte Carlo (MCMC)
    - e.g. COSMOMC (Lewis & Bridle
    - chains optimally explore parameter space
    - active area of research!

The problem (summary)…

- Mapmaking
  - large TOD sets (>1 TB, >10^{10} samples)
  - complicated, asymmetric beams (sidelobes)
  - complicated scan patterns
- Power Spectrum
  - large maps (10-100 MB, >10^7 pixels)
  - spherical sky geometry (spherical transforms)
  - complicated covariance matrices (non-sparse, ~n_{pix}^2!)
- Parameter Estimation
  - 10-20 parameters in "new standard model"
  - degeneracies between parameters
  - incorporation of prior information & other experiments

The solutions (sort of)…

- Mapmaking
  - parallel computing (I/O in particular)
  - approximate and Monte-Carlo methods
  - optimal scan strategies
- Power Spectrum
  - quadratic estimators, Monte-Carlo (e.g. MASTER/FASTER)
  - fast spherical harmonic transforms & convolution
  - compression of data, parallelization of linear algebra
- Parameter Estimation
  - sampling methods (MCMC, Gibbs samplers)
  - fast predictors (CMBFast)

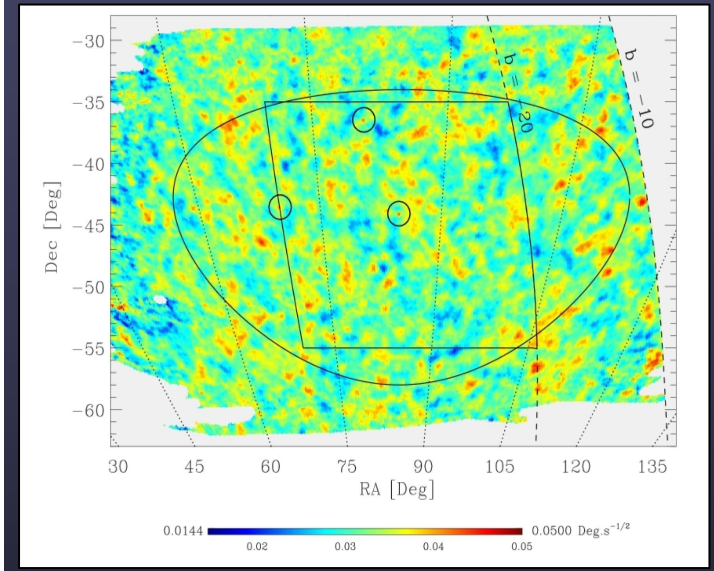
Talking point: MCMC &c

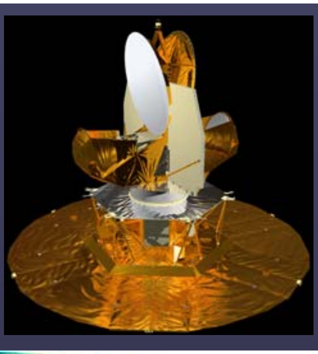
- Monte-Carlo methods are being developed:
  - MCMC for parameter estimation
  - MASTER for Boomerang power spectrum analysis
  - MAGIC Gibbs sampler for params+maps from TOD
- Possible uses:
  - replace Bond,Jaffe,Knox MLE for CBI analysis (see below)
  - mapmaking (see short presentation by Urvashi)
  - other?

CMB Imaging

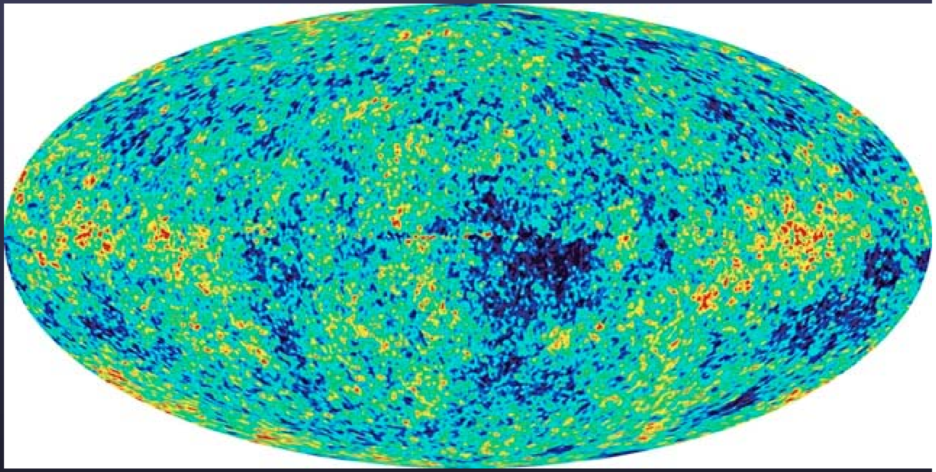
Images of the CMB

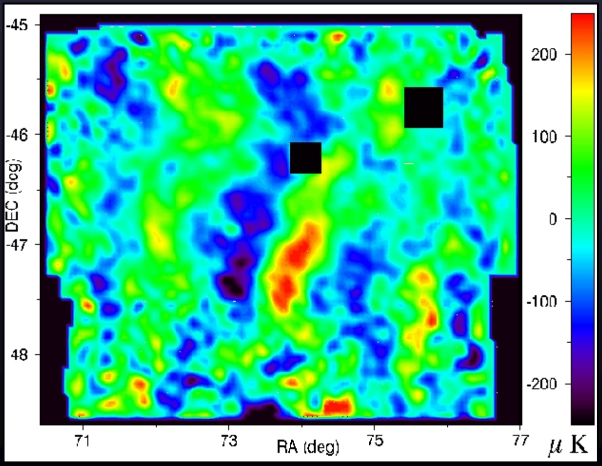
BOOMERANG Balloon



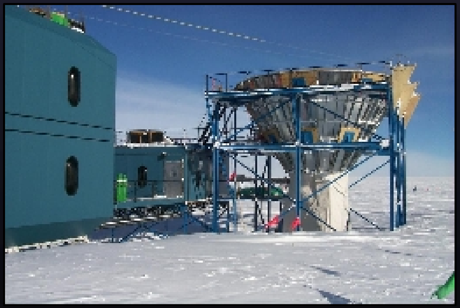


WMAP Satellite



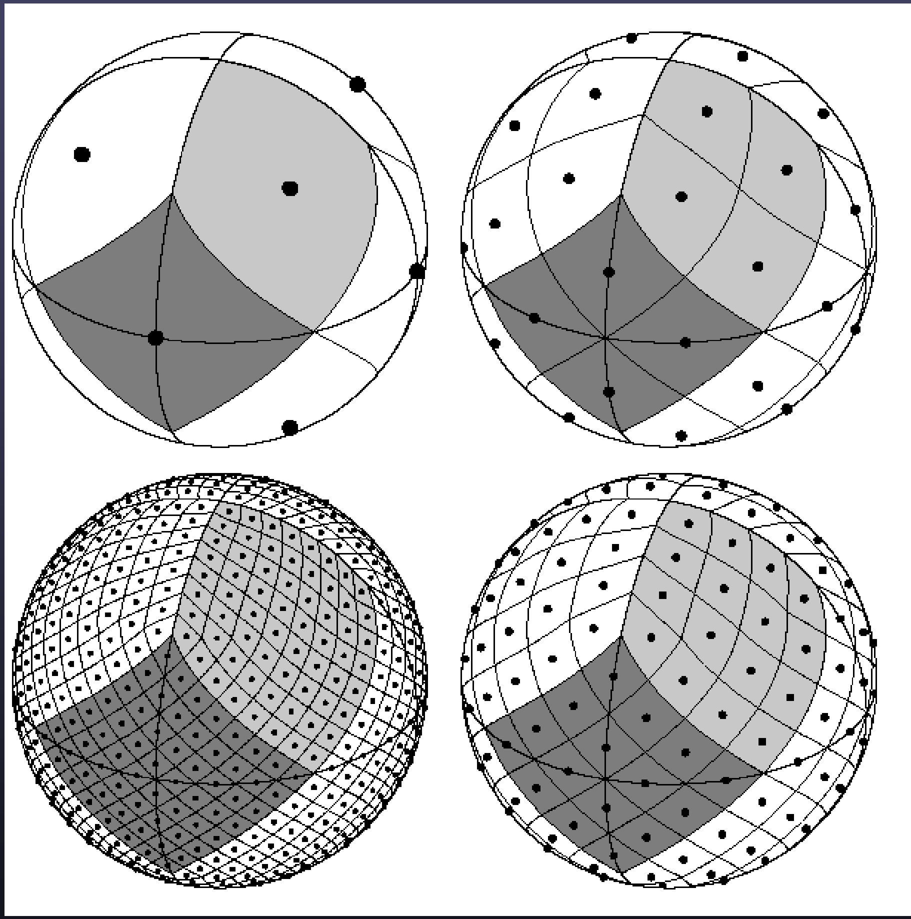


ACBAR South Pole



Spherical maps

- Need optimized map geometry and fast convolvers:

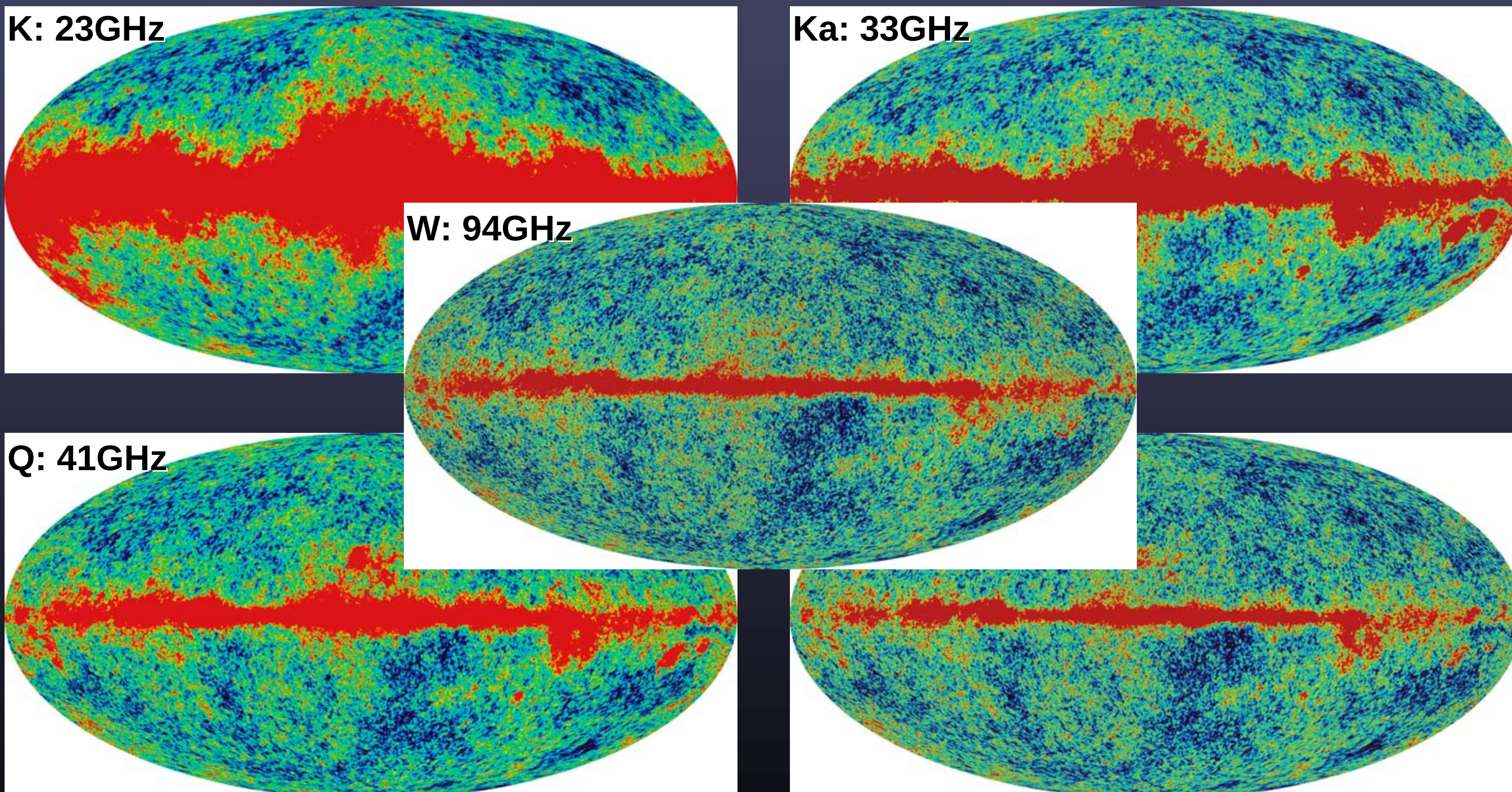


http://www.eso.org/science/healpix

– see Wandelt & Gorski (astro-ph/0008227) for convolution

WMAP: case study

- HEALpix maps:



CMB Interferometry

CMB Interferometers

- CMB issues:
  - Extremely low surface brightness fluctuations < 50 μK
  - Polarization <10%
  - Large monopole signal 3K, dipole 3 mK
  - No compact features, approximately Gaussian random field
  - Foregrounds both galactic & extragalactic
- Interferometry
  - Inherent differencing (fringe pattern), filtered images
  - Works in spatial Fourier domain
  - Element gain effect spread in image plane
  - Spherical sky can be ignored for small fields (but…)
  - Limited by need to correlate pairs of elements
  - Sensitivity requires compact arrays

The Cosmic Background Imager

- 13 90-cm Cassegrain antennas
  - 78 baselines
- 6-meter platform
  - Baselines 1m – 5.51m
- 10 1 GHz channels 26-36 GHz
  - HEMT amplifiers (NRAO)
  - Cryogenic 6K, Tsys 20 K
- Single polarization (R or L)
  - Polarizers from U. Chicago
- Analog correlators
  - 780 complex correlators
- Field-of-view 44 arcmin
  - Image noise 4 mJy/bm 900s
- Resolution 4.5 – 10 arcmin
- Rotatable platform

The CMB and Interferometry

- The sky can be uniquely described by spherical harmonics
  - CMB power spectra are described by multipole l
- For small (sub-radian) scales the spherical harmonics can be approximated by Fourier modes
  - The conjugate variables are (u,v) as in radio interferometry
  - The uv radius is given by $|\mathbf{u}| = l / 2\pi$
- An interferometer naturally measures the transform of the sky intensity in l space convolved with aperture

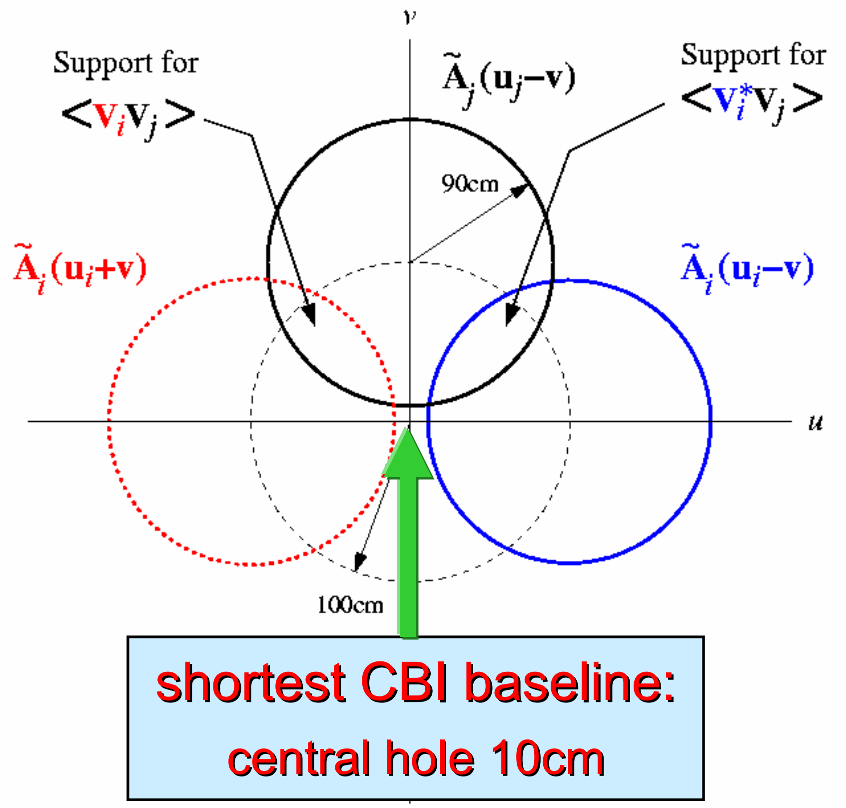
$$V(\mathbf{u}) = \int d^2\mathbf{x}\, A(\mathbf{x}-\mathbf{x}_p)\, I(\mathbf{x})\, e^{-2\pi i \mathbf{u}\cdot(\mathbf{x}-\mathbf{x}_p)} + \mathrm{e}$$

$$= \int d^2\mathbf{v}\, \tilde{A}(\mathbf{u}-\mathbf{v})\, \tilde{I}(\mathbf{v})\, e^{2\pi i \mathbf{v}\cdot\mathbf{x}_p} + \mathrm{e}$$

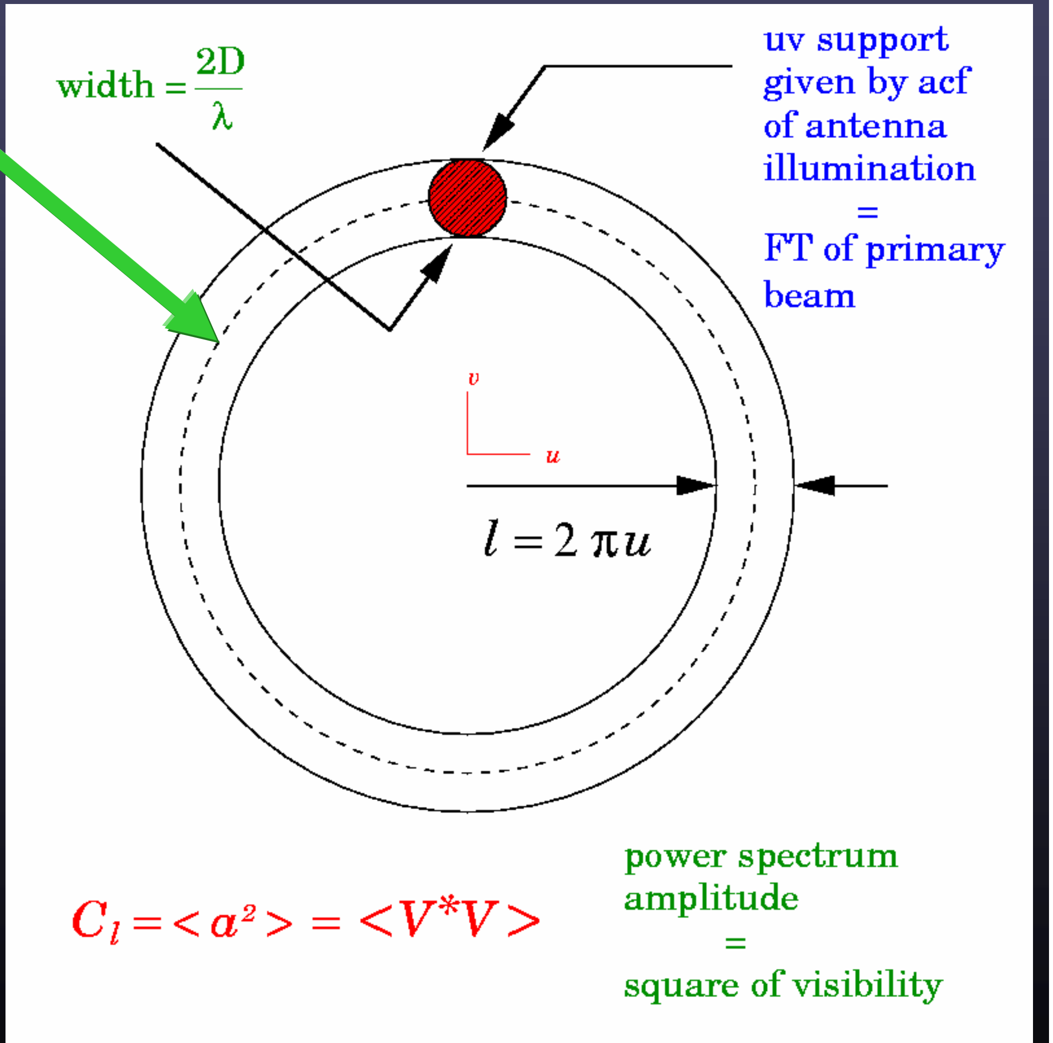
The *uv* plane

- The projected baseline length gives the angular scale

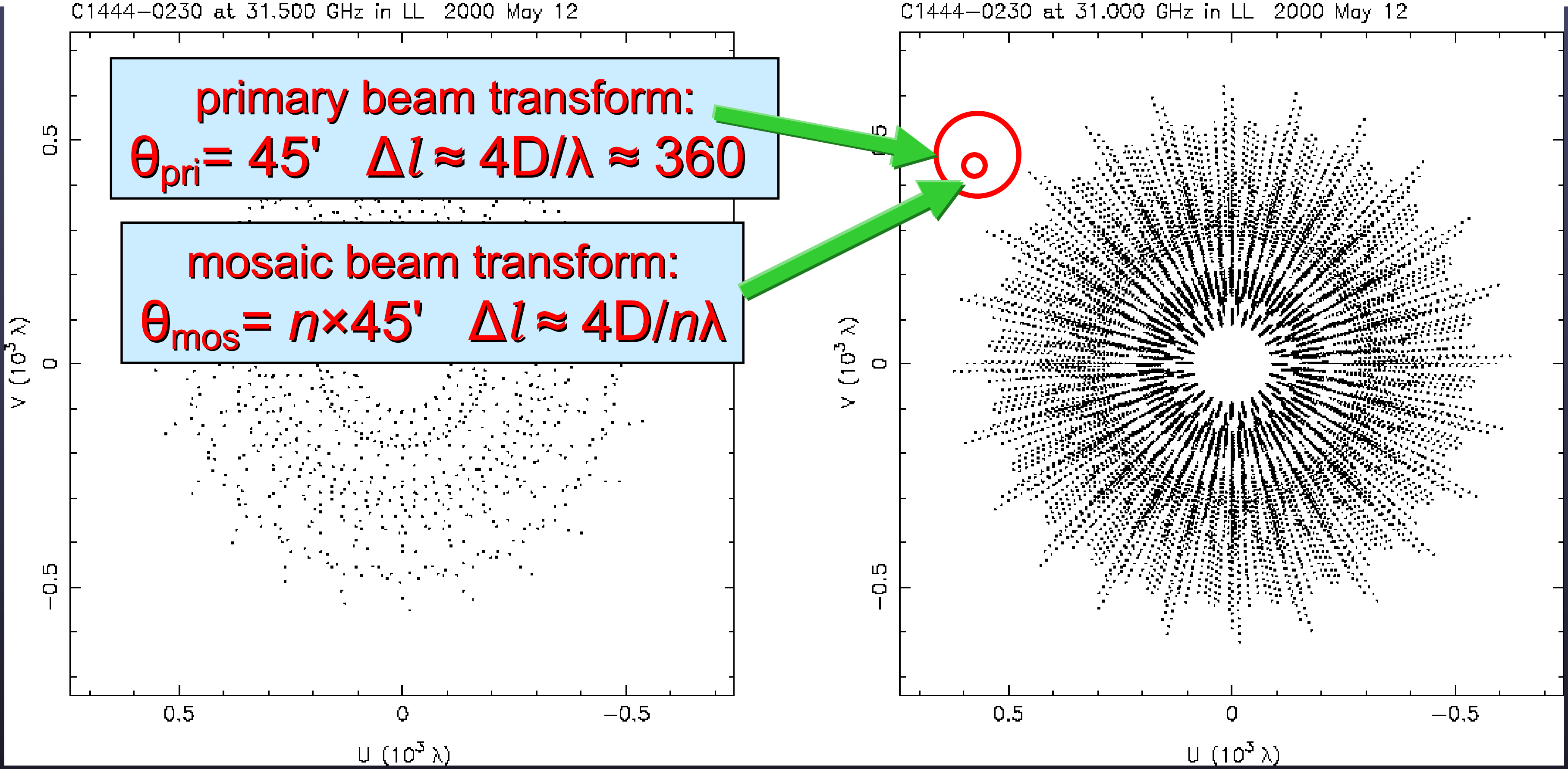
multipole:
$l = 2\pi B/\lambda = 2\pi|u_{ij}|$



shortest CBI baseline:
central hole 10cm



CBI Beam and *uv* coverage



- Over-sampled *uv*-plane
  - excellent PSF
  - allows fast gridded method (Myers et al. 2003)

Polarization Interferometry

- Observing with circularly polarized feeds (e.g. CBI):
  - correlation products RR, RL, LR, or LL from antenna pair
- Correlations to Stokes parameters (I,Q,U,V) :
- co-polar: RR = I + V LL = I – V
- cross-polar: $RL = [Q + i\,U]\,e^{-i2\Psi}$ $LR = [Q - i\,U]\,e^{i2\Psi}$
  - electric vector position angle EVPA = $\frac{1}{2}\tan^{-1}(U/Q)$
  - rotates with parallactic angle of detector Ψ on sky
- Stokes I,Q,U to E and B:
- $Q + i\,U = [E + i\,B]\,e^{i2\chi} \rightarrow RL = [E + i\,B]\,e^{i2(\chi-\Psi)}$
  - counter-rotates with wave vector angle $\Psi = \frac{1}{2}\tan^{-1}(v/u)$
- visibility covariances:
  - <RR RR*> = TT <RR RL*> = TE <RL RL*> = EE + BB

Interferometry Equations

- Our co-polar and cross-polar visibilities are:

$$V^{RR}_{ij\nu}(\mathbf{u}_{ij\nu}) = \int d^2\mathbf{v}\, P_{ij\nu}(\mathbf{v})\, \widetilde{T}(\mathbf{v}) + \mathrm{e}^{RR}_{ij\nu}$$

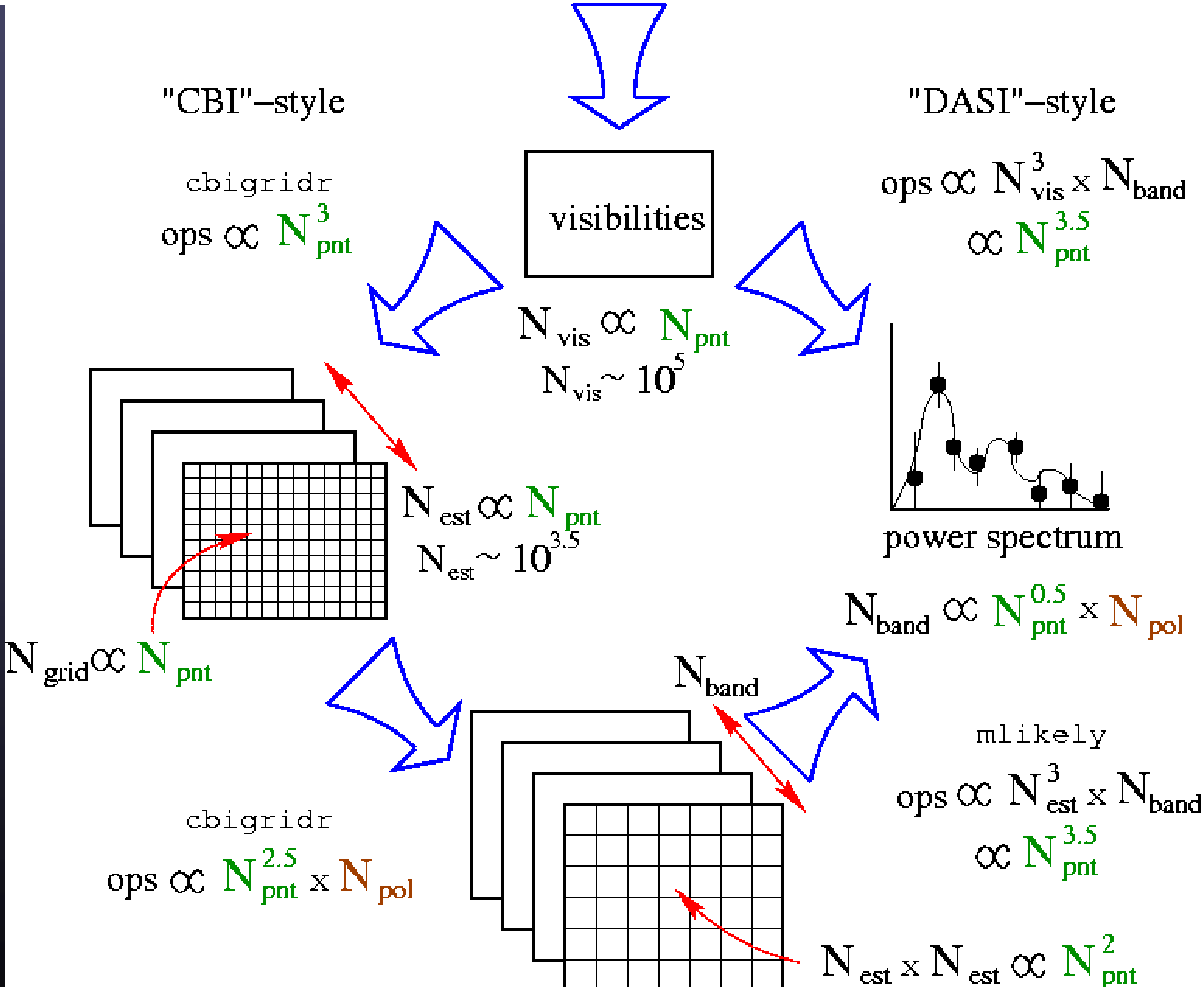
$$P_{ij\nu}(\mathbf{v}) = \widetilde{A}(\mathbf{u}_{ij\nu} - \mathbf{v})\, e^{2\pi i\, \mathbf{v}\cdot\mathbf{x}_p}$$

E & B response smeared by phase variation over aperture A

$$V^{RL}_{ij\nu}(\mathbf{u}_{ij\nu}) = \int d^2\mathbf{v}\, P_{ij\nu}(\mathbf{v})[\widetilde{E}(\mathbf{v}) + i\,\widetilde{B}(\mathbf{v})]\, e^{i2(\chi_{\mathrm{v}} - \psi_{ij})} + \mathrm{e}^{RL}_{ij\nu}$$

interferometer "directly" measures (Fourier transforms of) T,E & B!

The Computational Problem



Constraints & Projection

- Fit for CMB power spectrum bandpowers
- Terms for "known" effects
  - instrumental noise
  - residual source foreground
  - incorporate as "noise" matrices with known prefactors
- Terms for "unknown effects"
  - e.g. foreground sources with known positions
  - known structure in C
  - incorporate as "noise" matrices with large prefactors
  - equivalent to downweighting contaminated modes in data

$$\mathbf{C} = \mathbf{C}^{\mathrm{N}} + \sum_B q_B\, \mathbf{C}_B^{\mathrm{T}} + q_{\mathrm{src}}\, \mathbf{C}^{\mathrm{src}} + q_{\mathrm{res}}\, \mathbf{C}^{\mathrm{res}} + q_{\mathrm{scan}}\, \mathbf{C}^{\mathrm{scan}}$$

noise fitted projected

HM02: Another Approach

- Could also attempt reconstruction of Fourier plane
  - $v = \mathrm{P}\, t + e \quad \rightarrow \quad v = \mathrm{M}\, s + e$
- e.g. ML solution over $e = v - \mathrm{M}s$
  - $x = \mathrm{H}\, v = s + n \qquad \mathrm{H} = (\mathrm{M}^T\mathrm{N}^{-1}\mathrm{M})^{-1}\mathrm{M}^T\mathrm{N}^{-1} \qquad n = \mathrm{H}\, e$
- see Hobson & Maisinger 2002, MNRAS, 334, 569
  - applied to VSA data
  - same as optimal mapmaking!

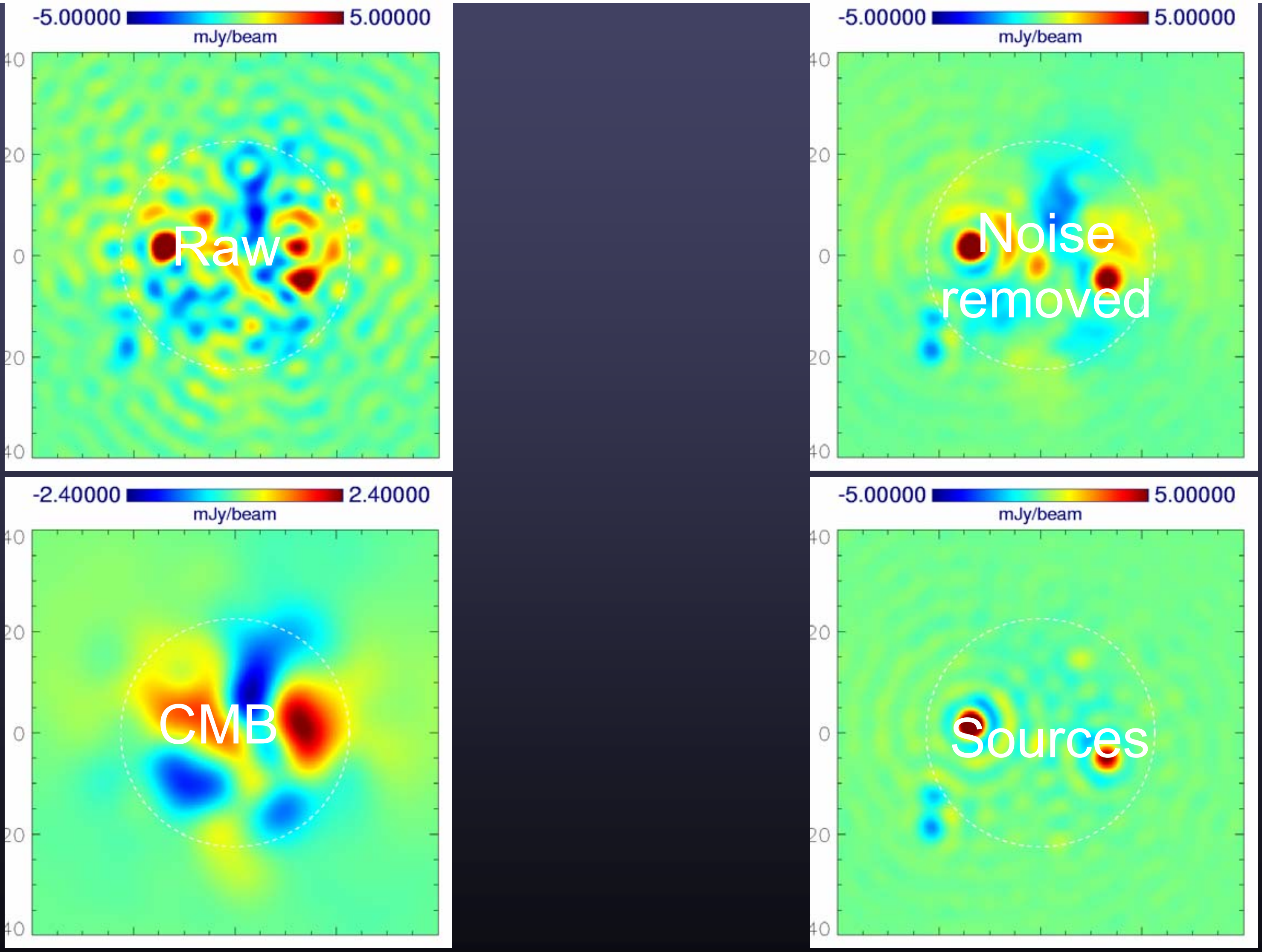
Wiener filtered images

- Covariance matrices can be applied as Wiener filter to gridded estimators

$$\mathbf{\Delta}^{\mathrm{X}} = \mathbf{C}^{\mathrm{X}}\,\mathbf{C}^{-1}\,\mathbf{\Delta}$$

- Estimators can be Fourier transformed back into filtered images
- Filters C^{X} can be tailored to pick out specific components
  - e.g. point sources, CMB, SZE
  - Just need to know the shape of the power spectrum

Example – Mock deep field



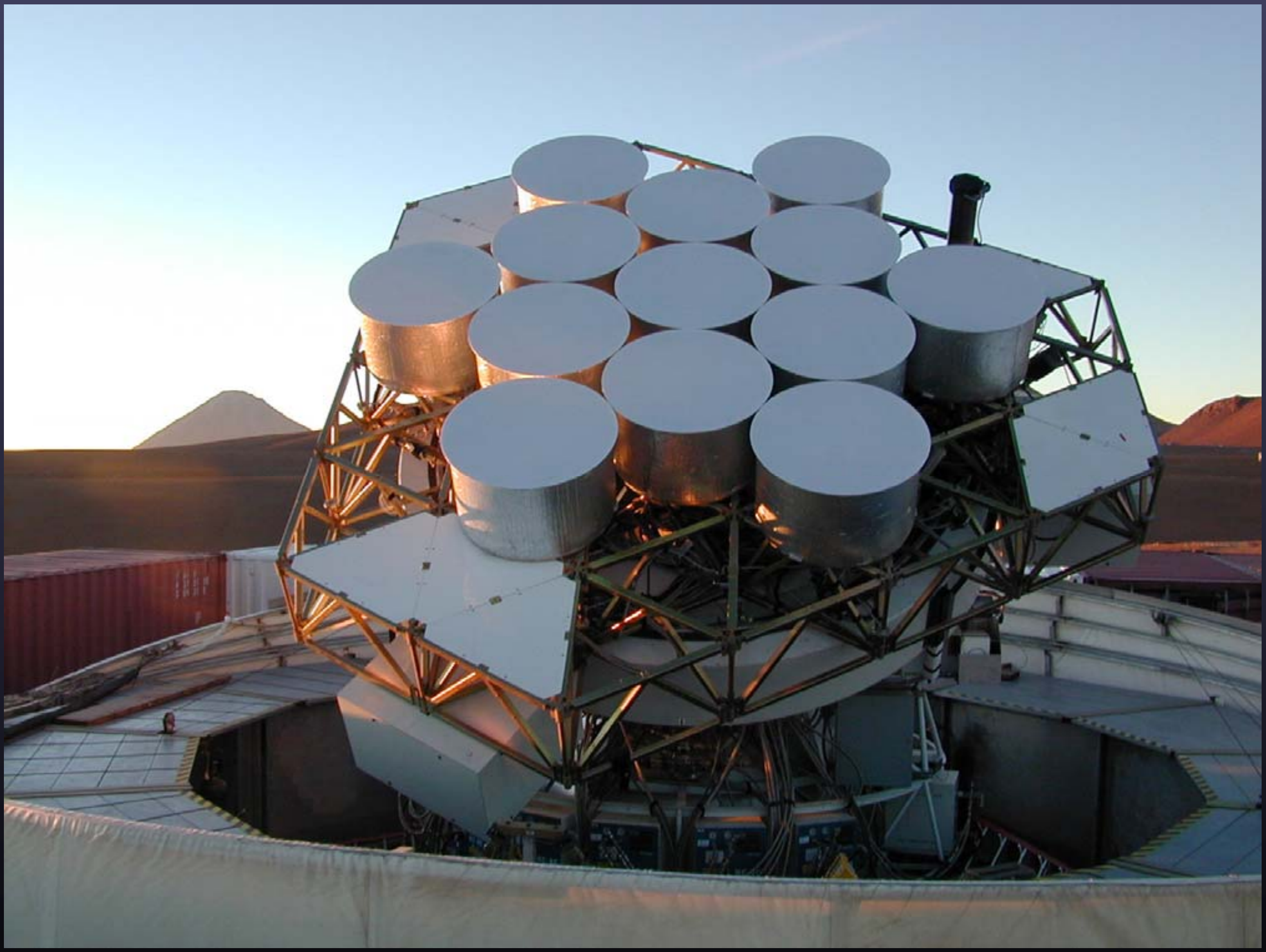
CBI Polarization New Results!

Brought to you by:

A. Readhead, T. Pearson, C. Dickinson (Caltech)
S. Myers, B. Mason (NRAO),
J. Sievers, C. Contaldi, J.R. Bond (CITA)
P. Altamirano, R. Bustos, C. Achermann (Chile)
& the CBI team!

CBI Current Polarization Data

- Observing since Sep 2002 (processed to May 2004)
  - compact configuration, maximum sensitivity



CBI Polarization Data Processing

- Massive data processing exercise
  - 4 mosaics, 300 nights observing, >10^6 visibilities!
  - scan projection over 3.5° requires fine gridding
    - more than 10^4 gridded estimators
- Method: Myers et al. (2003)
  - gridded estimators + max. likelihood
  - used for TT in CBI 2001-2003 papers
- Parallel computing critical
  - both gridding and likelihood now parallelized using MPI
    - using 256 node/ 512 proc McKenzie cluster at **CITA**
    - 2.4 GHz Intel Xeons, gigabit ethernet, 1.2 Tflops!
    - currently 4-6 hours per full run (!)
    - current limitation 1 GB memory per node

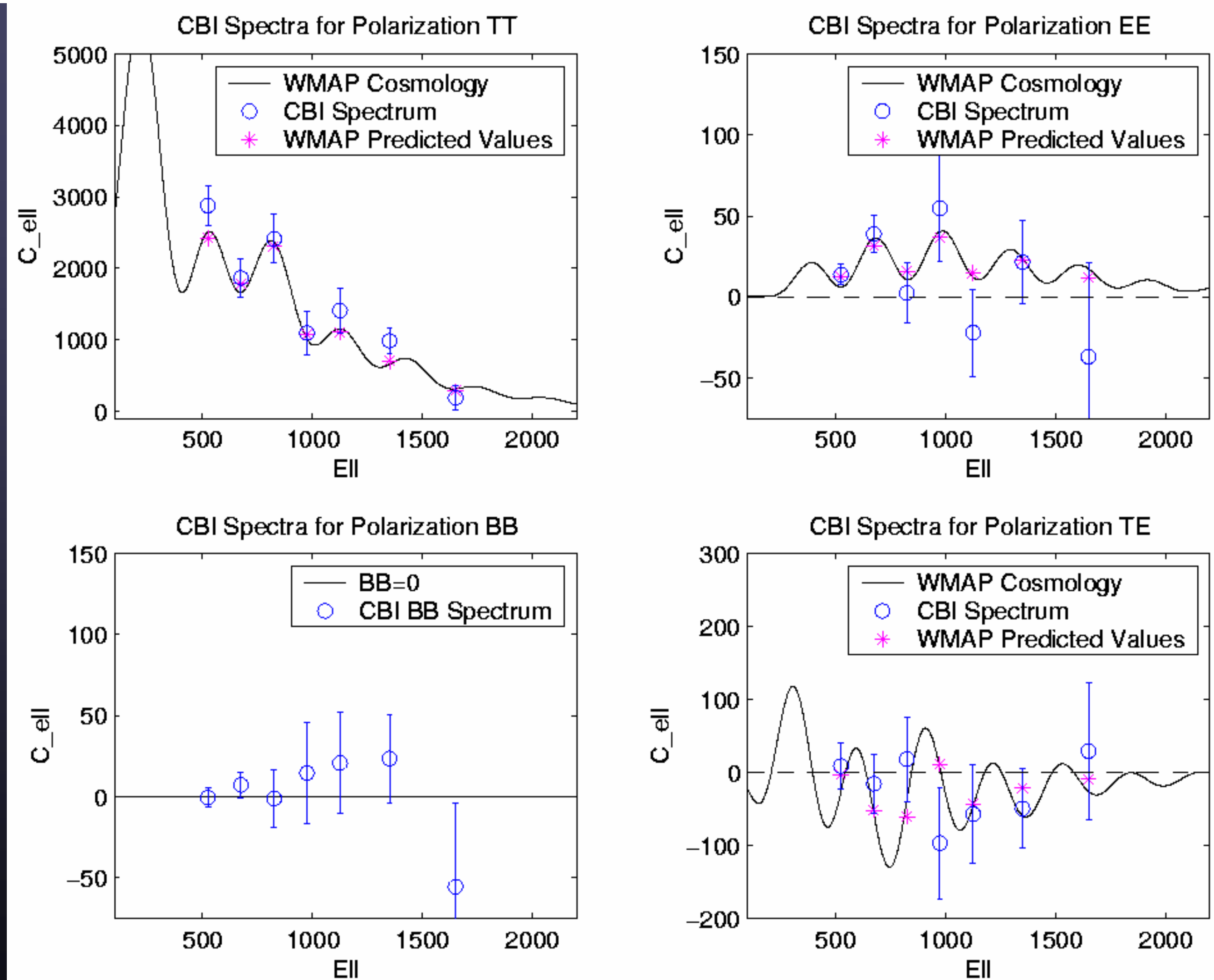
$$\Delta_i = \sum_k Q_{ik} V_k$$

Matched filter gridding kernel

$$Q_{ik} = \frac{w_k}{z_i} \tilde{A}_k^*(\mathbf{u}_k - \mathbf{u}_i)\, e^{-2\pi i \mathbf{u}_i \cdot \mathbf{x}_k}$$

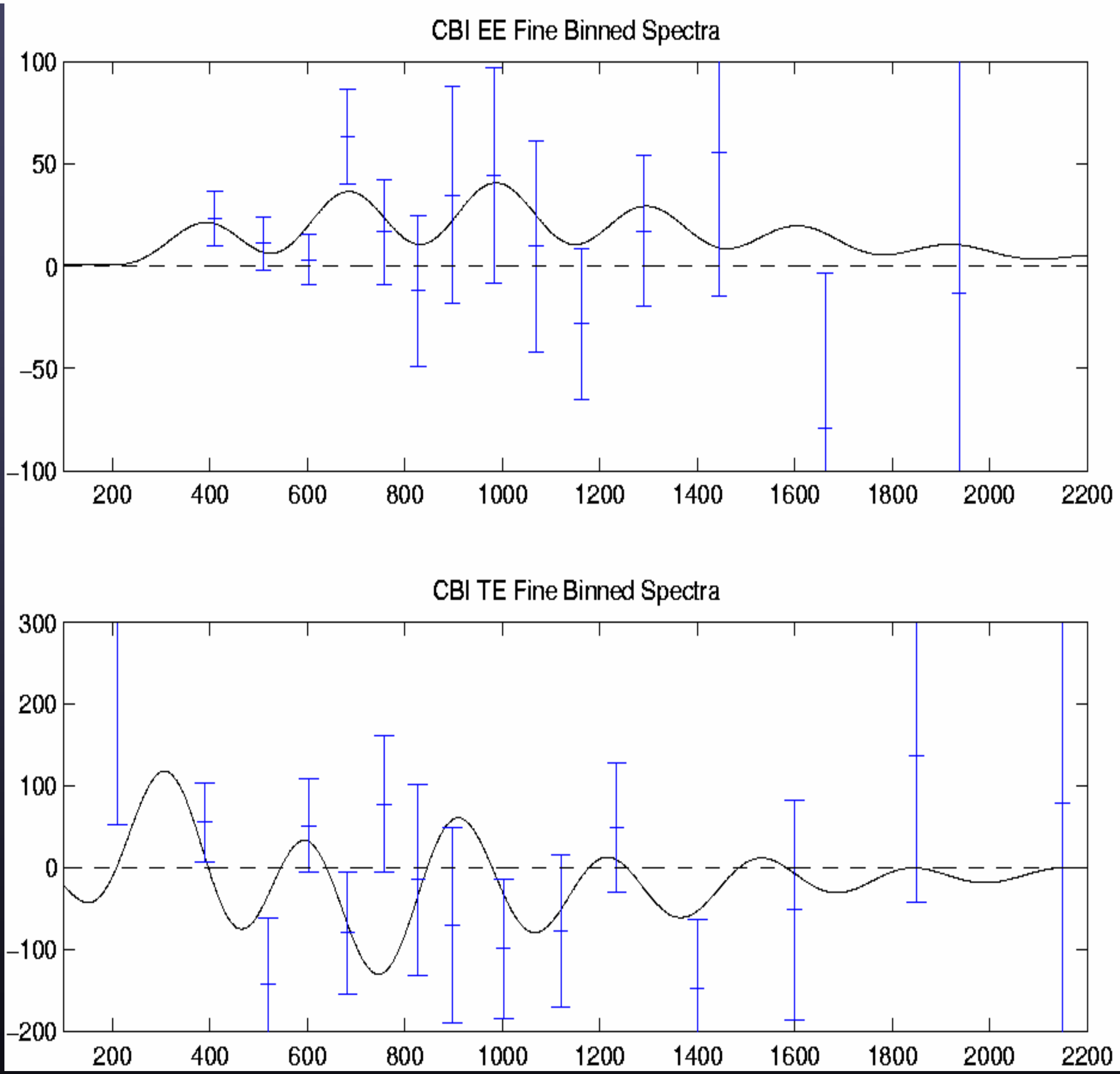
mosaicing phase factor

New: CBI Polarization Power Spectra



- matched to peaks ($\Delta l \approx 150$)

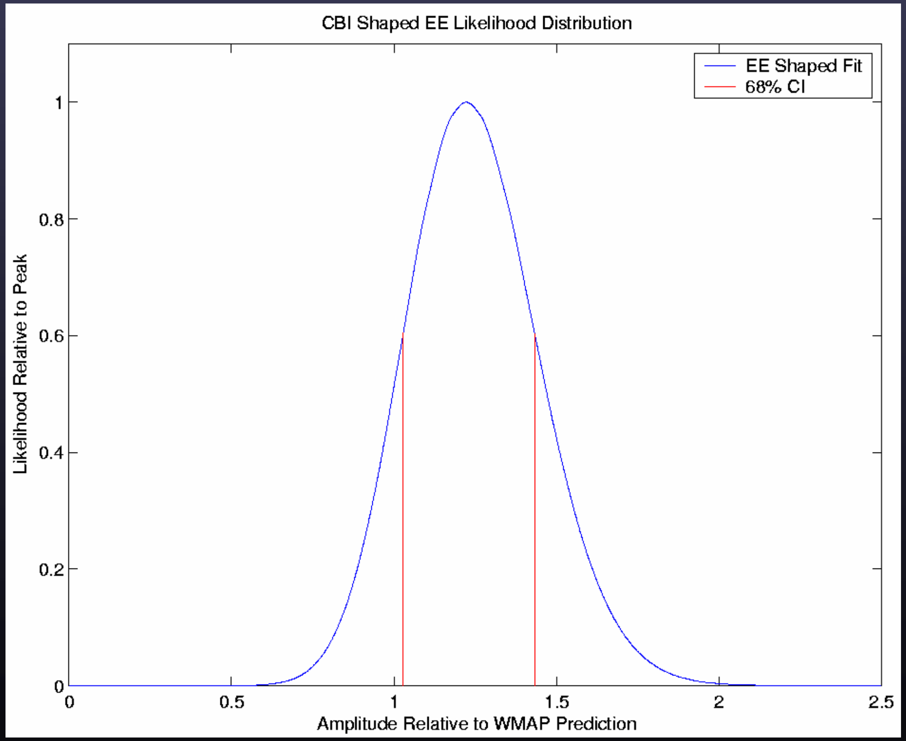
New: CBI Polarization Power Spectra



- finer resolution ($\Delta l \approx 75$ at $l = 800$)

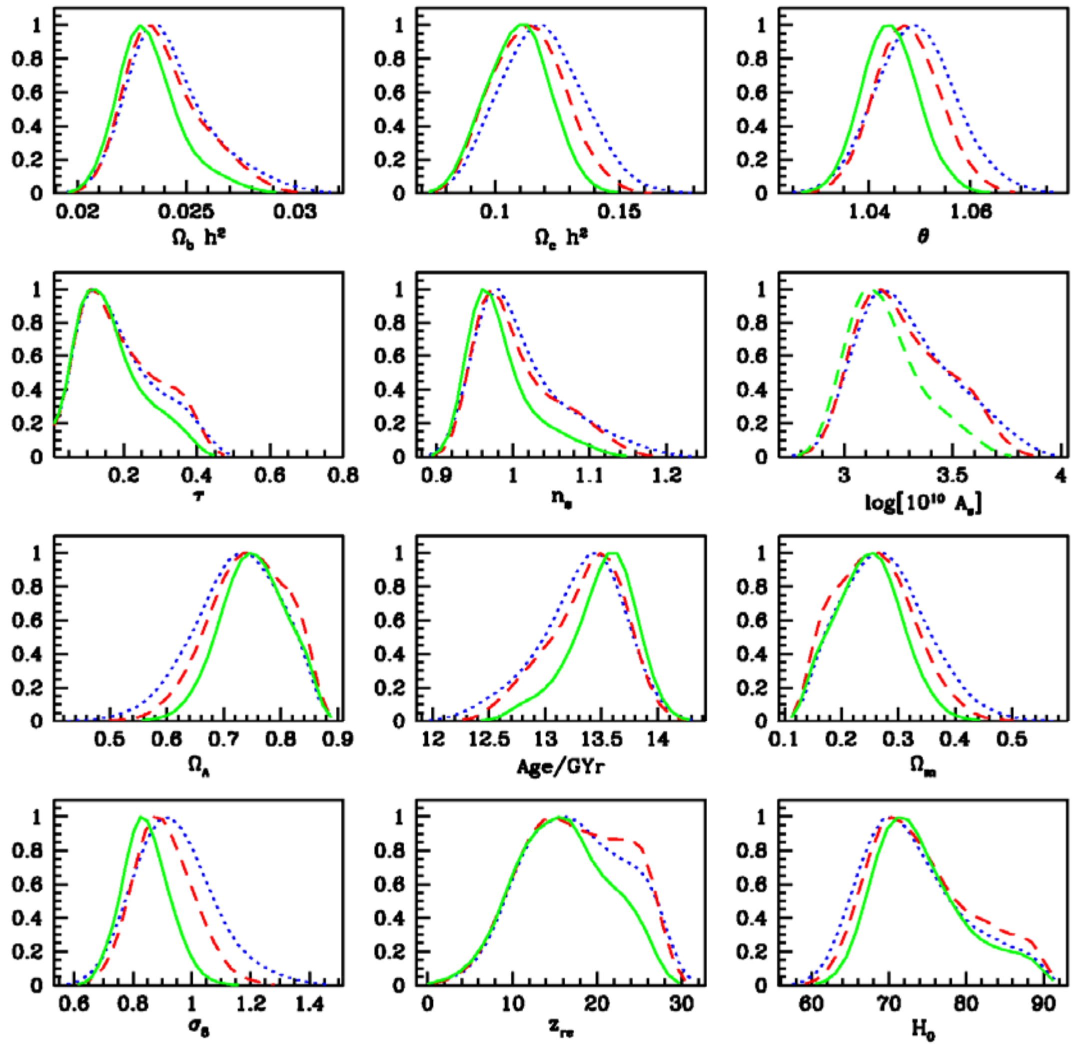
New: Shaped C_l fits

- Use WMAP'03 best-fit Cl in signal covariance matrix
  - bandpower is then relative to fiducial power spectrum
  - compute for single band encompassing all *l*s
- Results for CBI data (sources projected from TT only)
  - q_B = 1.22 ± 0.21 (68%)
  - EE likelihood vs. zero : equivalent significance 8.9 σ



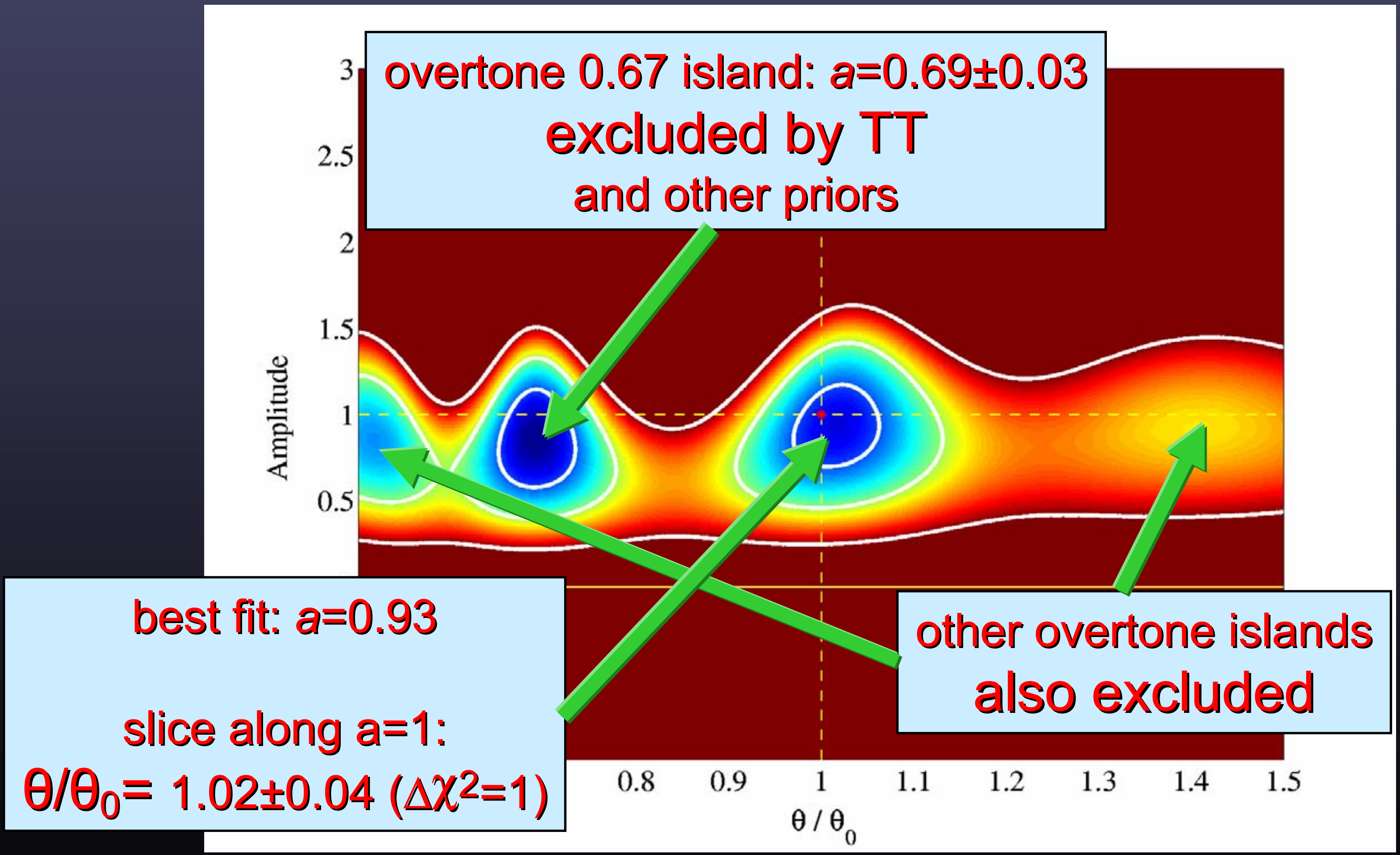
New: CBI Polarization Parameters

- use fine bins ($\Delta l = 75$) + window functions ($\Delta l = 25$)
- cosmological models vs. data using MCMC
  - modified COSMOMC (Lewis & Bridle 2002)
- Include:
  - WMAP TT & TE
  - WMAP + CBI'04 TT & EE (Readhead et al. 2004b = new!)
  - WMAP + CBI'04 TT & EE $l<1000$ + CBI'02 TT $l>1000$ (Readhead et al. 2004a) [overlaps '04]

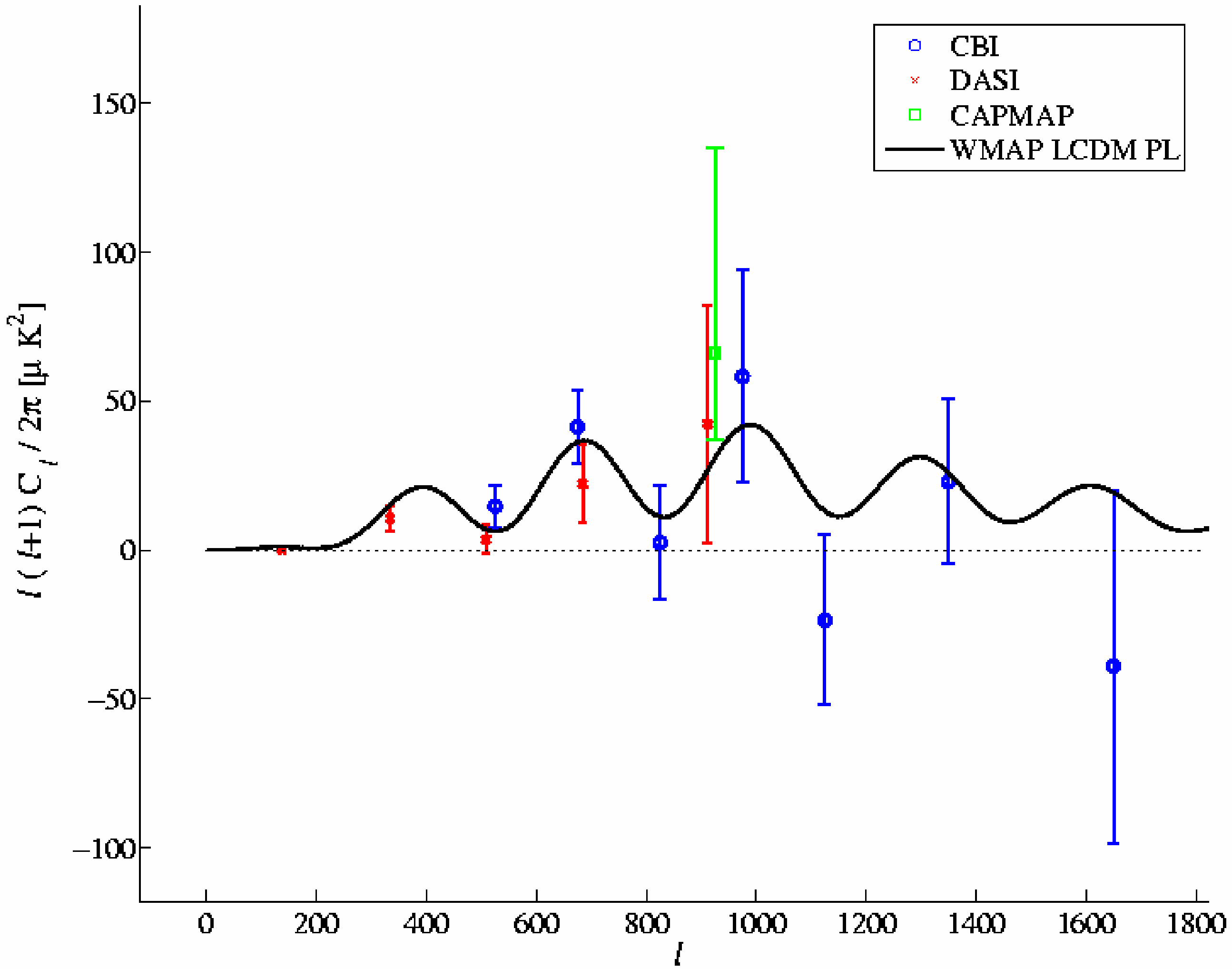


New: CBI EE Polarization Phase

- Scaling model: spectrum shifts by scaling l
  - allow amplitude *a* and scale θ to vary

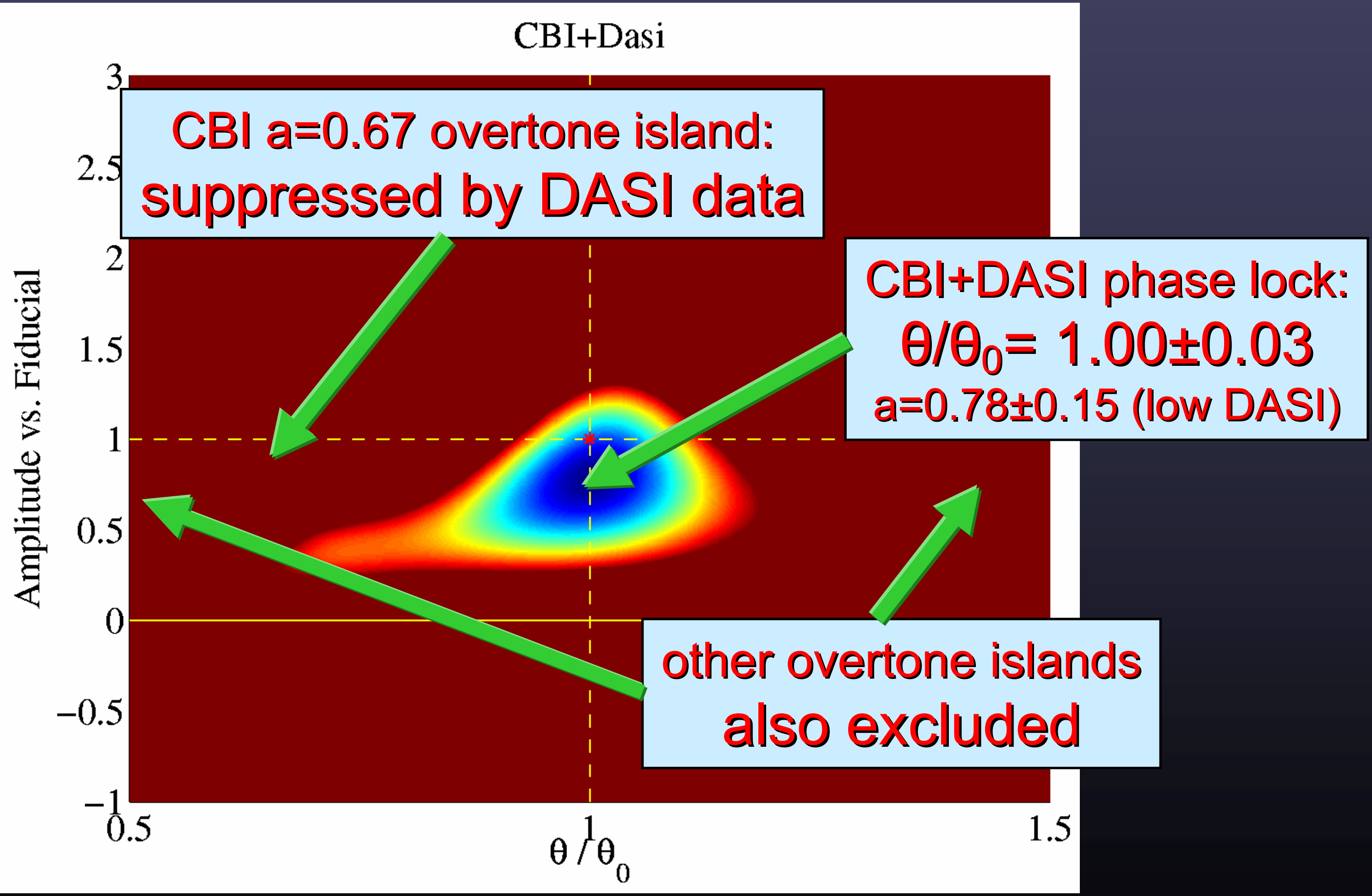


New: CBI, DASI, Capmap



New: CBI + DASI EE Phase

- Combined constraints on θ model:
  - DASI (Leitch et al. 2004) & CBI (Readhead et al. 2004)



CMB Imaging/Analysis Problems

- Time Stream Processing (e.g. calibration)
- Power Spectrum estimation for large datasets
  - MLE, approximate methods, efficient methods
  - optimal maps, efficient deconvolution
  - extraction of different components (foregrounds)
  - from PS to parameters (e.g. MCMC)
  - combined estimation (e.g. Gibbs sampling)
- Beyond the Power Spectrum
  - non-Gaussianity (test isotropy, amp & phase distributions)
  - bispectrum and beyond
- Other
  - "object" identification
  - topology
  - comparison of overlapping datasets

Selected References

- "Challenges" issues:
  - Wandelt 2004 (astro-ph/0401522)
  - Wandelt 2000 (astro-ph/0012416)
- Methods
  - CBI: Myers et al. 2002 (astro-ph/0205385)
  - maps: Armitage & Wandelt 2004 (astro-ph/0410092)
  - MAGIC: Wandelt 2004 (astro-ph/0401623)
  - MASTER: Hivon et al. 2001 (astro-ph/0105302)
- COSMOMC
  - Lewis & Bridle 2002 (astro-ph/0205436)
  - http://cosmologist.info/cosmomc/