Imperial College
London

The CMB: A case study in Bayesian probability

ICIC

The CMB: A Case Study

- Our underlying theories are statistical. How do we learn about cosmology from CMB observations?
  - predictions of power spectra (and higher moments): (quantum) noise
  - expand to include polarization
- Inferences in cosmology
- Measuring the spectrum, C_ℓ
  - temperature and polarization
- Measuring cosmological parameters
- Beyond the power spectrum
  - anisotropy [not small scales…]
    - sub-case study: topology
  - non-Gaussianity

} In the notes, but probably won't have time

ICIC

Data analysis as Radical Data Compression

- Radical Compression
  - Trillions of bits of data
  - Billions of measurements at 9 frequencies
  - 50 million pixel map of whole sky
  - 2 million harmonic modes measured
  - 2500 C_ℓ variances
    - 2000σ detection of CMB anisotropy power
  - Fit with just 6 parameters
    - Baryon density, CDM density, angular scale of sound horizon, reionization optical depth, slope and amplitude of primordial $P(k)$
    - $\Omega_b h^2$, $\Omega_c h^2$, θ_{MC}, τ, n_s, A_s
    - With no significant evidence for a 7th

ICIC

Parameters & C_ℓ

- What we really want:
  - P(theory | data)
    - theory = the parameters of LCDM
      - or perhaps even an indication of which overall theory is correct
    - data = our CMB data and any other information ("priors") we might consider.
  - Data compression
    - P(theory | raw TOI data) $\approx$ P(theory | noisy CMB map)
      $\approx$ P(theory | estimated $\hat{C}_\ell$)
    - Also need error bars (and/or full covariance matrix)
      - Even then, this is only approximate
        - effect of foreground removal on maps
        - $\hat{C}_\ell$ dist'n depends on more than just central value & covariance

ICIC

Bayesian methods: hierarchical models

- Timestream (d_t)
  - $\Rightarrow$ Map ($T_p \sim d_p$)
  - $\Rightarrow$ Spectrum ($C_l \sim d_l$)
  - $\Rightarrow$ cosmology

$$P(H \mid DI) = \frac{P(H \mid I) P(D \mid HI)}{P(D \mid I)}$$

$$\text{Posterior} \propto \text{Prior} \times \text{Likelihood}$$

- without loss of information? (~Sufficient Statistics)

- $P(\text{Cosmology}|d_t N_{tt'}) = P(\text{Cosmology} \mid \text{Map}_{pix}, N_{pp'})$
  $\approx P(\text{Cosmology} \mid D_l N_{ll'}, x_l)$

(Bond, AJ, Knox; WMAP)

- (assume that we can calculate $P(\text{Cosmology}|D_l N_{ll'}, x_l)$ even from non-Bayes estimators)

ICIC

CMB Data Analysis: mapmaking

- Model: data = signal + noise, as a function of time

$$d_t = A_{tp}T_p + n_t \qquad \langle n_t n_{t'}\rangle = N_{tt'} \quad \text{\textasciitilde stationary}$$

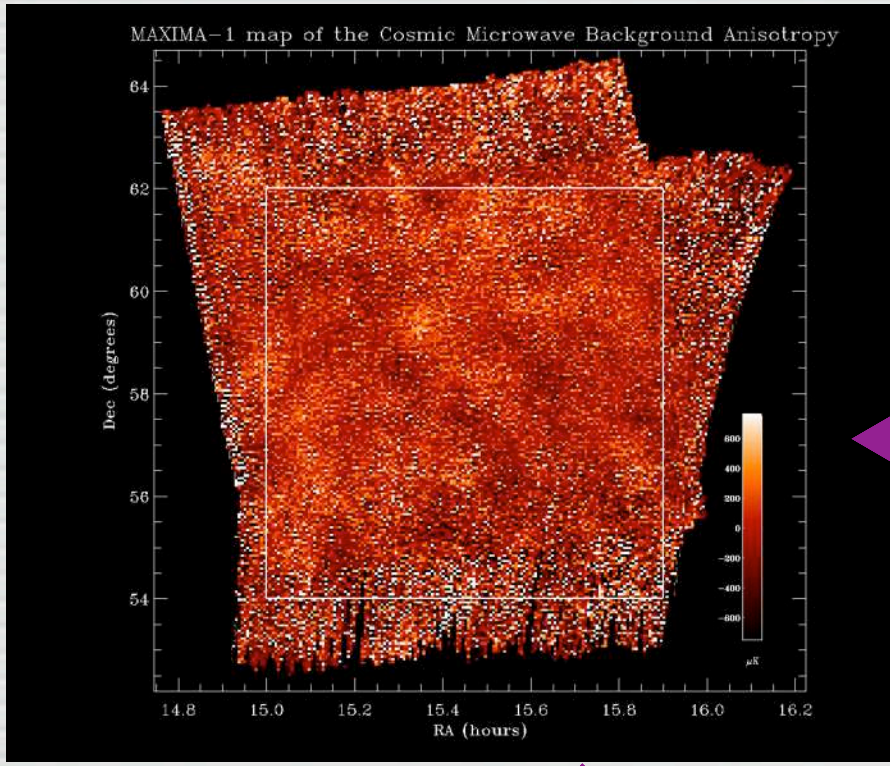
- Step 1: mapmaking (estimate T_p)
  - Gaussian noise ⇒ Gen'l least squares

$$\bar{T}_p = (A^T N^{-1} A)^{-1} A^T N^{-1} d \qquad \langle \delta T_p \delta T_{p'}\rangle = (A^T N^{-1} A)^{-1}$$

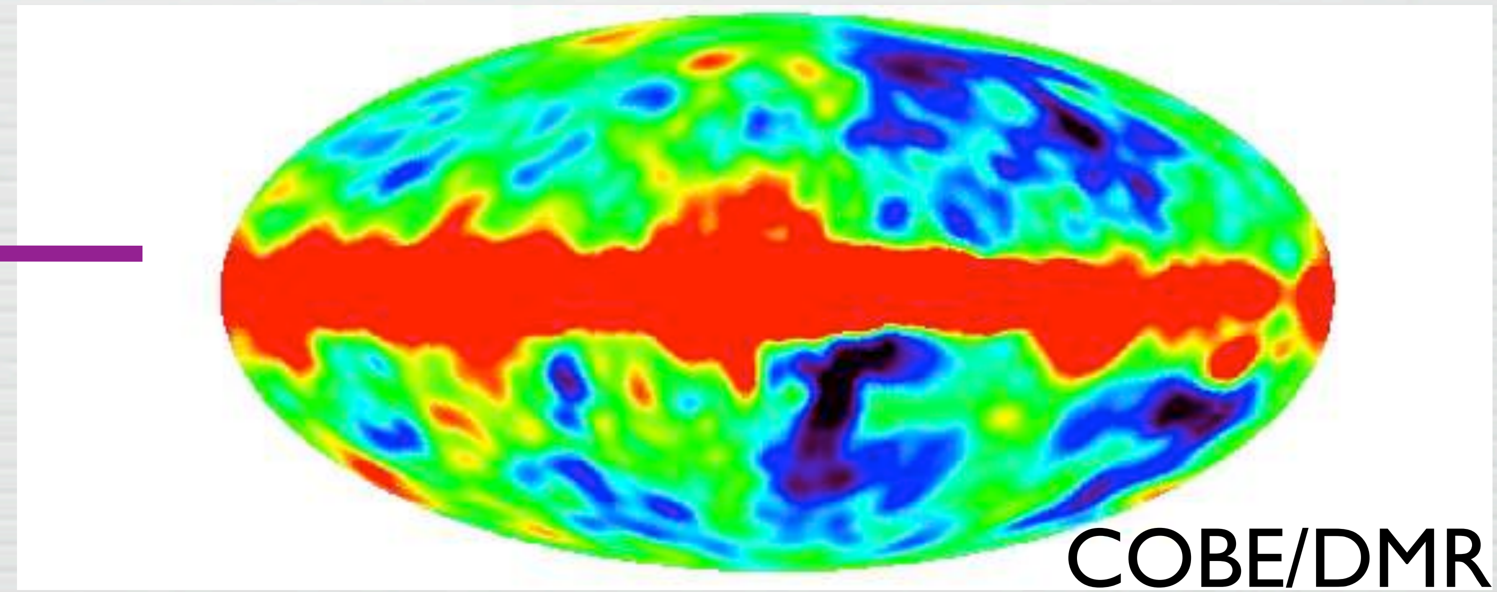
  - If we stop here, uniform prior gives a Gaussian posterior for the map with this mean and variance.
    - aside: Gaussian C_ℓ prior gives Wiener filter
  - But it is also a sufficient statistic
  - Algorithms:
    - Rely on simplicity/sparseness of A_{tp}
    - FFT methods to apply timestream (t) operations
    - Conjugate gradient least-squares soln (nb. doesn't give corr'n matrix)
    - Further simplifications for specific cases (1/f noise, observations in rings)

ICIC

Maps of the Cosmos

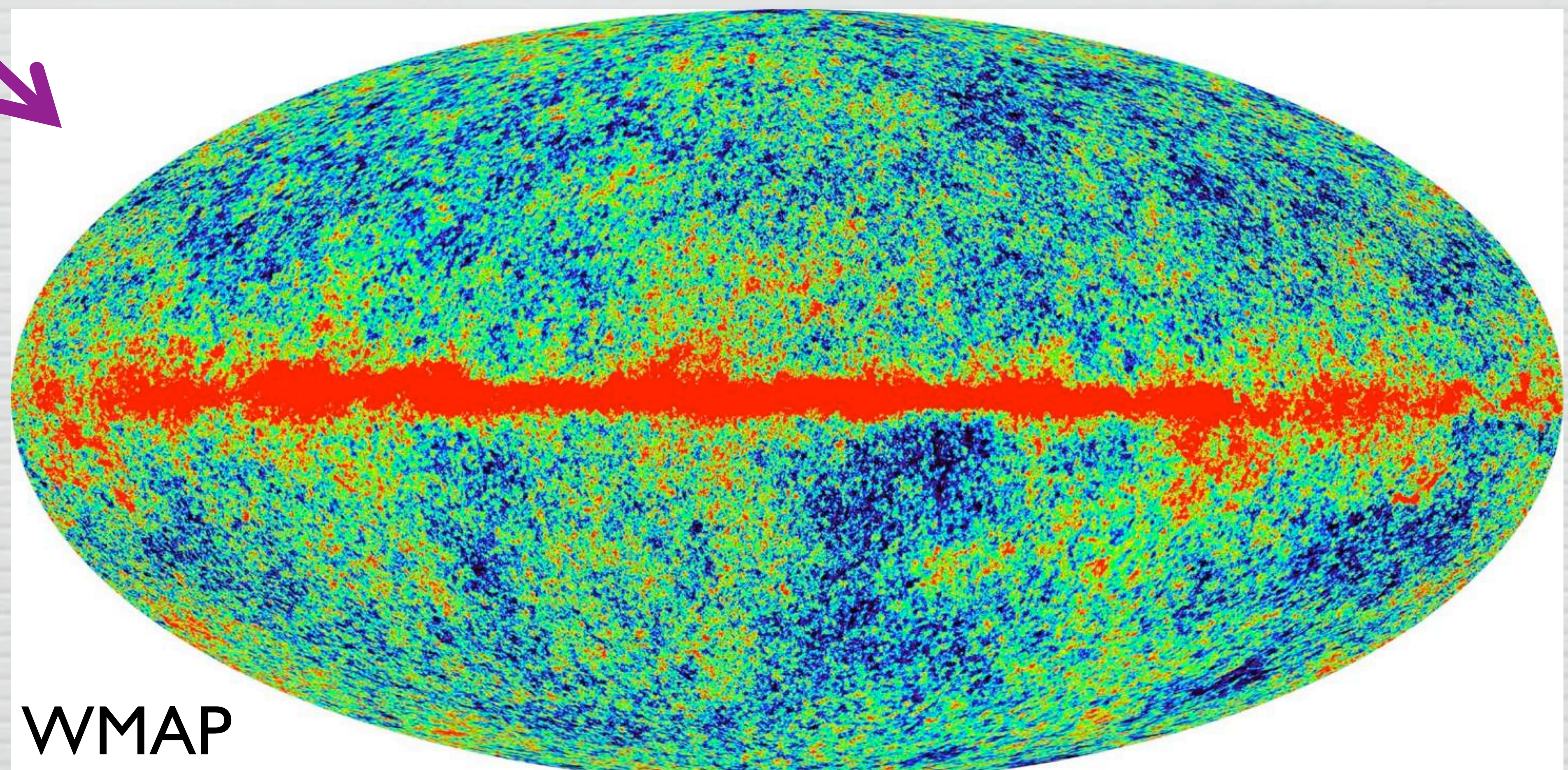


MAXIMA



COBE/DMR

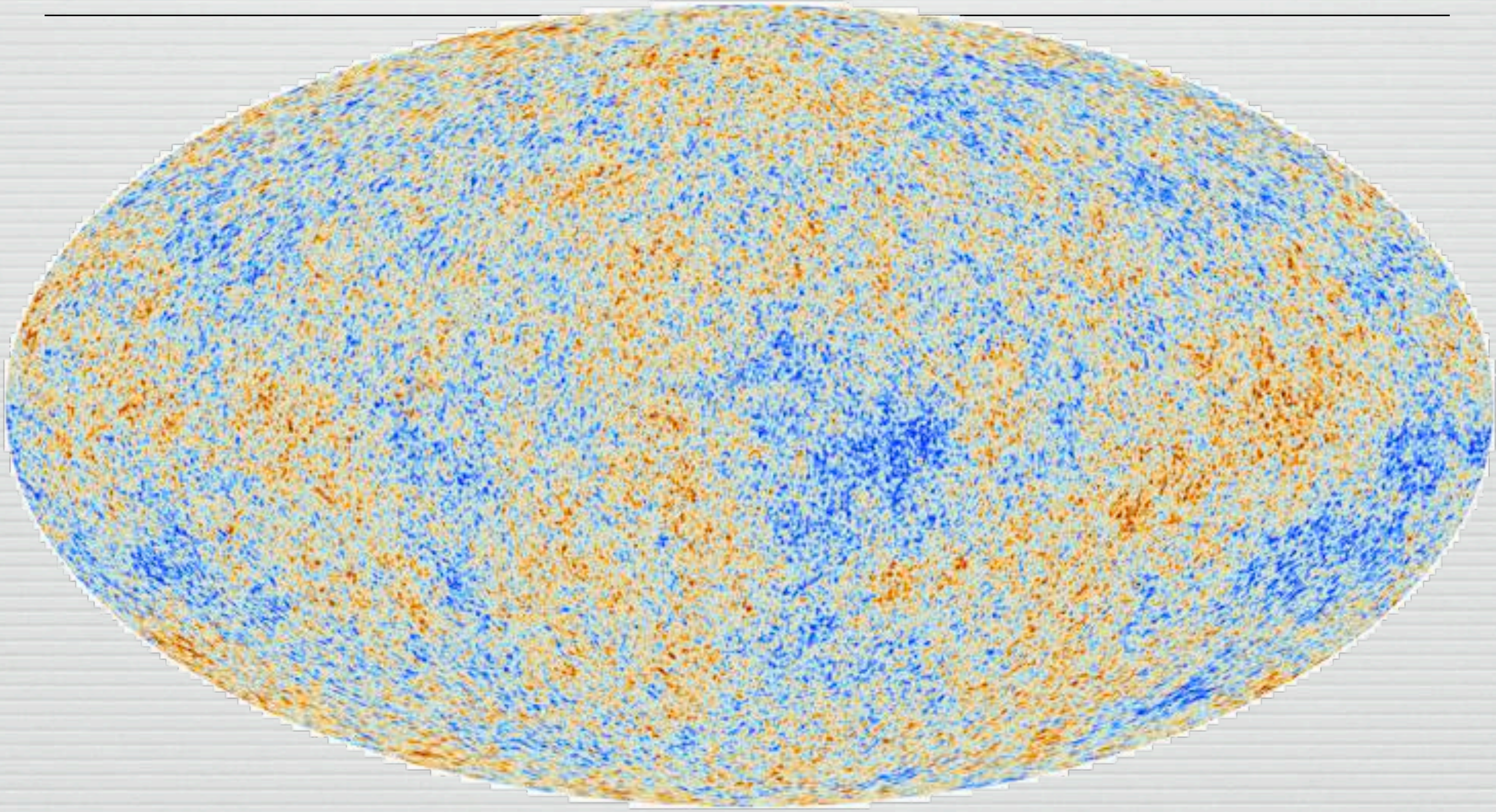
NB. pixelization on sphere non-trivial. CMB uses "HEALPix"



WMAP

ICIC

Planck



CMB Data Analysis: Spectrum estimation

- Step 2: Don't need to go back to the timeline to estimate the power spectrum, C_ℓ.
  - Model the sky as a correlated, statistically isotropic Gaussian random field

$$\frac{T(\hat{x})-\bar{T}}{\bar{T}} \equiv \frac{\Delta T}{T}(\hat{x}) = \sum_{\ell m} a_{\ell m} Y_{\ell m}(\hat{x}) \qquad \langle a^*_{\ell m} a_{\ell' m'} \rangle = \delta_{\ell\ell'}\delta_{mm'} C_\ell$$

spherical harmonic wavenumber ℓ

Parametric version of cov. mat. est'n: diag in ℓ basis

$$\langle T_p T_{p'} \rangle = S_{pp'} = \sum_\ell \frac{2\ell+1}{4\pi} C_\ell B_\ell^2 P_\ell(\hat{x}_p \cdot \hat{x}_{p'})$$

$$P(\bar{T}|C_\ell) = \frac{1}{|2\pi(S+N)|} \exp -\frac{1}{2}\bar{T}^T(S+N)^{-1}\bar{T}$$

  - complicated and expensive function of C_ℓ
    - Many practical issues in calculating this explicitly.
    - At low ℓ, use sampling (usu. Gibbs), Newton-Raphson, Copula
    - At high ℓ, approximate by a function of estimated (ML) C_ℓ and errors & some other information X_ℓ

ICIC

Expected errors

- Estimating the error (variance$^{1/2}$) on a variance (C_ℓ)
- $\langle \delta C_\ell \, \delta C_\ell \rangle = \langle a_{\ell m}\, a_{\ell m}\, a_{\ell m}\, a_{\ell m} \rangle - \langle a_{\ell m}\, a_{\ell m} \rangle \langle a_{\ell m}\, a_{\ell m} \rangle$
  - Wick's theorem: $\langle a^4 \rangle = 3\langle a^2 \rangle^2$
  - CMB case: Knox 95, Hobson & Magueijo 96
    - need to account for $(2\ell+1)f_{\rm sky}$ measurements of each ℓ

$$(\delta C_\ell)^2 \cong \frac{2}{(2\ell+1) f_{\rm sky}} (C_\ell + N_\ell)^2 \qquad N_\ell \approx w^{-1} = (\theta_p \sigma_p)^{-2}$$

# of modes

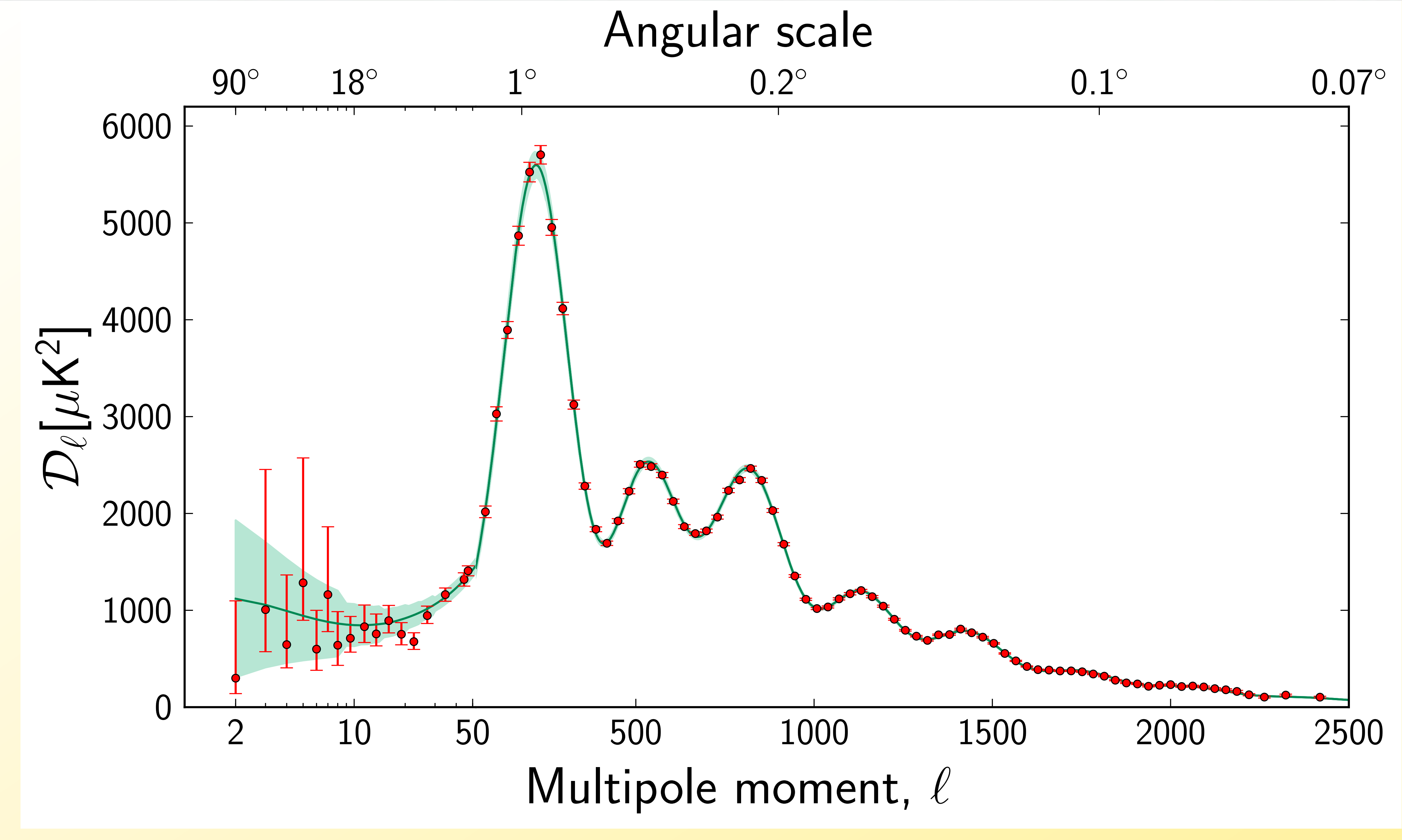
Sample (cosmic) Variance

Noise variance

- Bandpowers: bin in ℓ (weighted for specific C_ℓ shape) to reduce errors and decrease covariance

ICIC

Planck errors



Error band: cosmic variance estimate

error bars: cosmic + noise variance

ICIC

A toy model

- Consider all-sky observations with uniform white noise

$$d_p = T_p + n_p \qquad \langle T_p T_{p'} \rangle = S_{pp'} = \sum_\ell \frac{2\ell+1}{4\pi} C_\ell B_\ell^2 P_\ell(\hat{x}_p \cdot \hat{x}_{p'})$$

- Pixel-space likelihood $\langle n_p n_{p'} \rangle = N_{pp'} = \sigma^2 \delta_{pp'}$

$$P(d_p|C_\ell) = \frac{1}{|2\pi(S+N)|^{1/2}} \exp -\frac{1}{2} d^T (S+N)^{-1} d$$

- Work in harmonic space $d_{\ell m} \simeq \int d^2\hat{x}_p \, d(\hat{x}_p) Y_{\ell m}(\hat{x}_p)$
- White noise equiv to const. noise spectrum, $N_\ell = N \propto \sigma^2$

$$\langle n_{\ell m} n_{\ell' m'} \rangle = \delta_{\ell\ell'} \delta_{mm'} N$$

- Likelihood separates

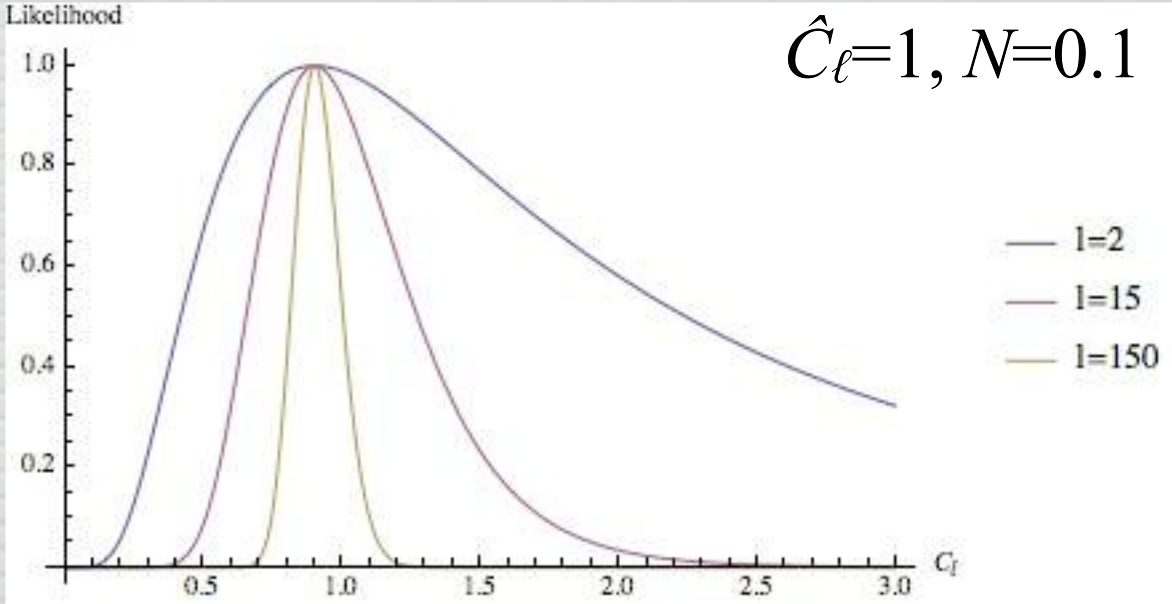
$$P(d_{\ell m}|C_\ell) = \prod_\ell \frac{1}{|2\pi(C_\ell + N)|^{\ell+1/2}} \exp\left(-\frac{2\ell+1}{2} \frac{\hat{C}_\ell}{C_\ell + N} \right)$$

- with pseudo spectrum $\hat{C}_\ell \equiv \frac{1}{2\ell+1} \sum_m |d_{\ell m}|^2$

Toy model

$$P(d_{\ell m}|C_\ell) = \prod_\ell \frac{1}{|2\pi(C_\ell + N)|^{\ell+1/2}} \exp\left(-\frac{2\ell+1}{2}\frac{\hat{C}_\ell}{C_\ell + N}\right) \qquad \hat{C}_\ell \equiv \frac{1}{2\ell+1}\sum_m |d_{\ell m}|^2$$

- Likelihood (as a function of C_ℓ) maximized at $C_\ell = \hat{C}_\ell - N$
- with curvature $\left.\frac{d^2 \ln P}{dC_\ell^2}\right|_{C_\ell=\hat{C}_\ell - N} = -\left(\frac{2\hat{C}_\ell}{2\ell+1}\right)^{-1}$
  - cf. Gaussian $\frac{d^2 \ln P}{dx^2} = -\left(\sigma^2\right)^{-1}$
    and Fisher information $F \equiv -\left\langle \frac{d^2 \ln P}{dx^2} \right\rangle$



- Skew positive likelihood
- more Gaussian as $\ell \to \infty$

ICIC

Bayesian methods: MADCAP/MADspec

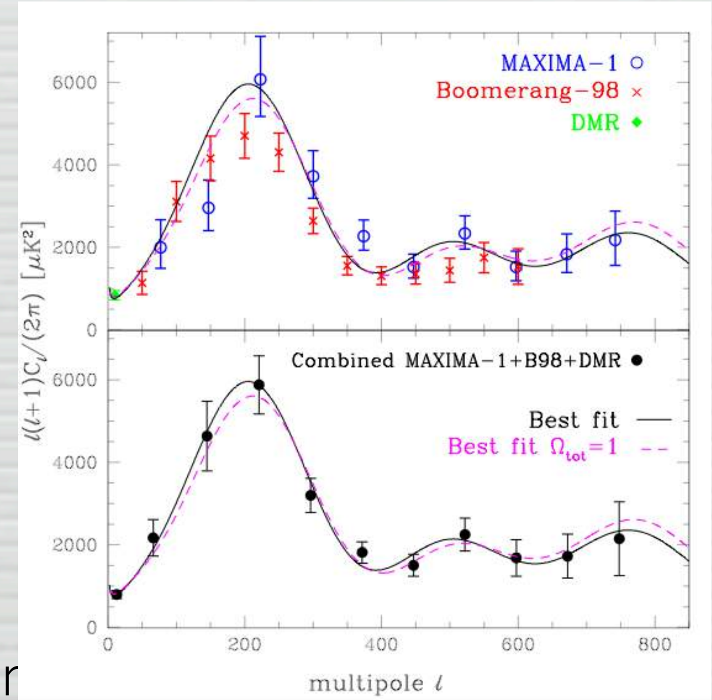
- (quasi-)Newton-Raphson iteration to Likelihood maximum
- Algorithm driven by matrix manipulation (iterated quadratic):

$$\delta C_\ell = \tfrac{1}{2} F^{-1}_{ll'} \mathrm{Tr}\left[\left(dd^T - C\right)\left(C^{-1}\frac{\partial C}{\partial C_\ell}C^{-1}\right)\right]$$

$$F_{ll'} = \tfrac{1}{2}\mathrm{Tr}\left[C^{-1}\frac{\partial C}{\partial C_\ell}C^{-1}\frac{\partial C}{\partial C_\ell}\right] \quad \text{Fisher matrix}$$

$$\mathrm{C} = \mathrm{S} + \mathrm{N}$$

- Fisher = approx. Likelihood curvature
- full polarization: signal matrix $S^{xx'}_{pp'}$
- Arbitrary (precomputed) noise spectrum
- Arbitrary linear filters
  - Stompor et al; Jaffe et al; Slosar et al
- $O(N^3)$ operations naïvely (matrix manipulations), speedup to ~$O(N^2)$ for spectrum estimates (potentially large prefactor)
  - Fully parallelized (MPI, SCALAPACK)
  - do calculations in the natural basis
  - no explicit need for full $N_{pp'}$ matrix in pixel basis (just noise spectrum or autocorrelation)
- e.g., MAXIMA, BOOMERANG



Borrill, Cantalupo, Stor

Frequentist Monte Carlo methods

- MASTER: quadratic pseudo-C_l estimate (Hivon et al)

$$d_{\ell m} = \sum_p d_p w_p \Omega_p Y_{\ell m}(\hat{x}_p)$$

$$\hat{C}_\ell = \frac{1}{2\ell+1} \sum_m |d_{\ell m}|^2 \qquad \text{pseudo-}C_l$$

$$\hat{C}_\ell \approx \langle \hat{C}_\ell \rangle = \sum_{\ell'} C_\ell M_{\ell\ell'} F_\ell B_\ell^2 + N_\ell$$

(Will discuss Bayesian sampling for C_ℓ later on)

where
N is noise bias
M is mode coupling depending on sky coverage
F is experimental filter

- SPICE: transform of correlation function estimate (Szapudi et al)
- Issues: filters, weights, noise estimation/iteration, input maps — optimal or naïve?

Hybrid Methods: FASTER

- Key insight: MASTER covariance formalism allows calculation of diagonal part of pseudo-a_{lm} covariance — use for likelihood maximization
  - (nb. this has maximum entropy and so is conservative!)
  - *Diagonal* likelihood:
  - $$P(d_{\ell m} \mid C_\ell I) = \frac{1}{\left[2\pi\left\langle \hat{C}_\ell + N_\ell \right\rangle\right]^{1/2}} \exp\left[-\tfrac{1}{2}\frac{\left|d_{\ell m}\right|^2}{\left\langle \hat{C}_\ell + N_\ell \right\rangle}\right]$$
  - MC evaluation of means;
  - Newton-Raphson iteration towards maximum
  - Easy calculation of Likelihood shape parameters

B98, CBI; Contaldi et al

(related suggestions from Delabrouille et al)

ICIC

WMAP (etc.): Cross-correlations

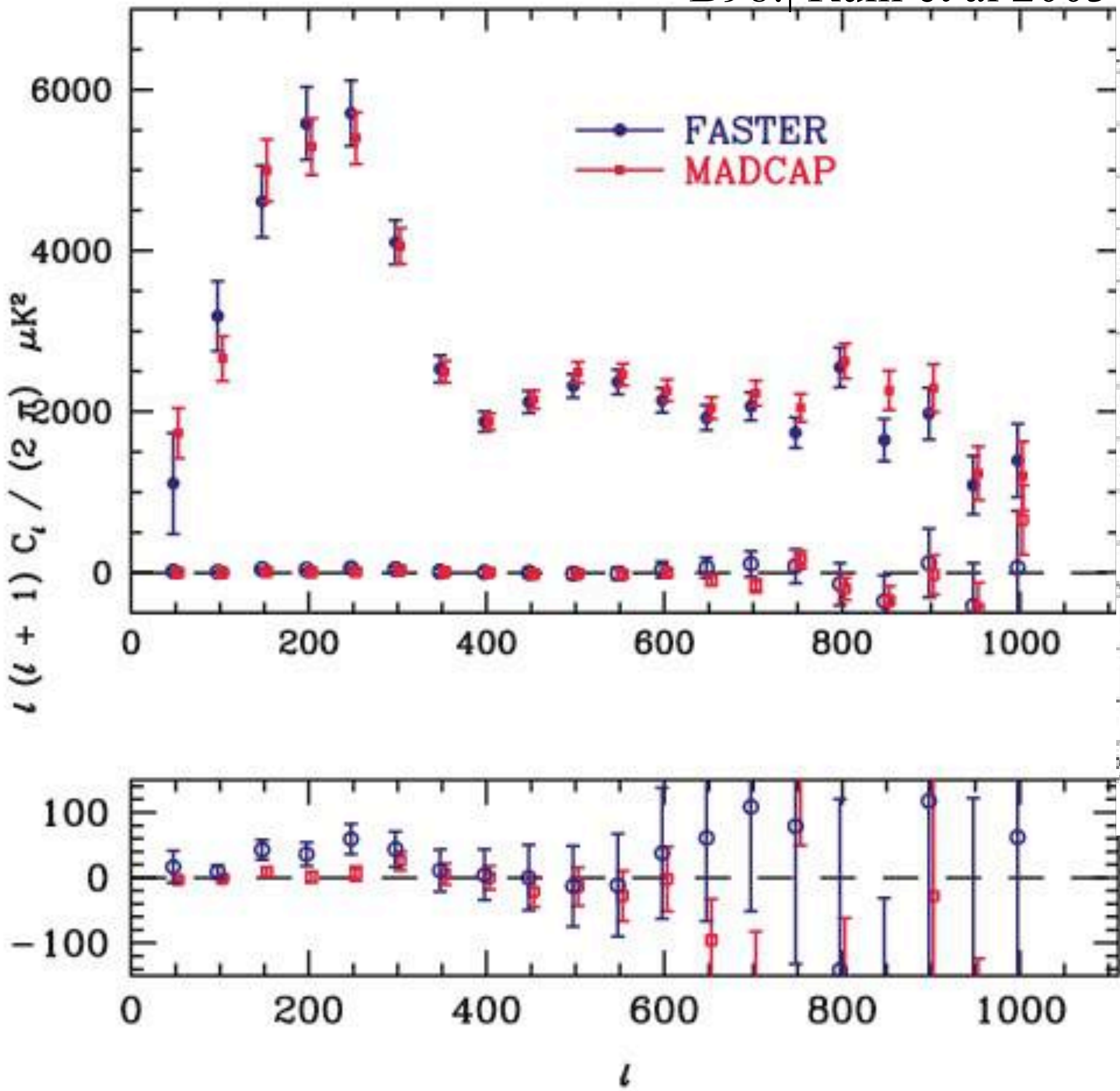
- Take advantage of uncorrelated noise between different detectors
- $$\left\langle d_p^1 d_{p'}^2 \right\rangle = \left\langle \left(s_p^1 + n_p^1\right)\left(s_{p'}^2 + n_{p'}^2\right) \right\rangle = S_{pp'}^{12} + N_{pp'}^{12} = S_{pp'}$$

  Monte Carlo method — without need for noise bias removal

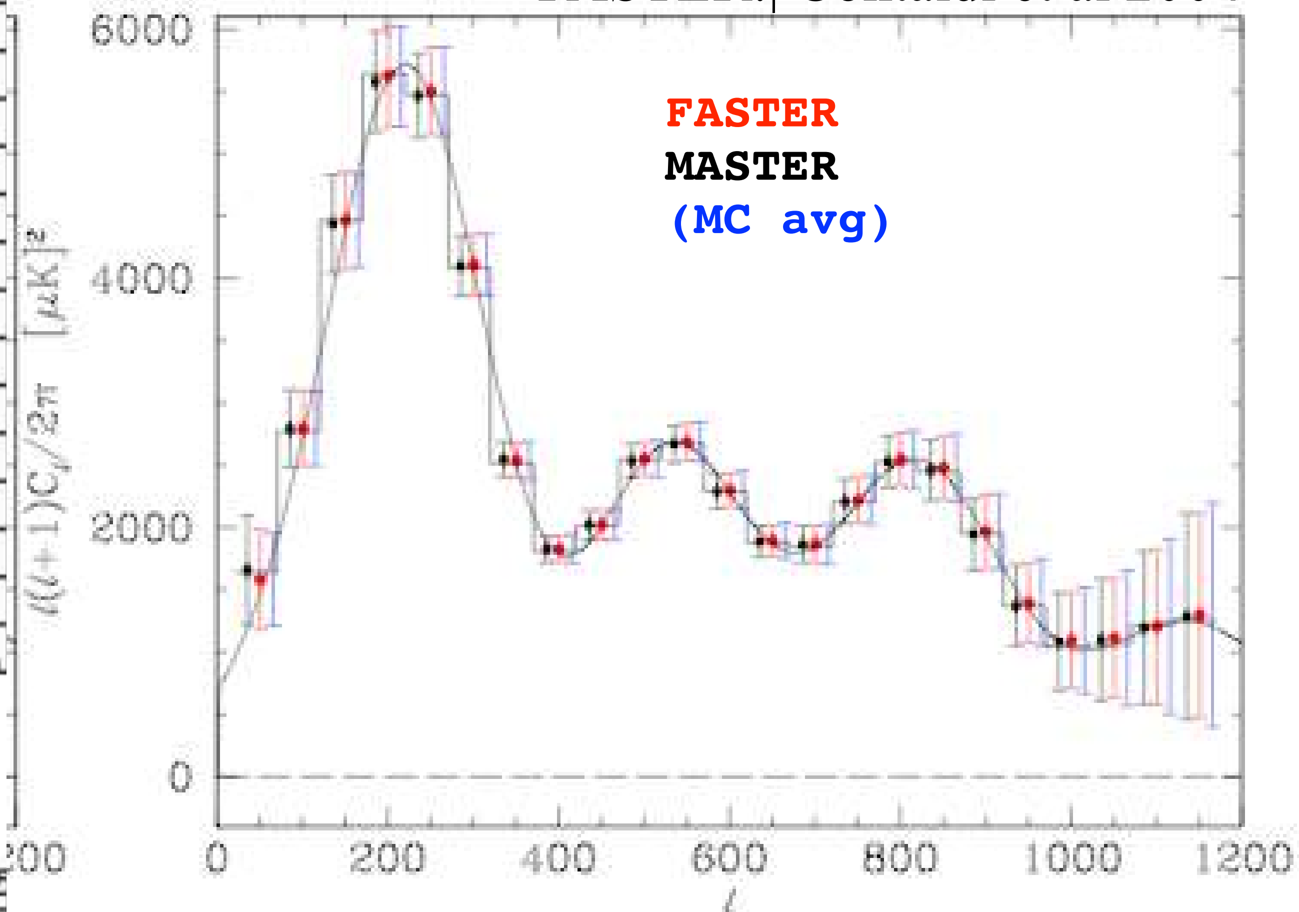
-

Comparisons

B98:| Ruhl et al 2003



FASTER:| Contaldi et al 2004

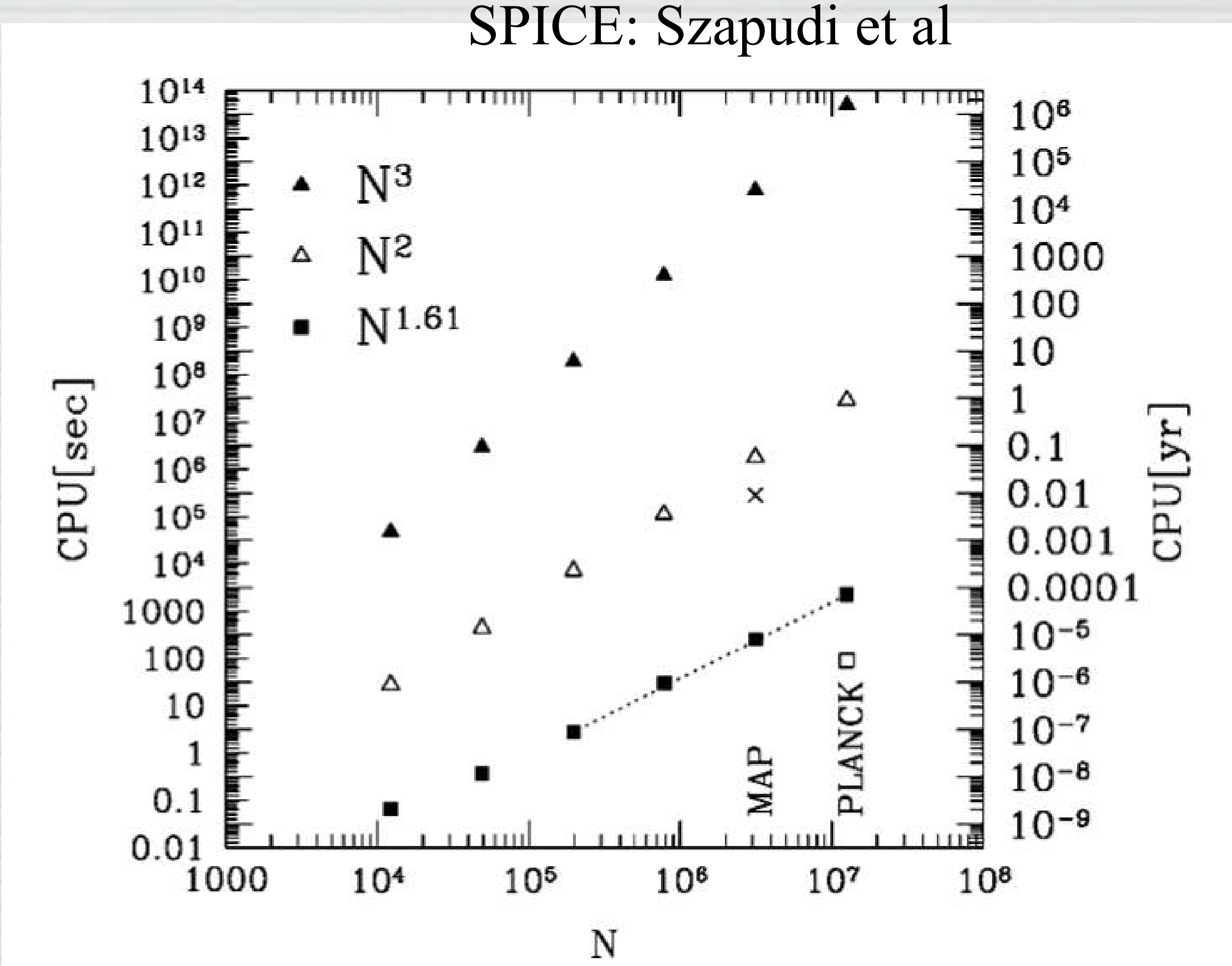


Timing and efficiency

- time
  - optimal/bayes: N_p^3
  - monte carlo: $N^{1.5}$
  - prefactors: N_{MC}, N_{bin}, …
- Space
  - TOI: 50 GB/yr @200Hz
  - maps: 384 Mb @ N_{side}=2048
  - noise matrix: $N^2/2$ entries

    ~9 petabytes @ N_{side}=2048

SPICE: Szapudi et al



resource management will become an issue even for cheapest methods

Bayesian/Frequentist Correspondence

- Why do both methods seem to work?
- frequentist mean ~ likelihood maximum
  frequentist variance ~ likelihood curvature
- Correspondence is *exact* for
  - linear gaussian models (mapmaking)
  - variance estimation with no correlations and "iid" noise — simple version of C_l problem
    - e.g., all sky, uniform noise
    - likelihood only function of d_{lm}^2
    - breaks down in realistic case of correlations, finite sky, varying noise
  - "asymptotic limit"
    - ~ high l iff noise correlations not "too strong"

- But we still want to bootstrap from point estimates to the full likelihood function

ICIC

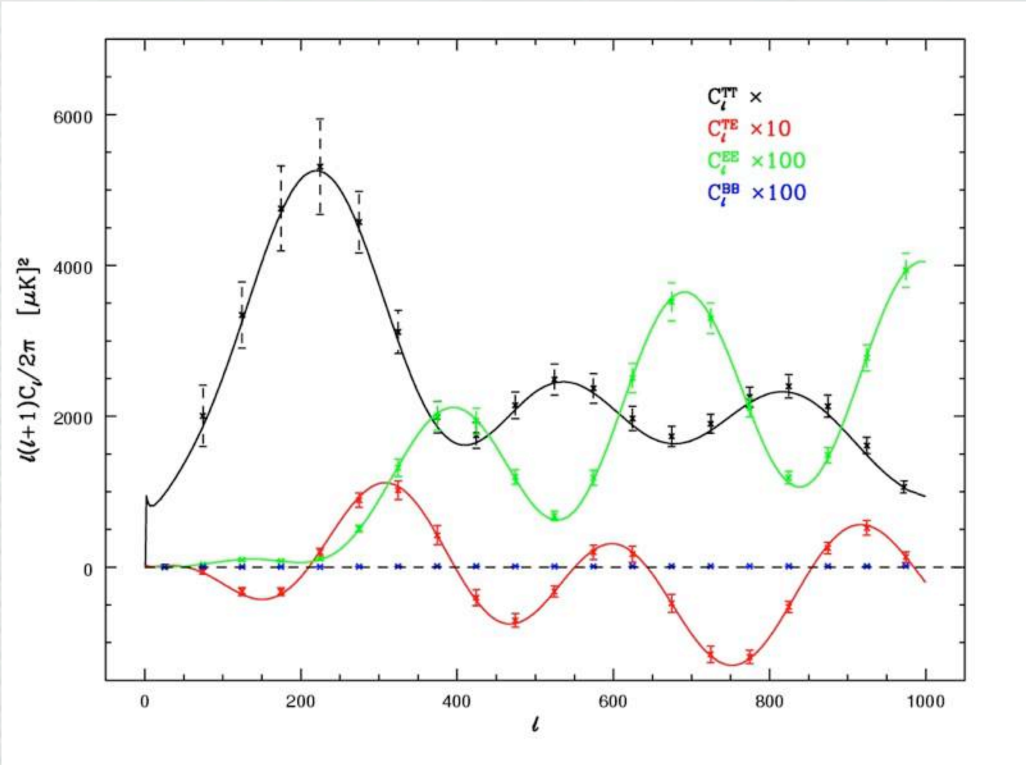
Polarization

- Formally the same problem:
  - $d_p \Rightarrow (i,q,u)_p = d_{i,p} = d_q$
  - $\langle d_q d_{q'} \rangle = N_{qq'} + S_{qq'}$
    - low S/N, large systematics
    - complicated correlations:
      - $N_{qq'}$: pixel differences
      - $S_{qq'} = S^{ij}_{qq'}$: linearly dependent on all of $C_l^{XX'}$ (X=T,E,B)
  - e.g., Seljak, Zaldarriaga; Kamionkowski, Kosowsky, Stebbins; &c.
- E/B leakage (= T/E/B correlation)
  - in principle, don't need extra separation step if full correlations/distributions is known
  - in practice, E/B characteristics impose specific correlation structure — easier to "separate"
    - Wiener filter for map from C_l.

ICIC

Polarization

- Formally the same problem:
  - $d_p \Rightarrow (i,q,u)_p = d_{i,p} = d_q$
  - $\langle d_q d_{q'} \rangle = N_{qq'} + S_{qq'}$
    - low S/N, large systematics
    - complicated correlations:
      - $N_{qq'}$: pixel differences
      - $S_{qq'} = S^{ij}_{qq'}$: linearly dependent on all of $C_l^{XX'}$ (X=T,E,B)
  - e.g., Seljak, Zaldarriaga; Kamionkowski, Kosowsky, Stebbins; &c.



- E/B leakage (= T/E/B correlation)
  - in principle, don't need extra separation step if full correlations/distributions is known
  - in practice, E/B characteristics impose specific correlation structure — easier to "separate"
    - Wiener filter for map from C_l.

ICIC

From C_ℓ to cosmology

- Step 3: Calculate & characterize posterior prob over some space of cosmological models and imposed priors
- For simplest [?] theories, C_ℓ is a deterministic function of the cosmological parameters $\theta=\{H_0, n_s, \Omega_m, \Omega_{DE}, \ldots\}$
  - $P(\theta|DI) = \int dC_\ell\, P(\theta|I)\, P(C_\ell|\theta I)\, P(C_\ell\,|DI)$
    $= P(\theta|I)\, P(C_\ell[\theta]\,|\,DI)$
    $= P(\theta|I)\, P(C_\ell[\theta]\,|\,\hat{C}_\ell, \sigma_\ell, \text{shape}, I)$

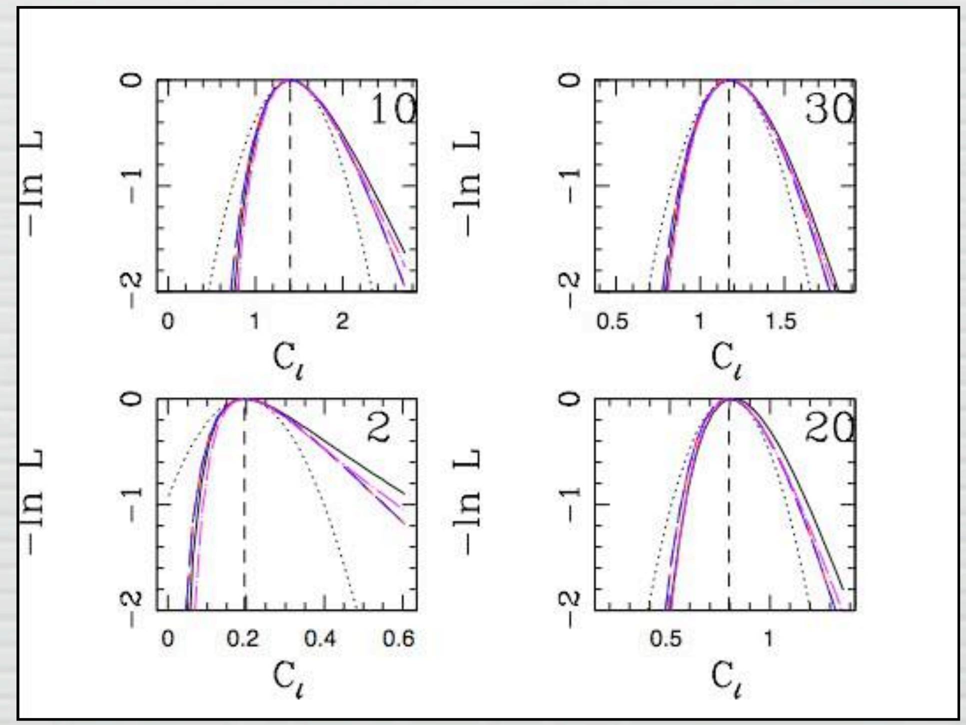
    ML est.   Variance
- So est'd C_ℓ is [approximately] a *sufficient statistic*
  - Only approximate, so not really a separate step
  - $P(\theta|d_t) = P(\theta|T_p) \approx P(\theta|C_\ell)$
  - can explore the likelihood — or finally assign meaningful priors on θ and calculate the posterior
    - MCMC, etc.

ICIC

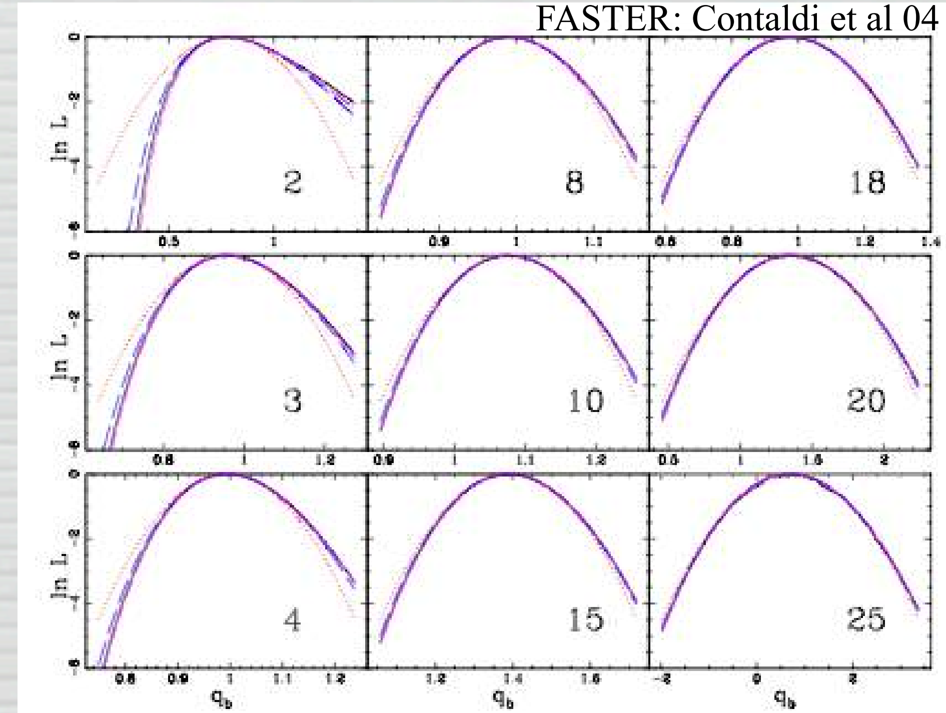
The shape of the likelihood function

$$P(\bar{T}|C_\ell) = \frac{1}{|2\pi(S+N)|} \exp -\frac{1}{2}\bar{T}^T(S+N)^{-1}\bar{T}$$

- Complicated function of C_ℓ [through S(C_ℓ)]
- *not* a Gaussian in C_ℓ
  - big effect at low ℓ
- ~Offset lognormal (BJK 00)
  - Gaussian in $\ln(C_\ell + x_\ell)$
  - Other approximations better at moderate ℓ
    - e.g., Hamimeche & Lewis
      - include polarization
      - treat T, Q, U on same footing



FASTER: Contaldi et al 04



Sampling from the posterior

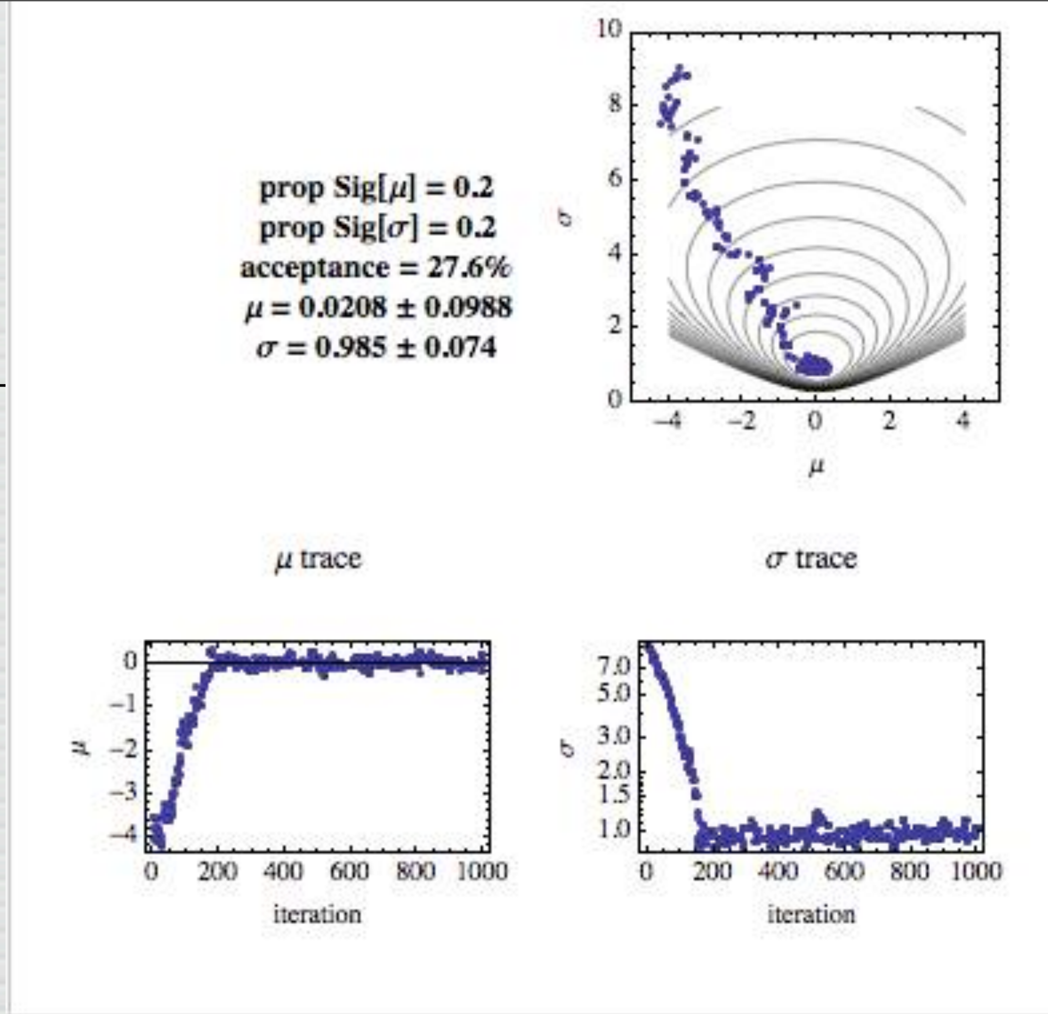
- Infeasible to directly explore P(θ|data) for many parameters θ
  - e.g., even the 6-parameter base LCDM model would require ~$100^6=10^{12}$ evaluations for 100 grid points in each direction…
- Instead, *generate samples* θ_i from the distribution.
  - Easy to evaluate moments (means, variances)
    - $\langle\theta\rangle = \frac{1}{N}\sum_i \theta_i$ or, more generally $\langle f(\theta)\rangle = \frac{1}{N}\sum_i f(\theta_i)$

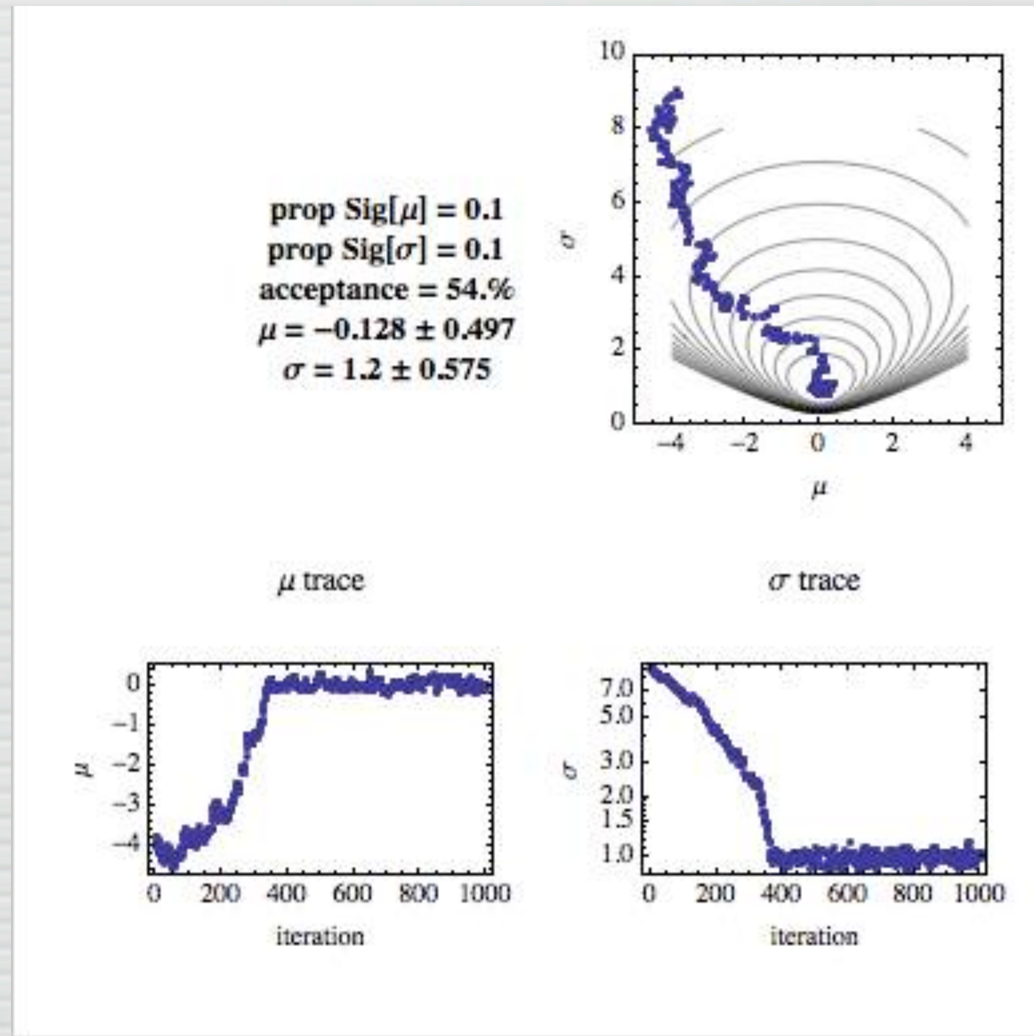
ICIC

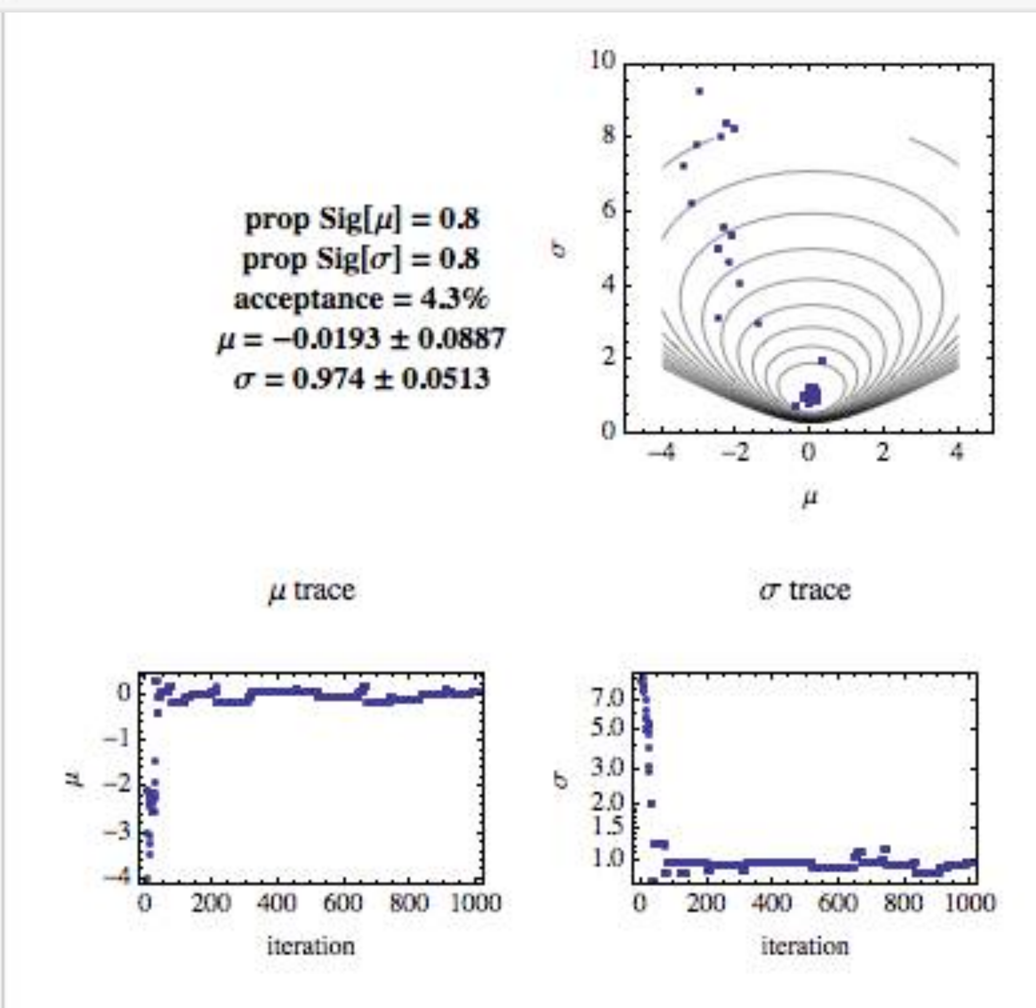
MCMC

- Generate samples from posterior P(x)
- Most methods require being able to generate samples from some simpler distribution
- e.g., Markov Chain Monte Carlo
  - Start with proposal distribution $Q(x^*|x)$: probability of proposing point x^* if starting at point x
  - often $Q(x|y) = Q(|x\text{-}y|)$ (Metropolis)
    - Metropolis Algorithm:
      - given point $x^{(i)}$, generate x^* from $Q(x^*|x^{(i)})$
      - accept x^* as $x^{(i+1)}$ with probability min[1,P(x^*)/P($x^{(i)}$)];
      - otherwise $x^{(i+1)} = x^{(i)}$
      - repeat…

ICIC

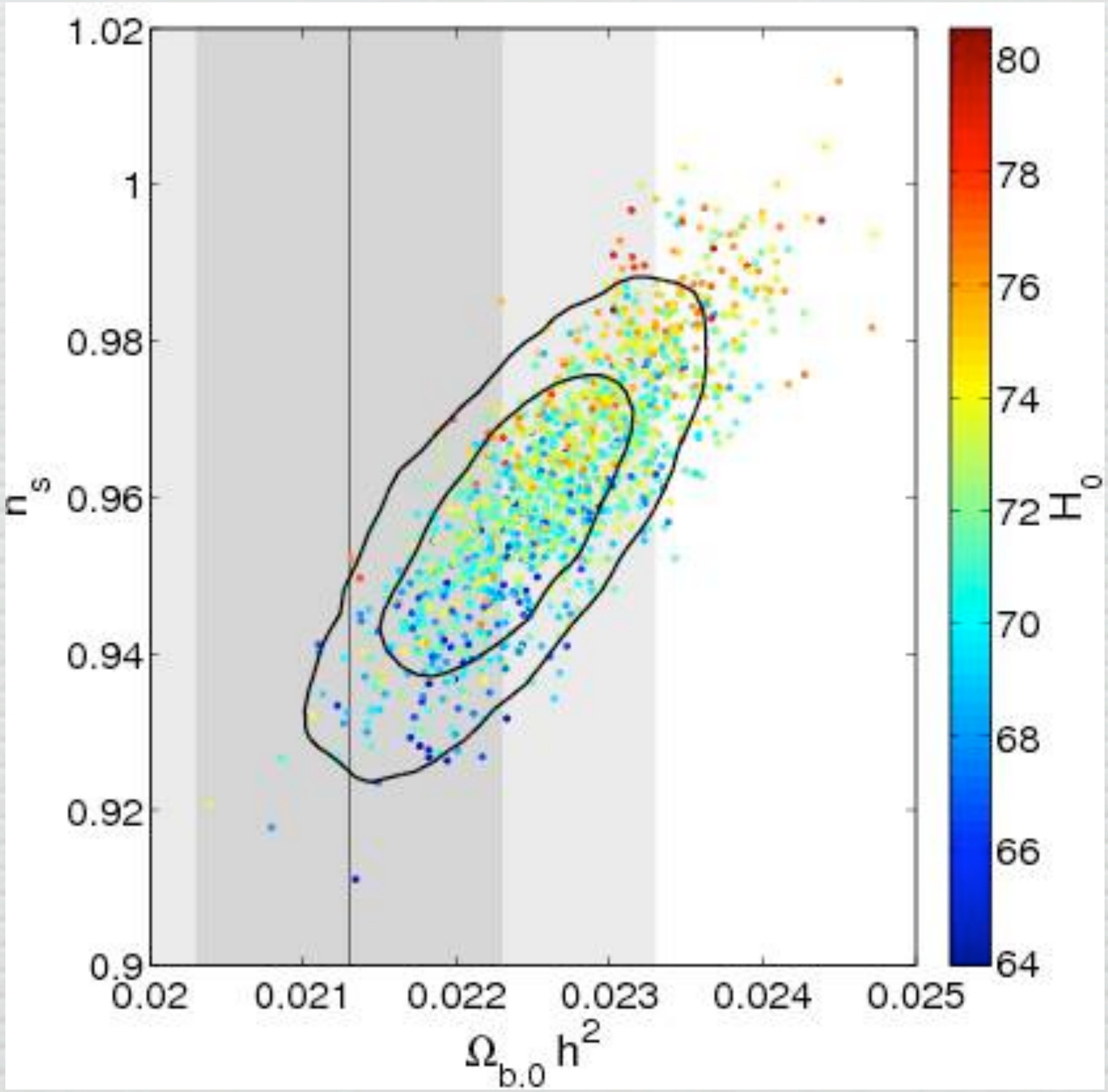

prop Sig[μ] = 0.2
prop Sig[σ] = 0.2
acceptance = 27.6%
μ = 0.0208 ± 0.0988
σ = 0.985 ± 0.074
μ trace
σ trace
iteration
iteration


prop Sig[μ] = 0.1
prop Sig[σ] = 0.1
acceptance = 54.%
μ = −0.128 ± 0.497
σ = 1.2 ± 0.575
μ trace
σ trace
iteration
iteration


prop Sig[μ] = 0.8
prop Sig[σ] = 0.8
acceptance = 4.3%
μ = −0.0193 ± 0.0887
σ = 0.974 ± 0.0513
μ trace
σ trace
iteration
iteration

Monte Carlo methods for the CMB

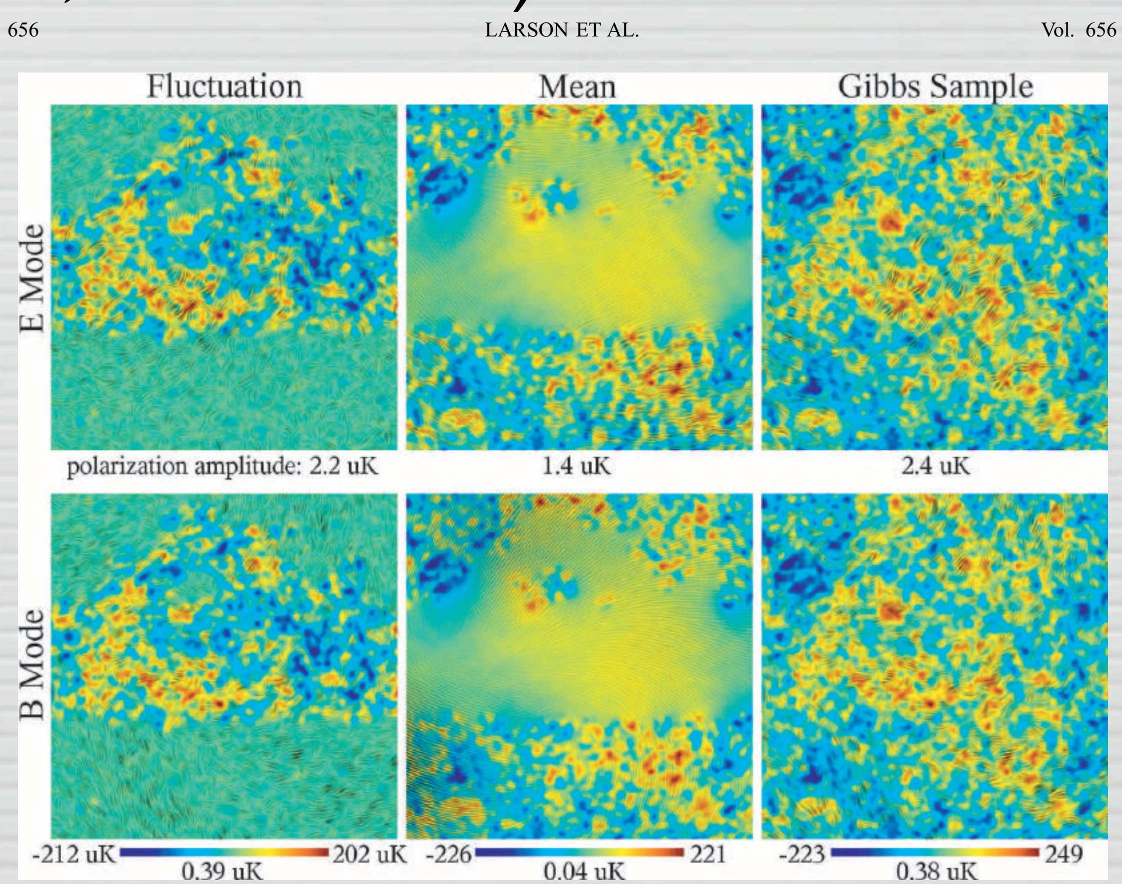
- Markov Chain Monte Carlo: A. Lewis' CosmoMC
  - coupled with fast deterministic calculation of power spectrum as fn of cosmological parameters
  - e.g. CMBFAST, CAMB, CLASS
  - Other techniques
    - e.g., Skilling's "nested sampling" which also allows fast calc'n of model likelihoods ("evidence")



ICIC

Aside: Gibbs Sampling

- Combine parametric models of foregrounds with power spectrum estimation
  - *Jewell et al; Wandelt et al; Eriksen et al; Larson et al;*
  - draw [full-sky] map realization given C_ℓ and foreground parameter (Wiener filter)
  - draw foreground realization given C_ℓ and map
  - draw C_ℓ realization given map (Wishart, Gamma dists)
- Output is sample maps and samples of C_ℓ
  - not always useful for subsequent *parameter estimation*
  - construct approx. likelihood by averaging over samples
    - Blackwell-Rao estimator



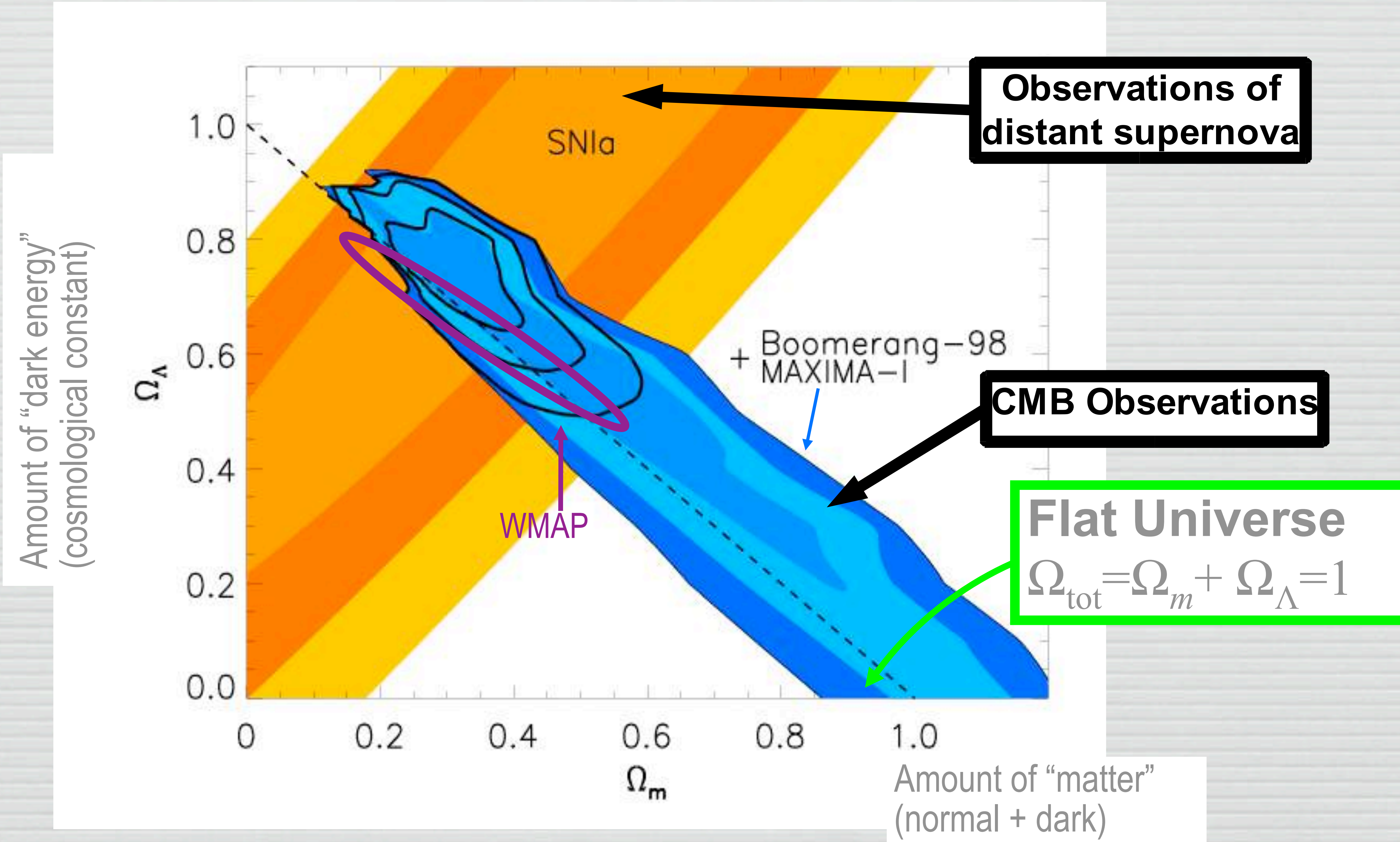
ICIC

The Planck likelihood

- High ℓ
  - Start with pseudo-C_ℓ of each detector, with conservative masks
    - for cosmology, consider
      100x100, 143x143, 217x217, 143x217
  - Foregrounds:
    - Use 353 GHz as a dust template
    - Explicit power spectral templates for unresolved point sources, SZ, CIB
  - Instrument:
    - relative calibration between 100, 143, 217
    - beam errors
  - Use Gaussian approximation assuming a fiducial models gives the signal covariances (Hamimeche & Lewis)
- low ℓ
  - Temperature: Planck 30-353 GHz
  - polarization: WMAP
    - needed to fix optical depth τ

ICIC

Measuring the geometry of the Universe



ICIC

Measuring the geometry of the Universe

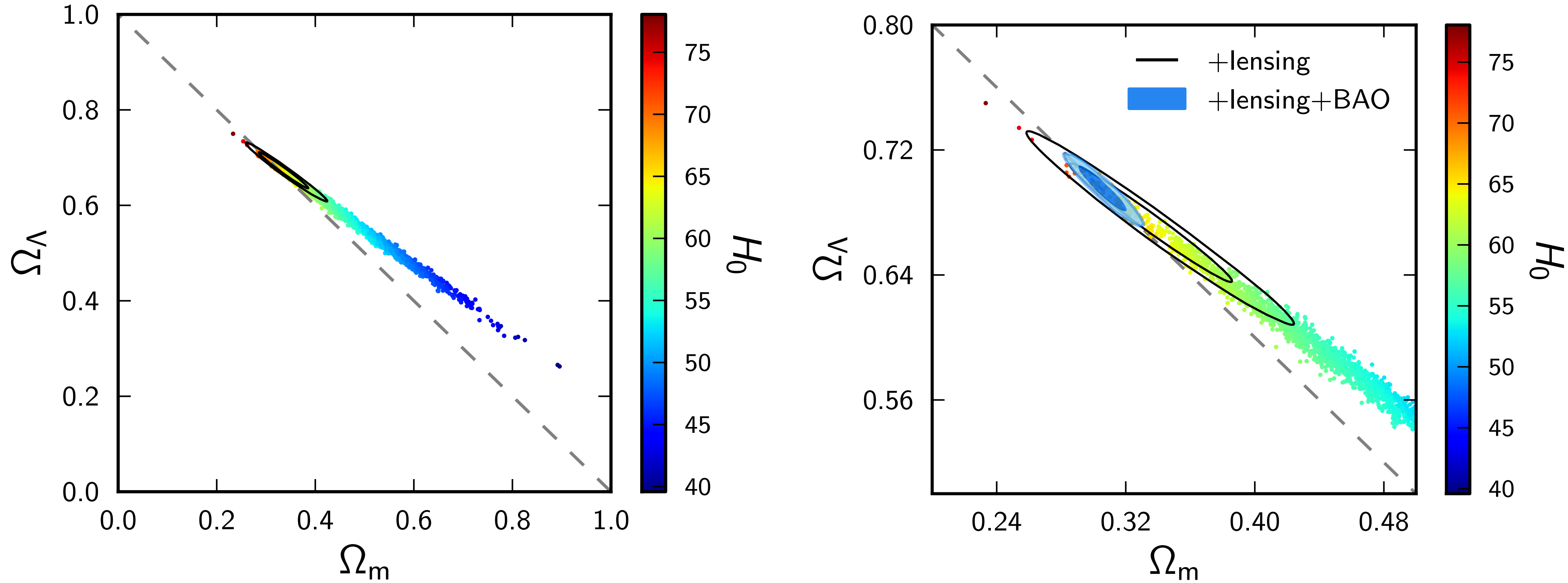


Fig. 25. The *Planck*+WP+highL data combination (samples; colour-coded by the value of H_0) partially breaks the geometric degeneracy between Ω_m and Ω_Λ due to the effect of lensing in the temperature power spectrum. These limits are significantly improved by the inclusion of the *Planck* lensing reconstruction (black contours). Combining also with BAO (right; solid blue contours) tightly constrains the geometry to be nearly flat.

ICIC

Planck Params

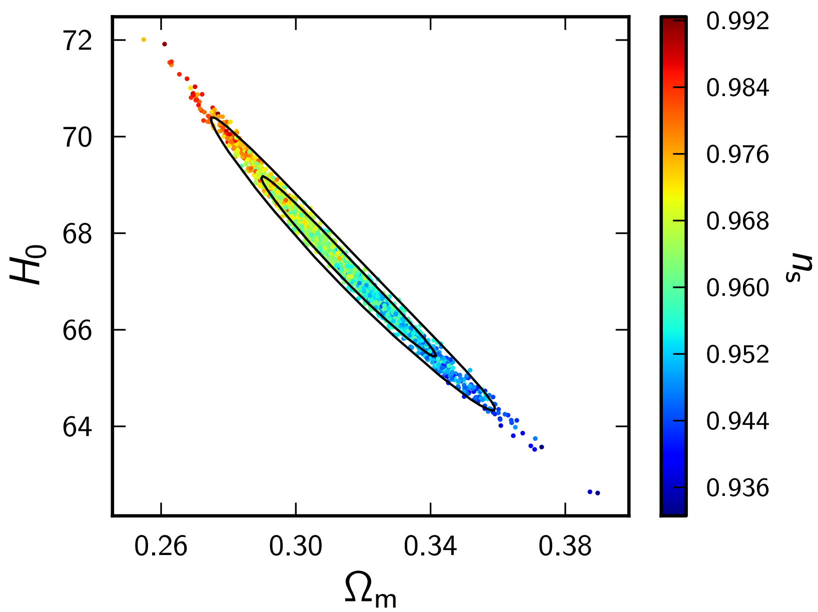


Fig. 3. Constraints in the Ω_m–H_0 plane. Points show samples from the *Planck*-only posterior, coloured by the corresponding value of the spectral index n_s. The contours (68% and 95%) show the improved constraint from *Planck*+lensing+WP. The degeneracy direction is significantly shortened by including WP, but the well-constrained direction of constant $\Omega_m h^3$ (set by the acoustic scale), is determined almost equally accurately from *Planck* alone.

Planck Collaboration: Cosmological parameters

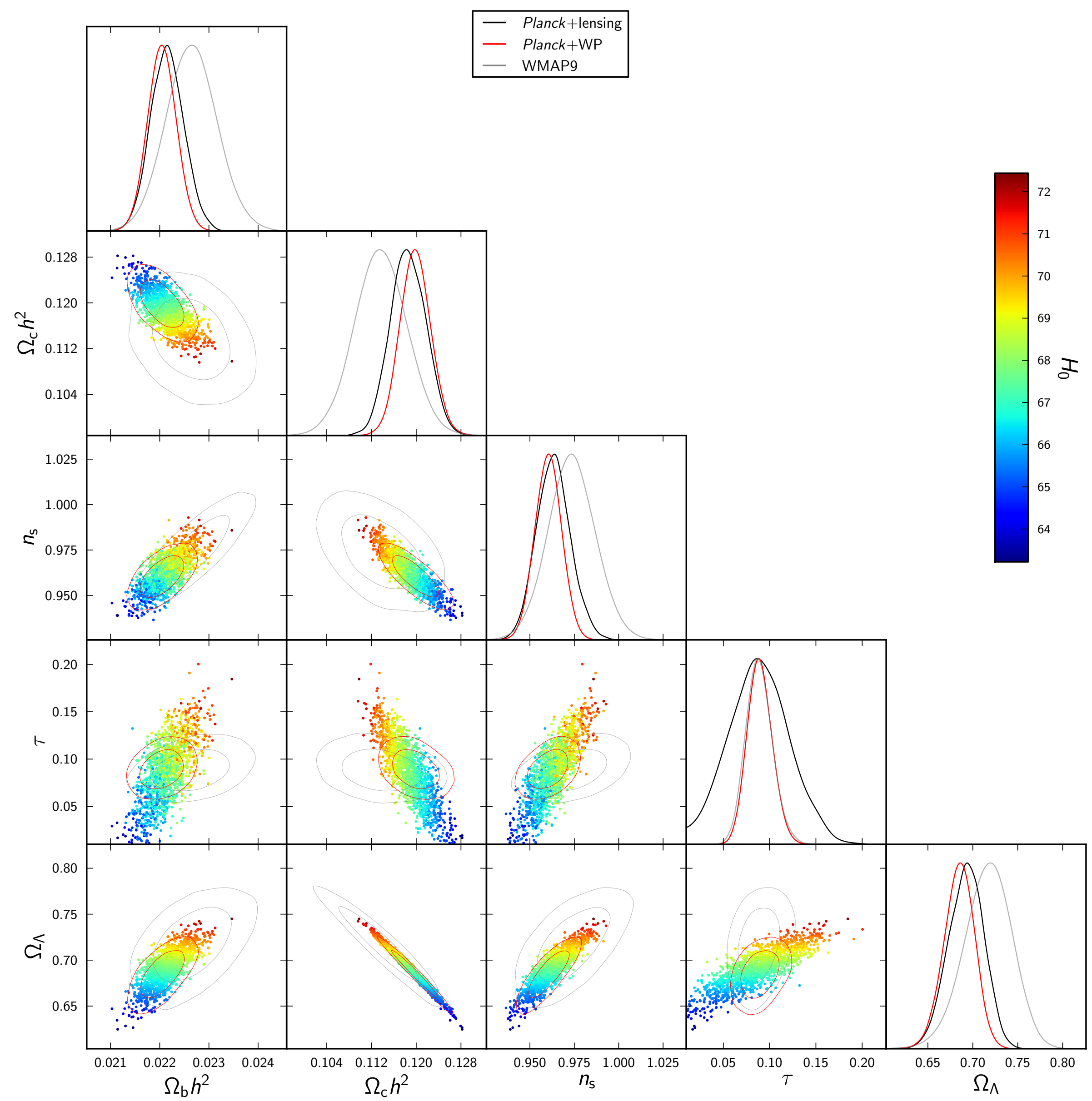


Fig. 2. Comparison of the base ΛCDM model parameters for *Planck*+lensing only (colour-coded samples), and the 68% and 95% constraint contours adding *WMAP* low-ℓ polarization (WP; red contours), compared to *WMAP*-9 (Bennett et al. 2012; grey contours).

ICIC

Hierarchical Models

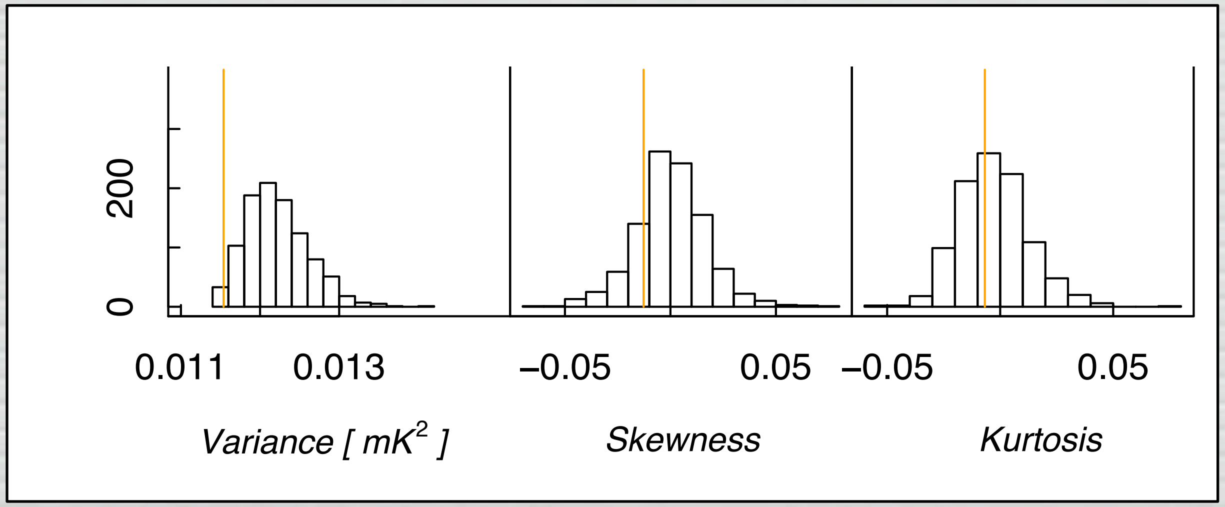
- So we have a hierarchical model
  - ask progressively more complicated questions of the data, with (approximately) no dependence on the details of previous results
  - Timelines ⇒ maps ⇒ spectra ⇒ parameters
- Each is a "nuisance parameter" for the next step w/ an uncontroversial prior *defining* that step
  - e.g., $\langle T_p T_{p'} \rangle = S_{pp'}(C_\ell)$ $P(C_\ell|\theta) = \delta[C_\ell - C_\ell(\theta)]$
- But in the realistic case there may be other nuisance parameters for which the priors are relevant:
  - timeline systematics, foregrounds, &c.

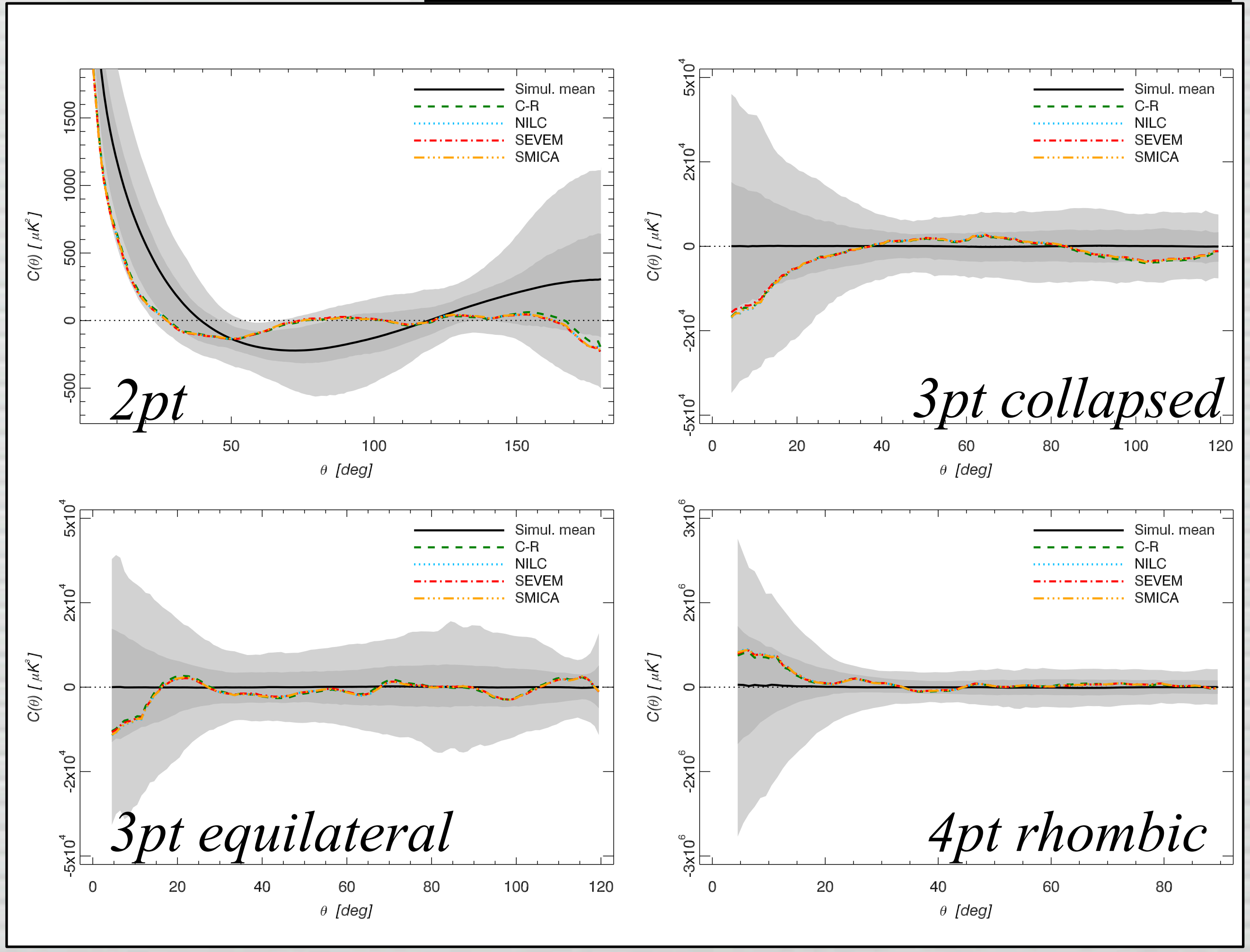
ICIC

Testing assumptions

- We have been calculating the posterior
  $P(\Omega_b h^2, \Omega_c h^2, \theta_{MC}, \tau, n_s, A_s \mid \text{Planck}, I)$
- Background information "I"
  - (testable) assumptions:
    - LCDM in general
    - gaussianity, isotropy
    - by some measures, it obeys these assumptions very well

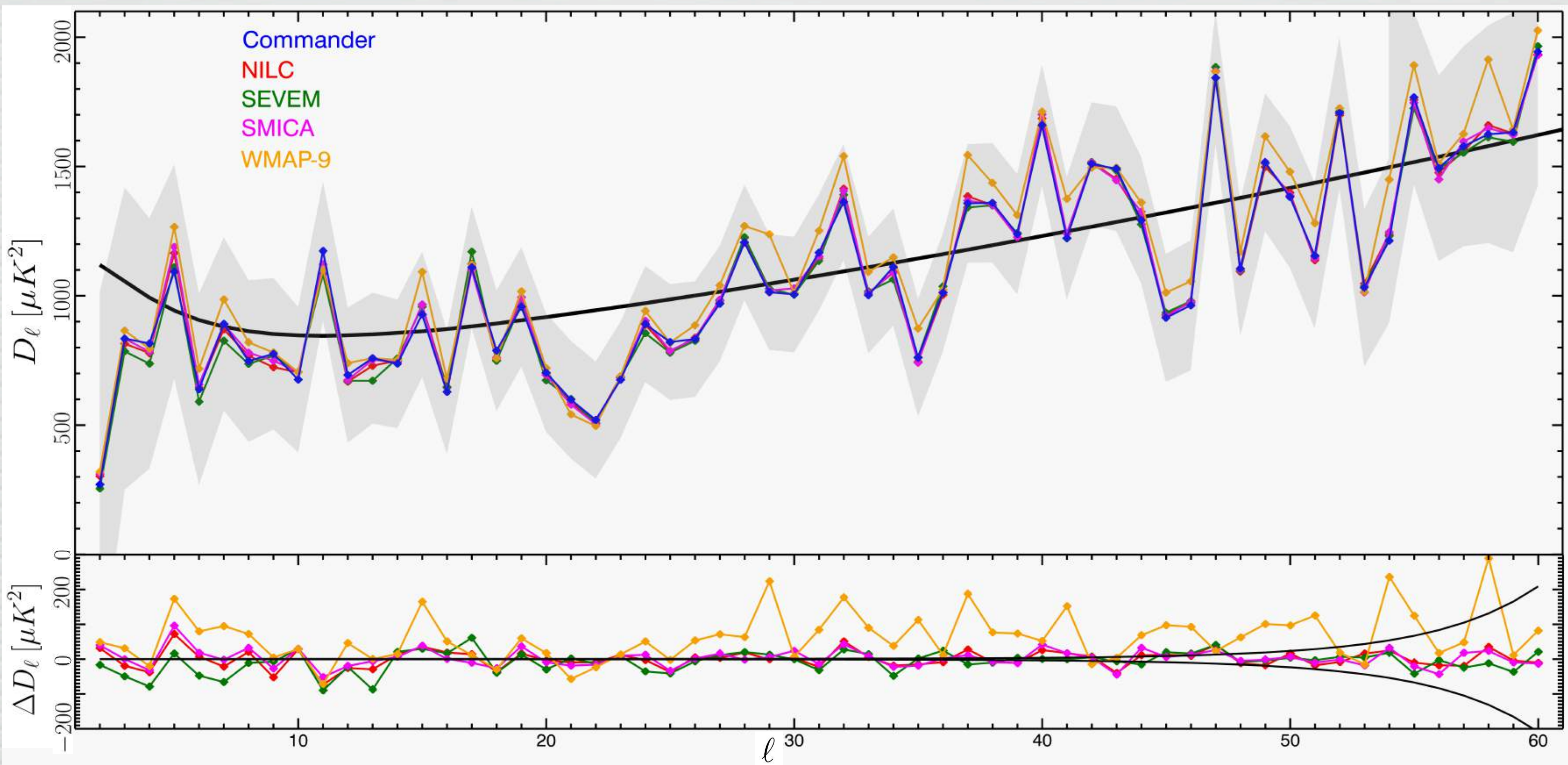
Very difficult to test these assumptions absent a specific alternative, in a Bayesian way





ICIC

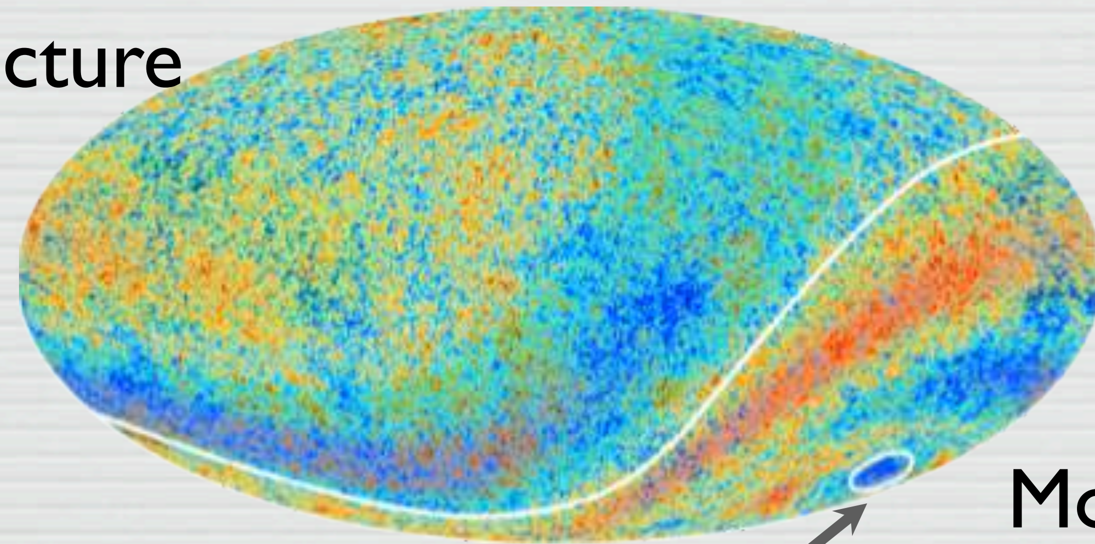
Low power on large scales



ICIC

Anomalies?

Less structure

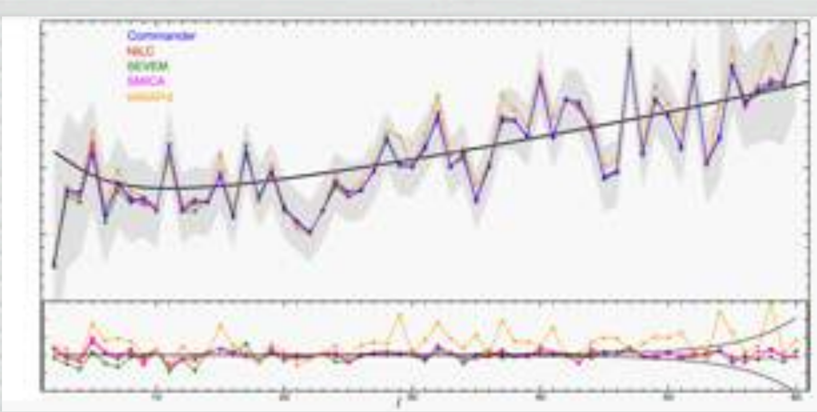


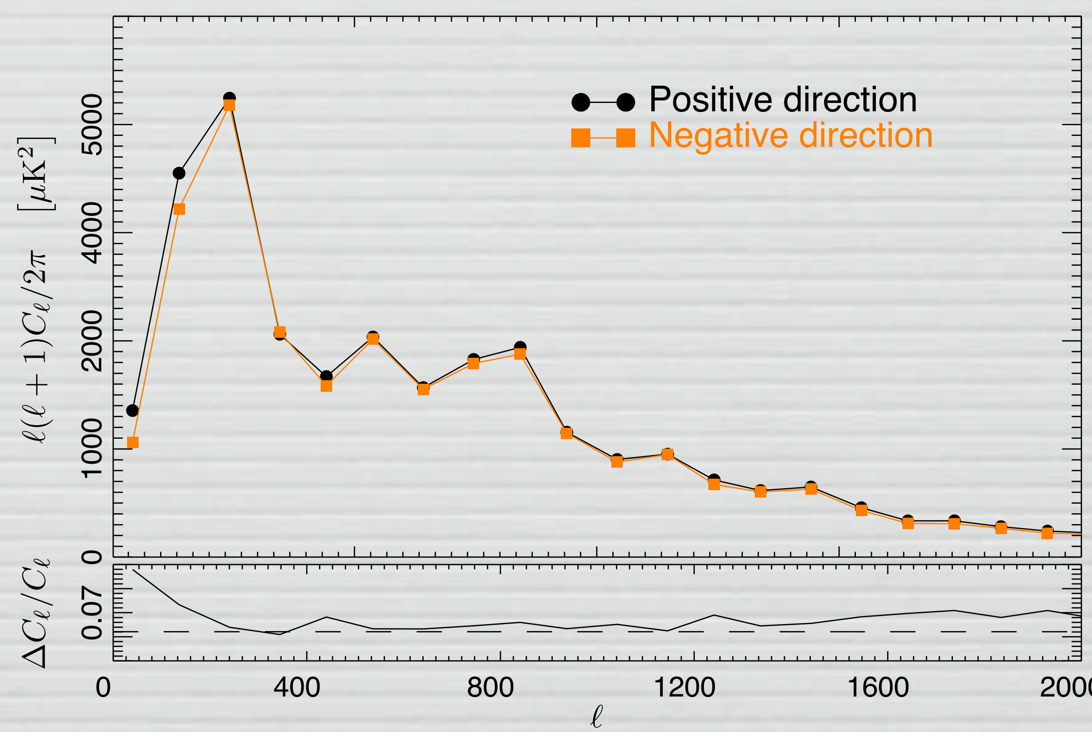
nb. there is also a *known* asymmetry from CMB dipole aberration.

More structure

"cold spot"

Small (but statistically significant) difference between the power in the hemispheres





Overall low amplitude at large scales

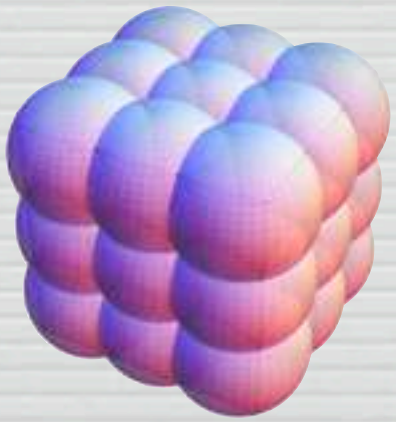
Large-scale anisotropy

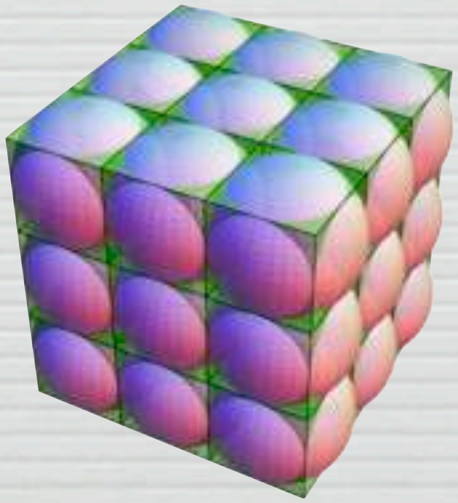
- Hemispherical differences: how can we arrange anisotropy on the scale of the horizon?
  - initial conditions: anisotropic inflation?
  - the large-scale structure of spacetime
    - change the geometry: Bianchi
      - homogeneous + anisotropic spacetimes
    - change the topology

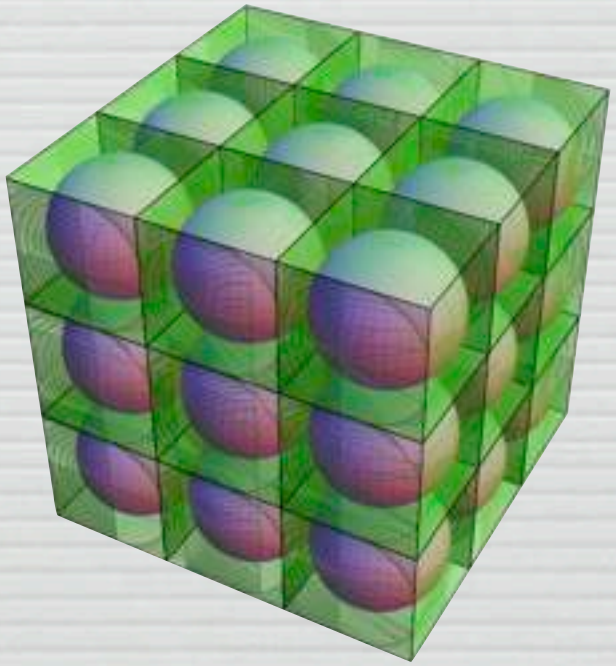
ICIC

The shape of the Universe

- General relativity determines the curvature of the Universe, but not its topology (holes and handles)
- Most theories of quantum gravity (and quantum cosmology) predict topological change on small scales and at early times.
- Does this have cosmological implications?
  - E.G., small universe $\Rightarrow$ fewer large-scale modes available $\Rightarrow$ low power on large scales?

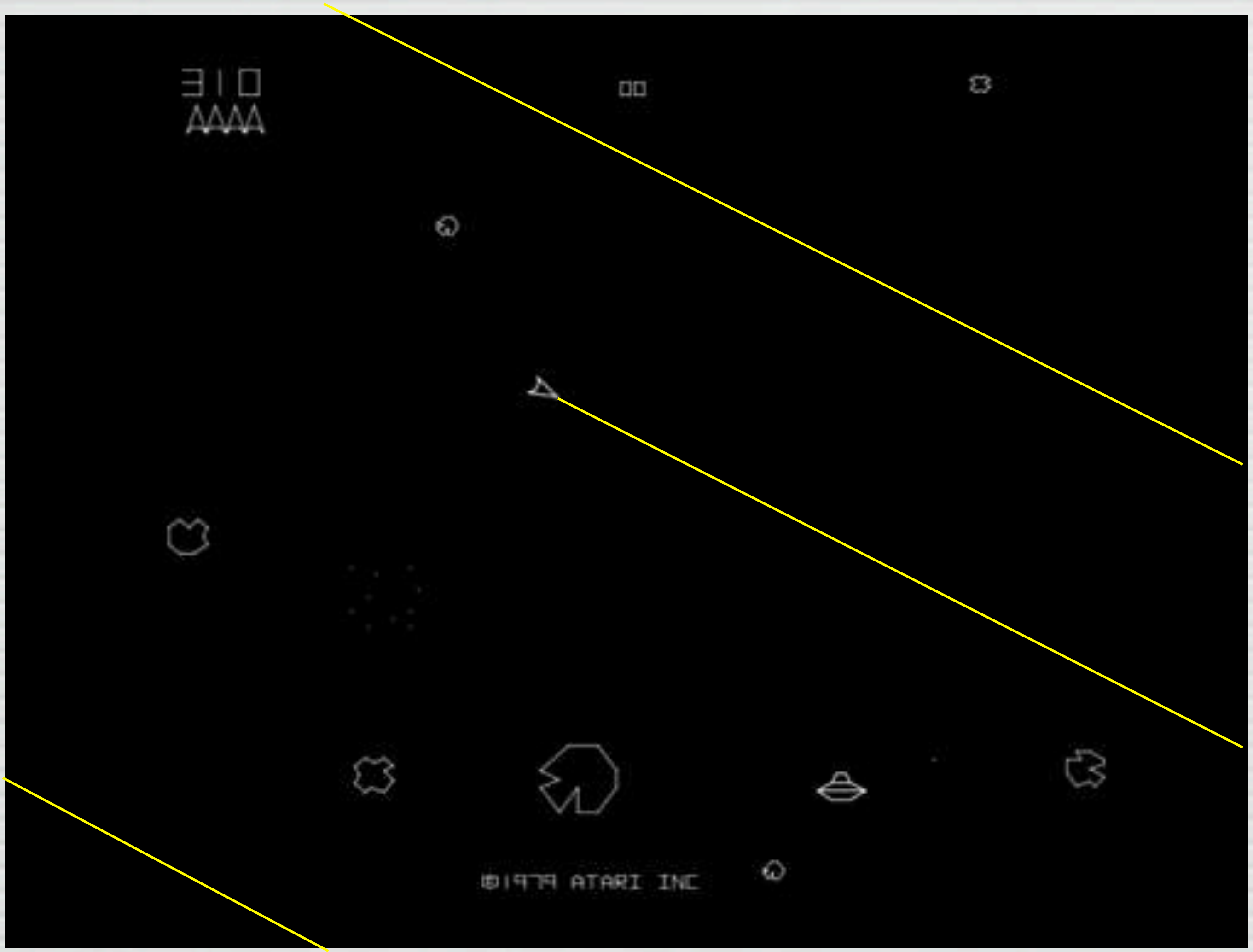


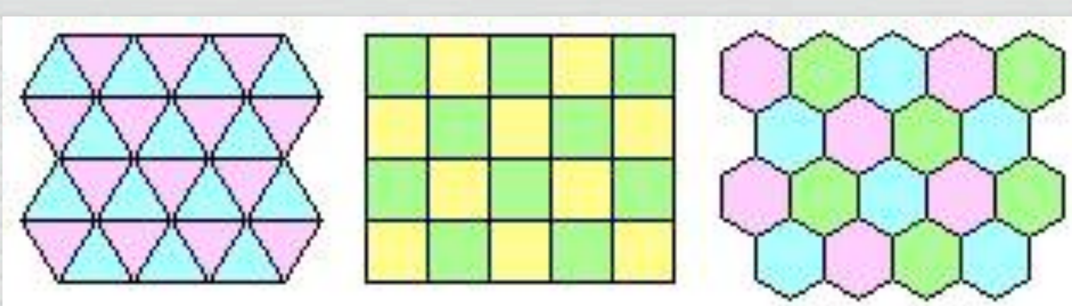




ICIC

Topology in a flat “universe”

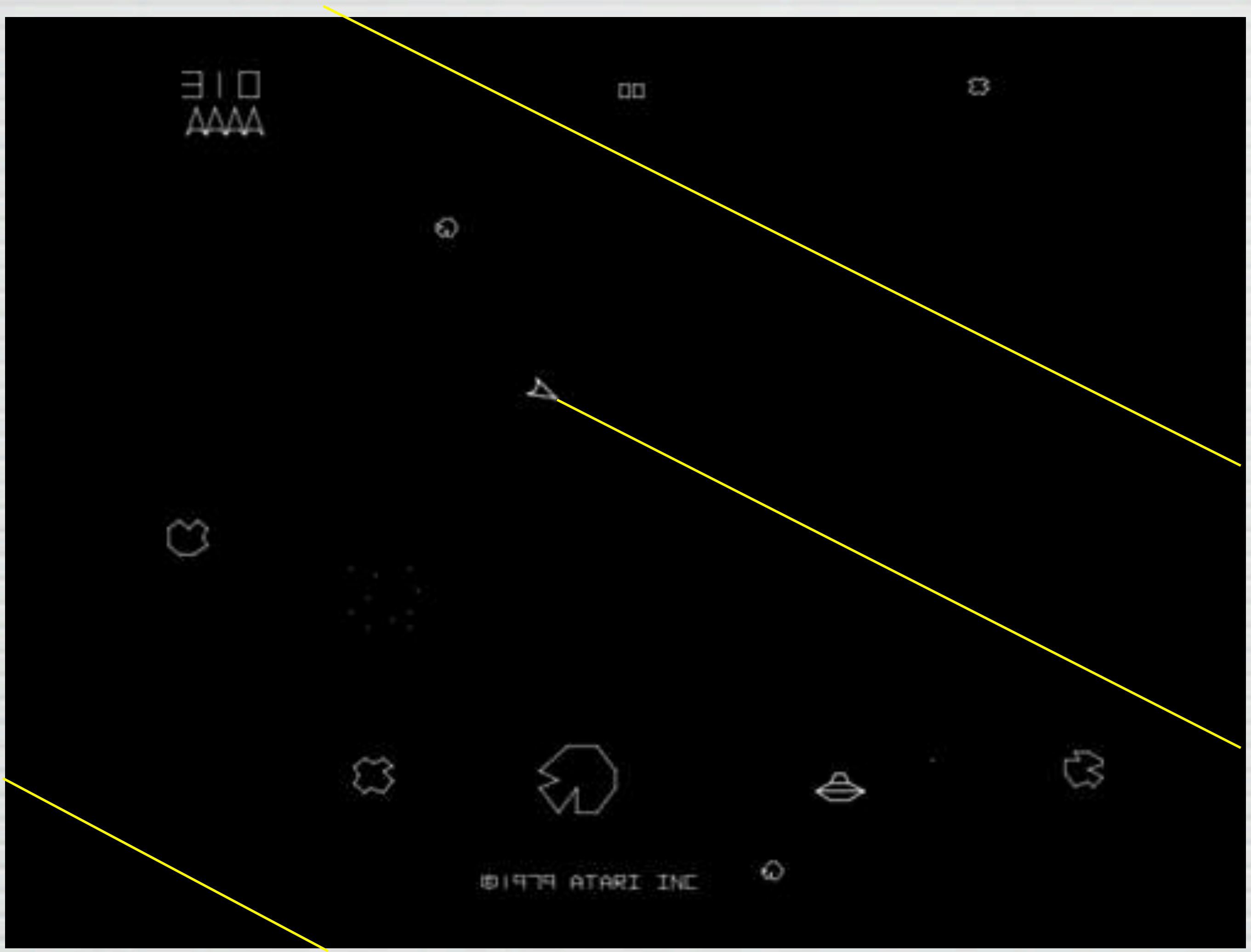


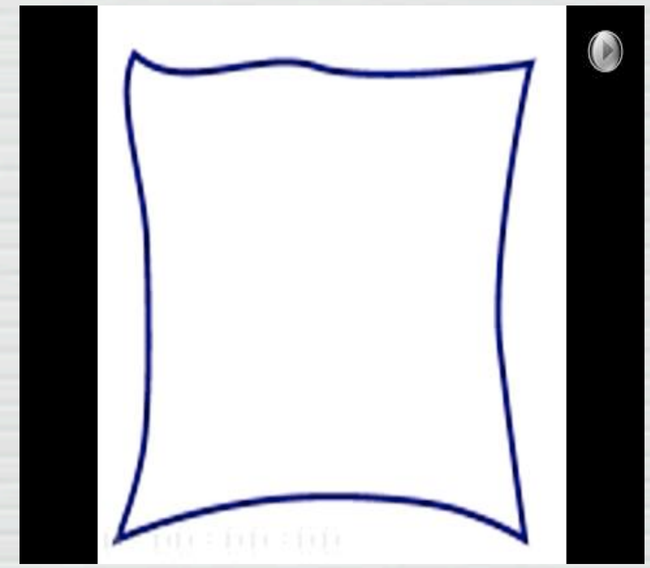


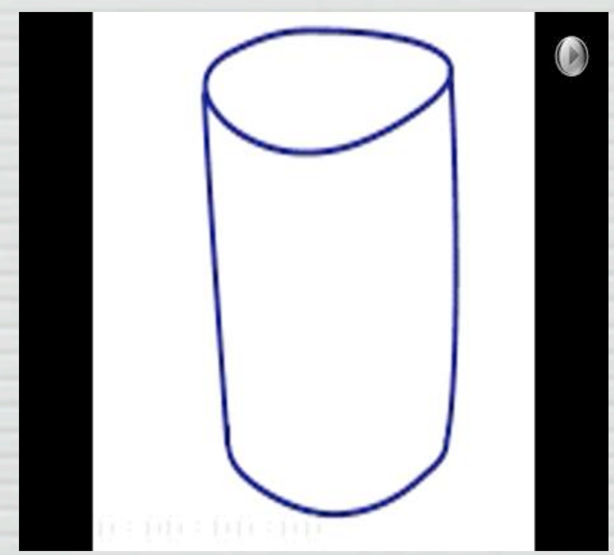
“tiling the plane”

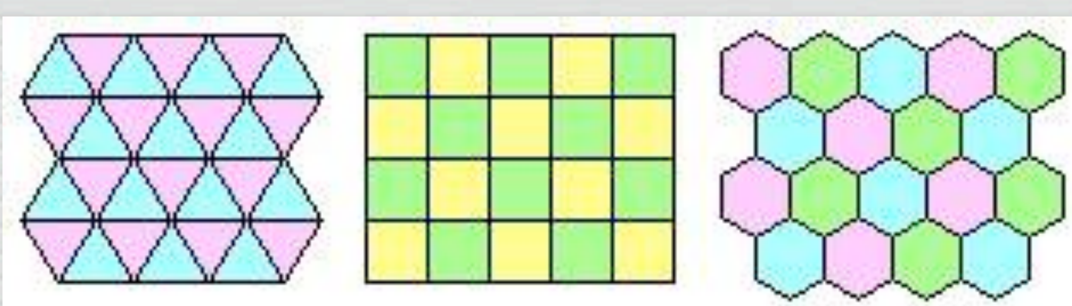
Don’t *need* to “embed” the square to have a connected topology.

Topology in a flat "universe"







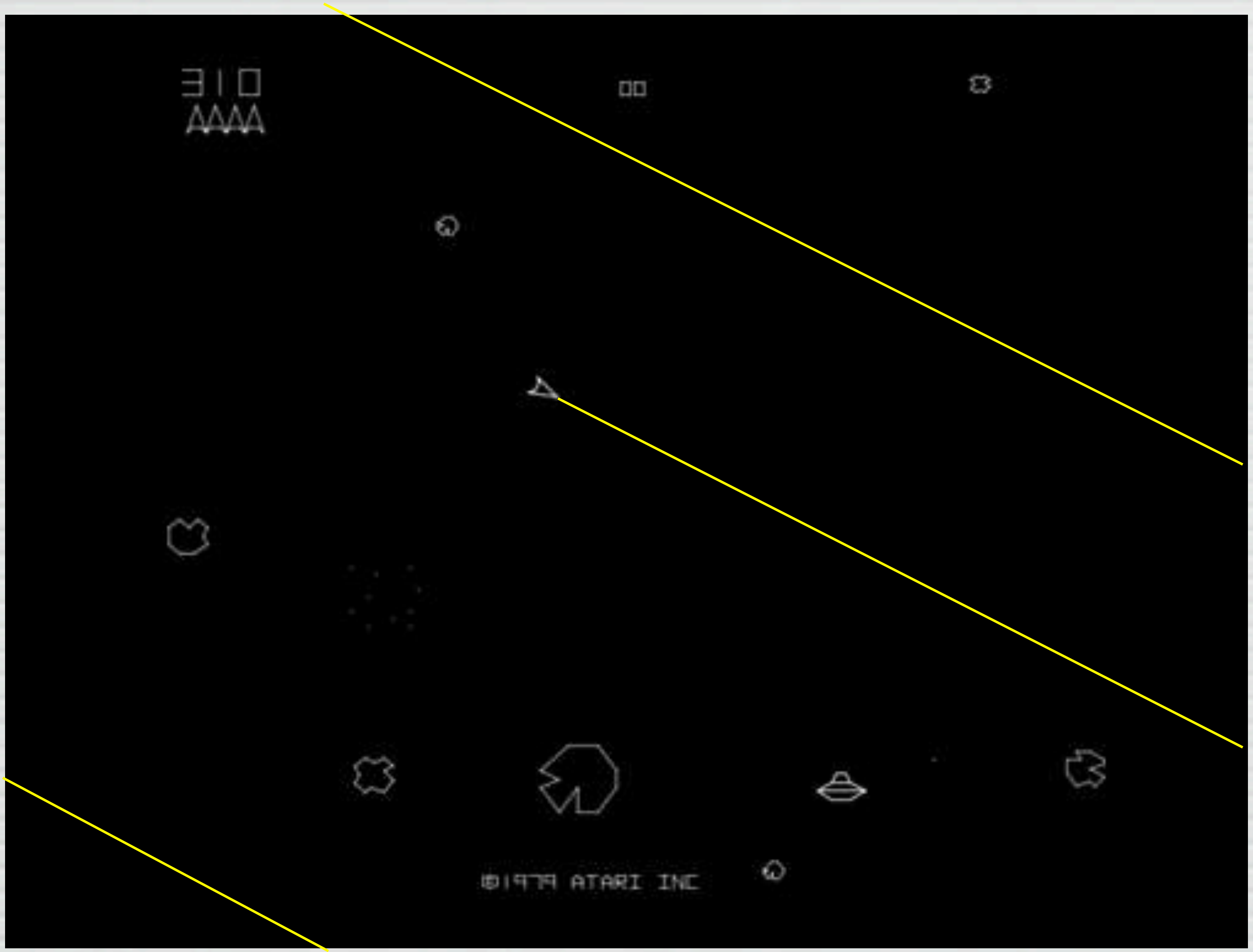


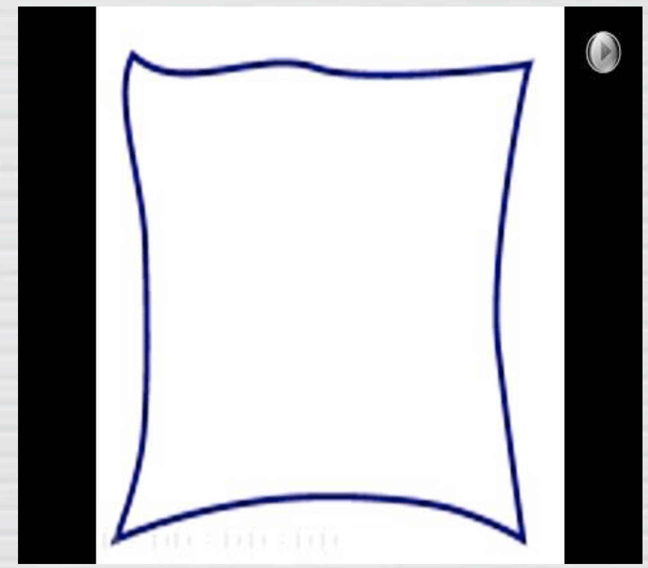
"tiling the plane"

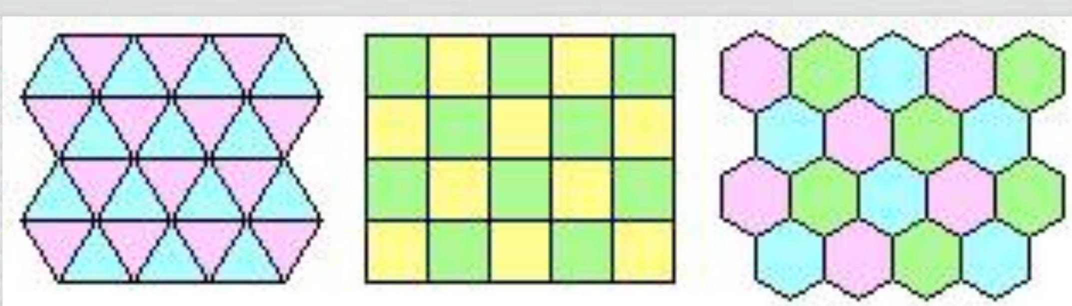
Don't *need* to "embed" the square to have a connected topology.

ICIC

Topology in a flat "universe"



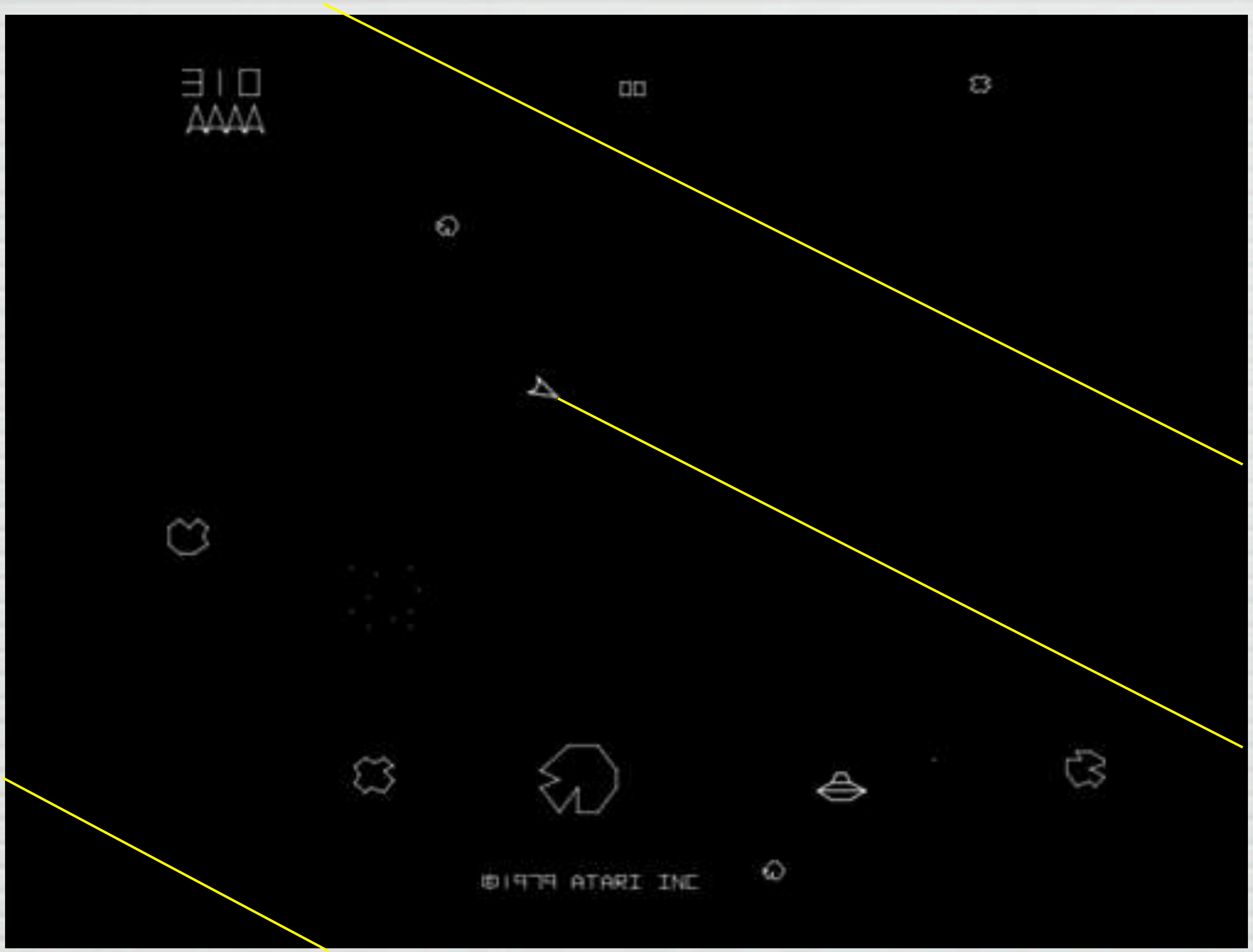


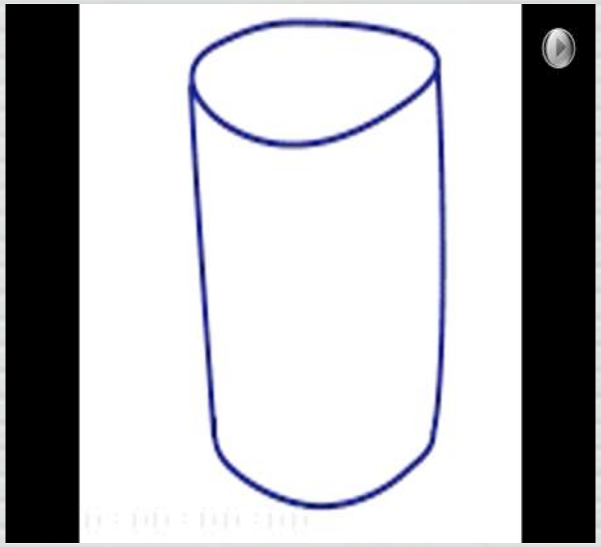


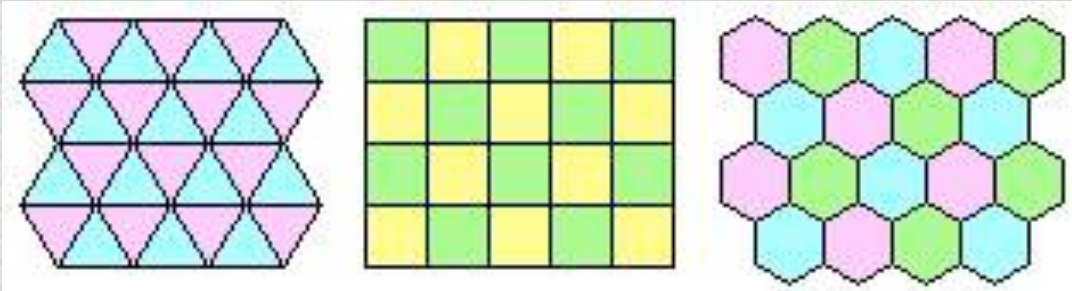
"tiling the plane"

Don't *need* to "embed" the square to have a connected topology.

Topology in a flat "universe"



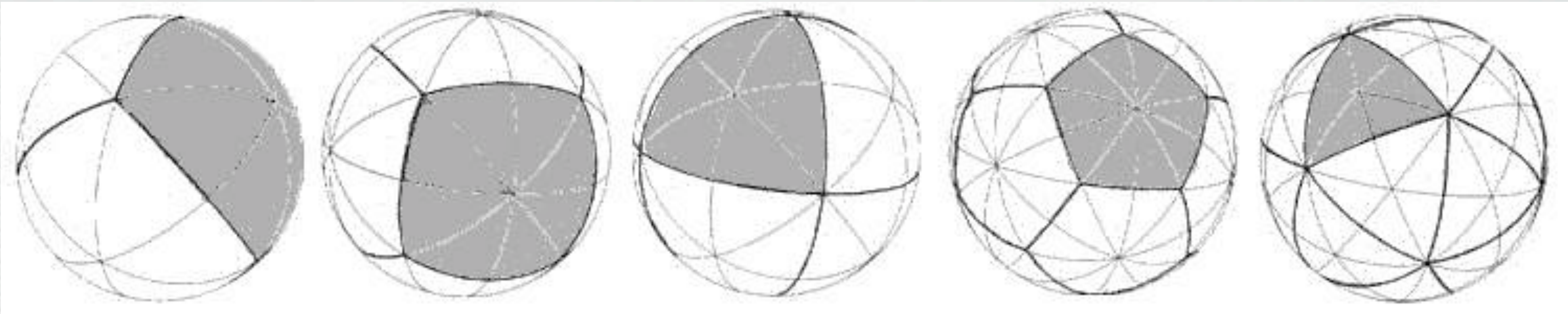




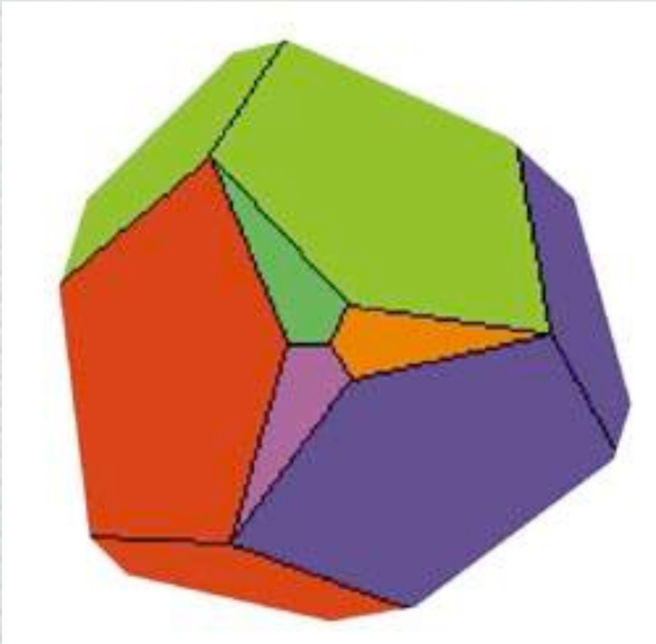
"tiling the plane"

Don't *need* to "embed" the square to have a connected topology.

Topology + geometry

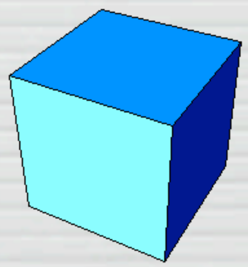
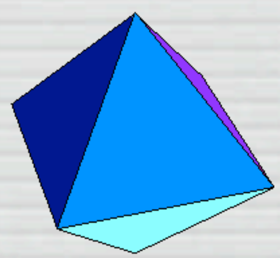
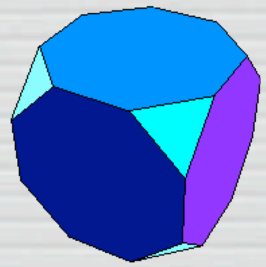
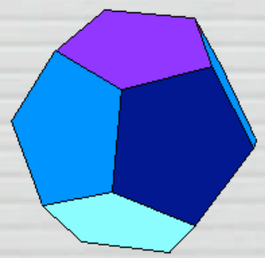


- Tile the 2-sphere with different fundamental domains
  - (Each of these has a 3-sphere analogy)
- Can also tile the hyperbolic universe:
  - (Bond, Pogosyan, etc.)



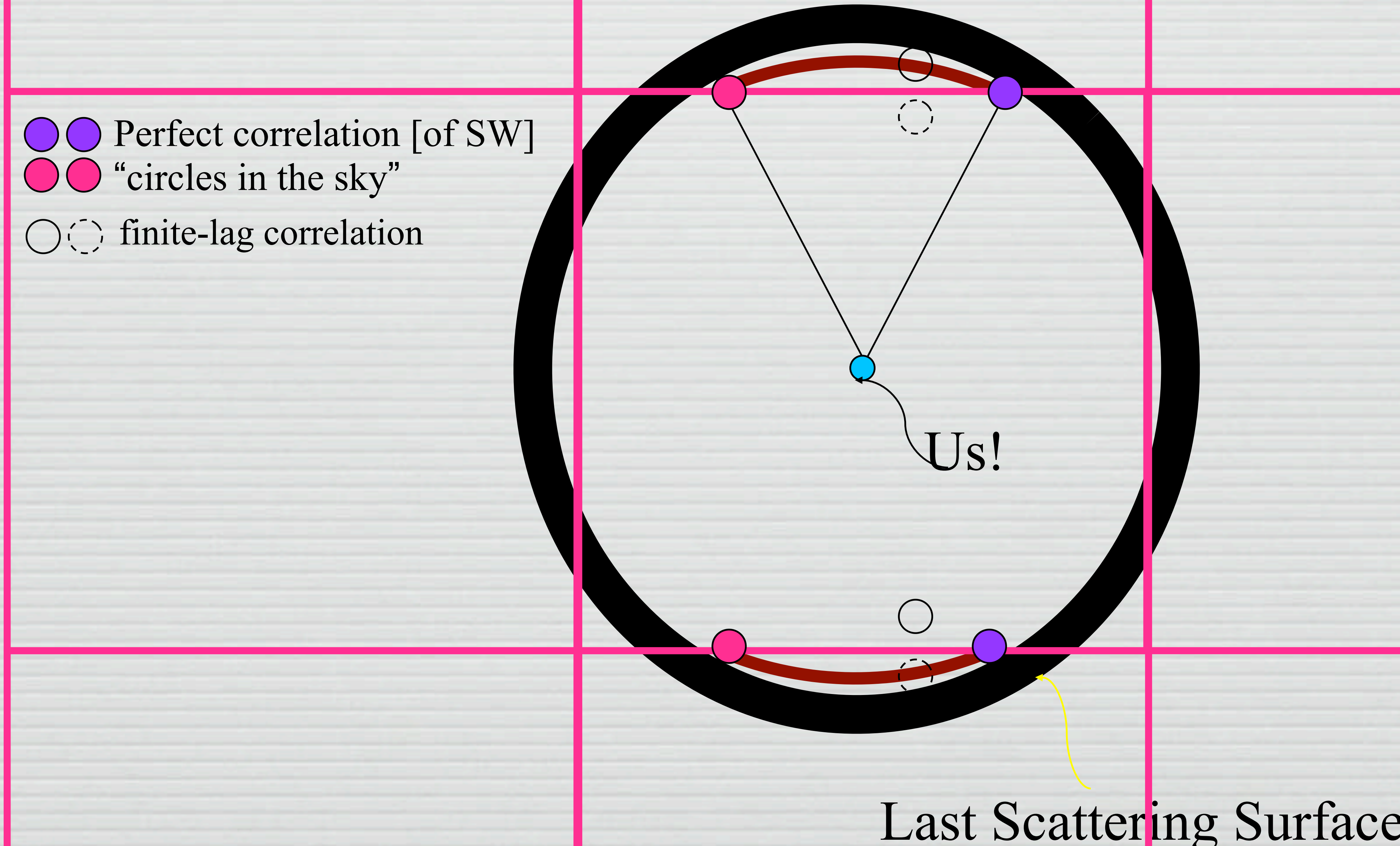
http://www.sciencenews.org/pages/sn_arc98/2_21_98/bob1.htm

Multiply-connected Spherical Topologies

Space	Fundamental group	Order	Elements	F.P.
Quaternionic	Binary Dihedral	8	order 2 rotations about 2 perpendicular axes	
Octahedral	Binary Tetrahedral	24	symmetries of r. tetrahedron	
Truncated Cube	Binary Octahedral	48	symmetries of r. octahedron	
Poincaré	Binary Icosahedral	120	symmetries of r. icosahedron	

ICIC

Measuring Topology with the CMB

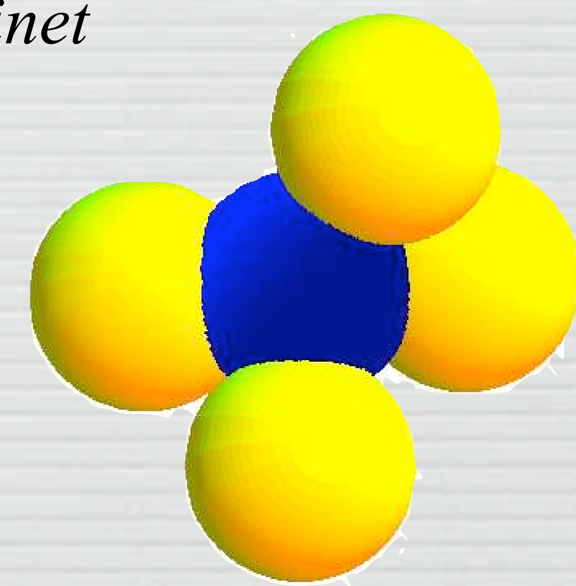


ICIC

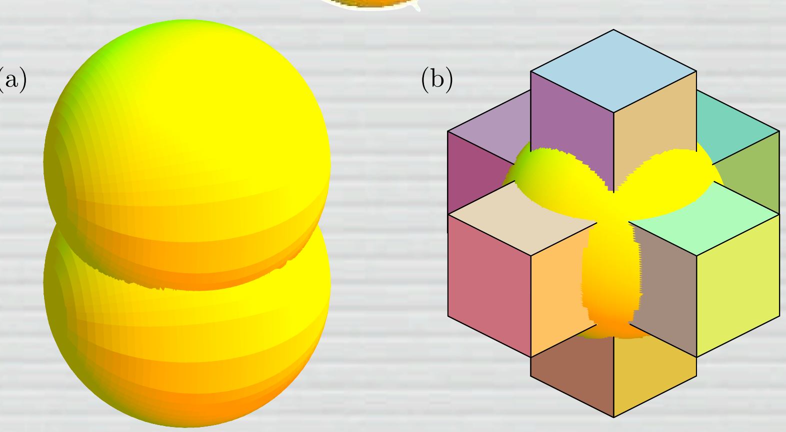
Topology in the CMB

Reviews:
Levin; Lachieze-Rey, Lehoucq, Luminet

- Look for repeated patterns
- Generic & specific methods
- matching patches (e.g., Levin et al)
  - method of images (e.g., Bond et al)
  - assumes infinitely thin LSS
  - mostly open Universes



- Circles in the sky (Cornish, Spergel, Starkman)
  - looks for LSS structure; ignores different views of the same point
  - nb. generic methods work as frequentist null tests but need comparison w/ specific topologies to get *statistics*
    - even Bayesians need to do exploratory statistics
  - Cornish et al '04: "fewer than 1 in 100 random skies generate a false match" [??]: limit out to 24 Gpc



Topology: methods

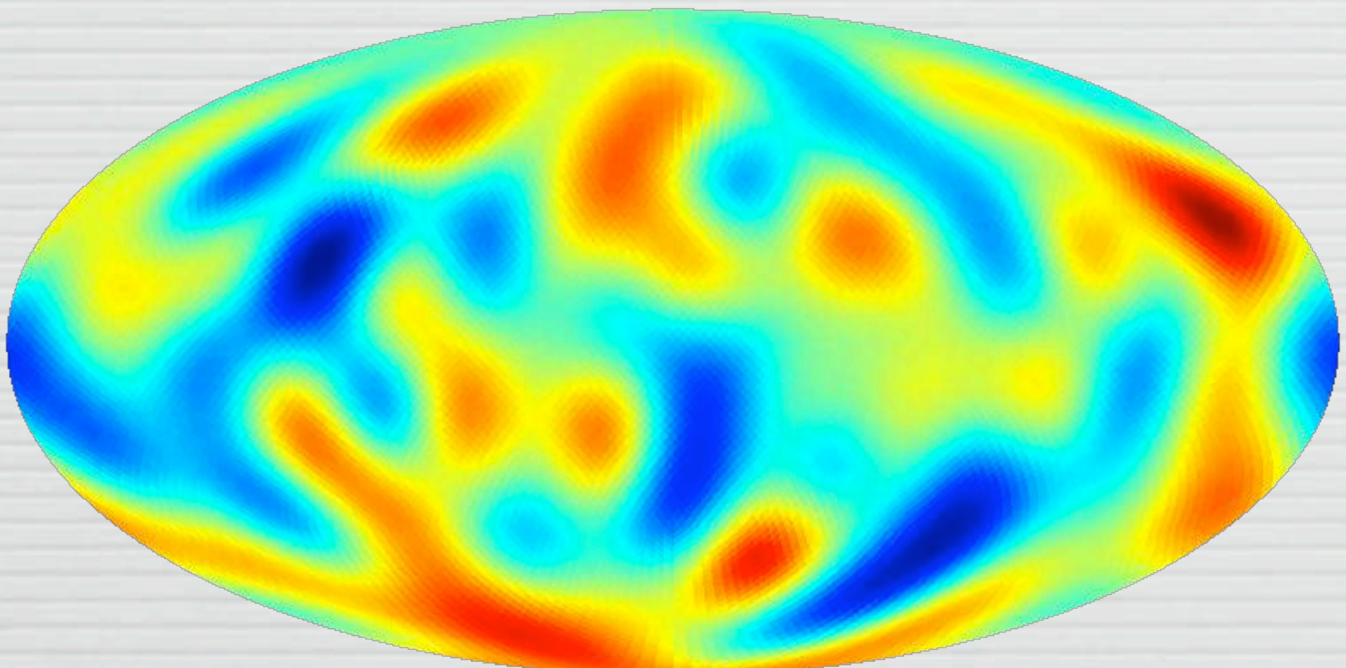
- When topological scale ≲ Horizon scale, induce anisotropic correlations (and suppress power) on large scales

- Direct search for matched circles
  - sensitive to topology with parallel matched surfaces
- Explicit Likelihood
  - calculate correlation matrix for specific topologies.
  - 3d Gaussian with $\langle \delta_{\boldsymbol{k}} \delta_{\boldsymbol{k'}} \rangle = (2\pi)^3 \delta_D(\boldsymbol{k}+\boldsymbol{k'}) P(k)$ w/ $\boldsymbol{k}$ restricted to fundamental domain with boundary conditions
  - induced CMB correlations depend on topology (incl. orientation)

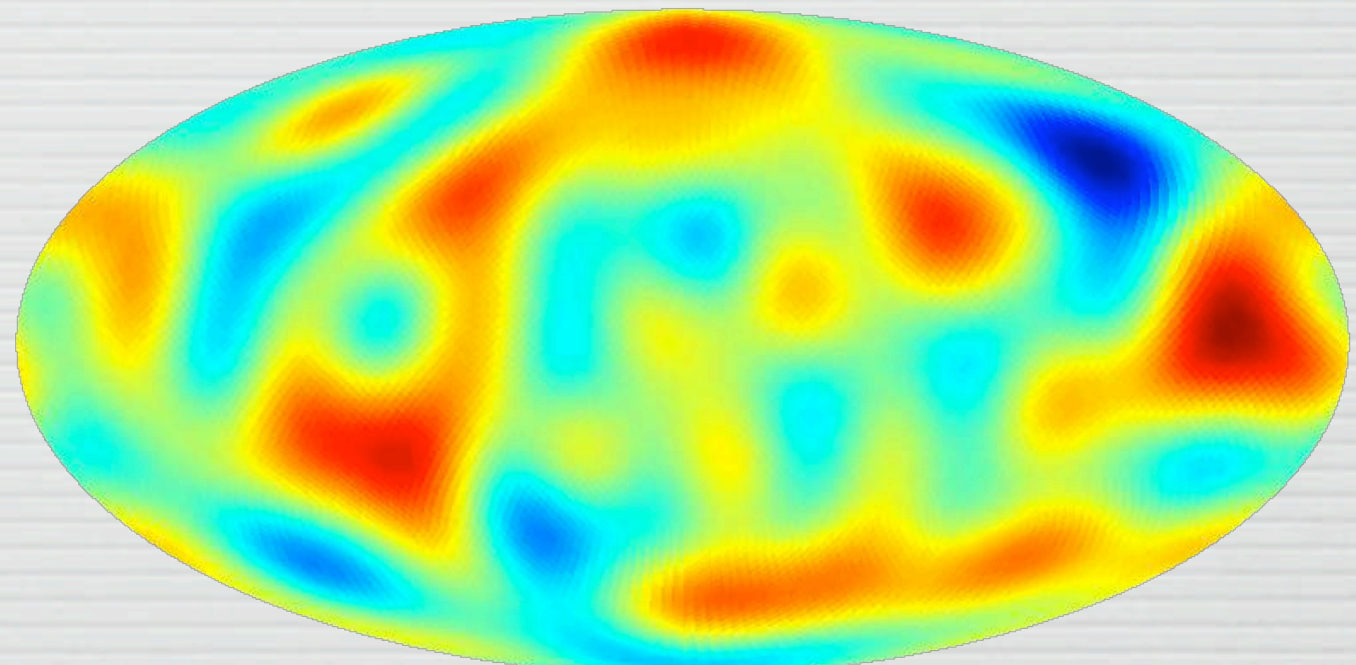
ICIC

Simulated Maps (Ω_k = - 0.063)

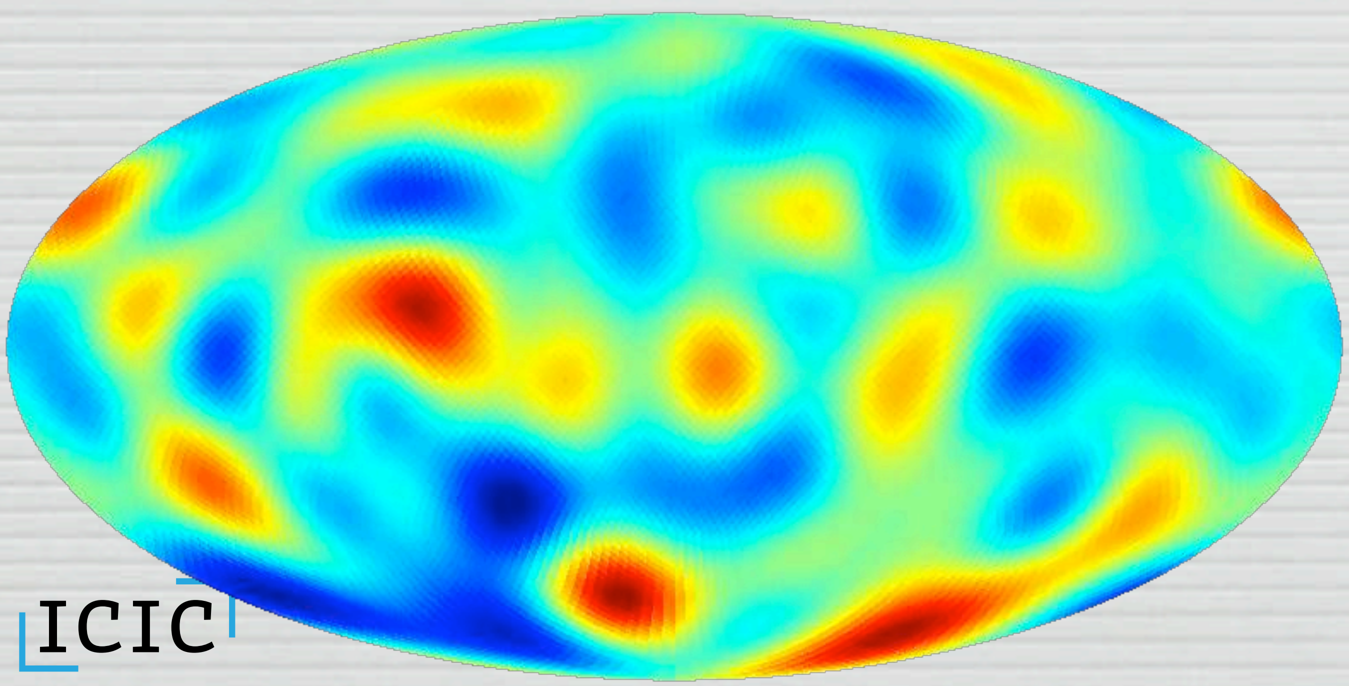
Quaternionic/bi-dehedral



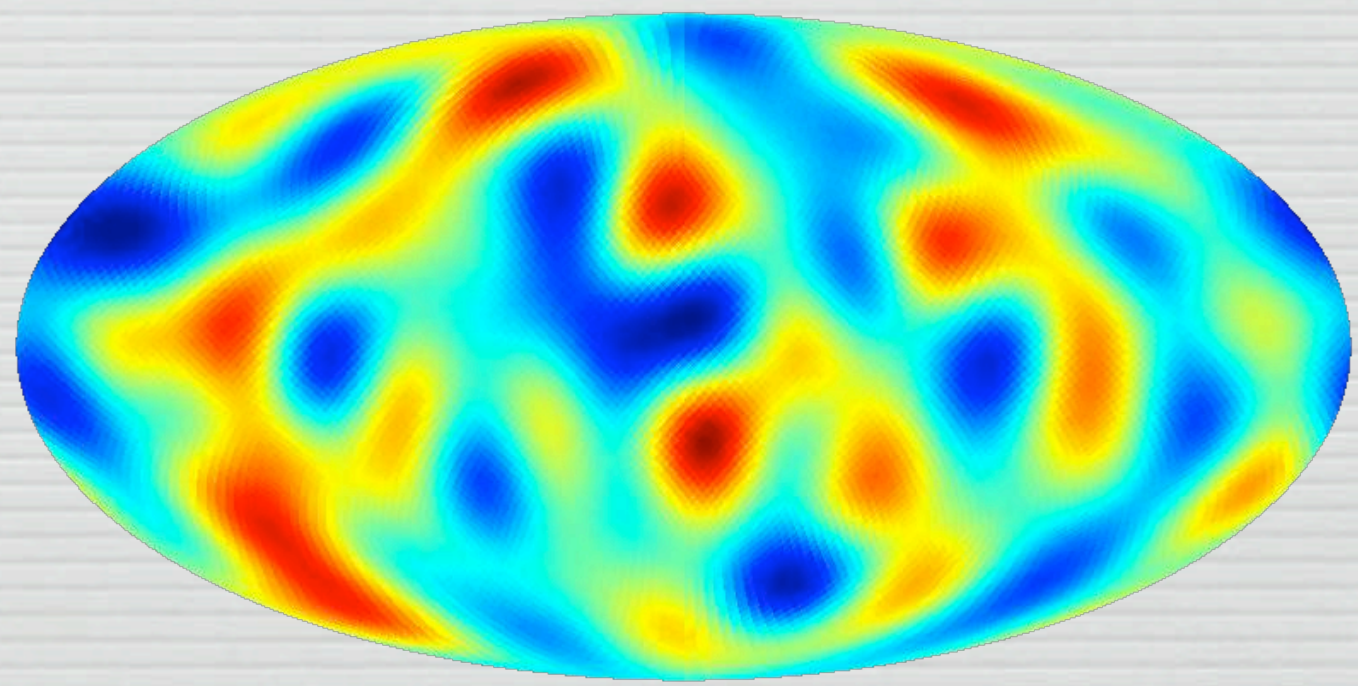
Octahedral/bi-tetrahedral



Truncated cube/bi-octahedral

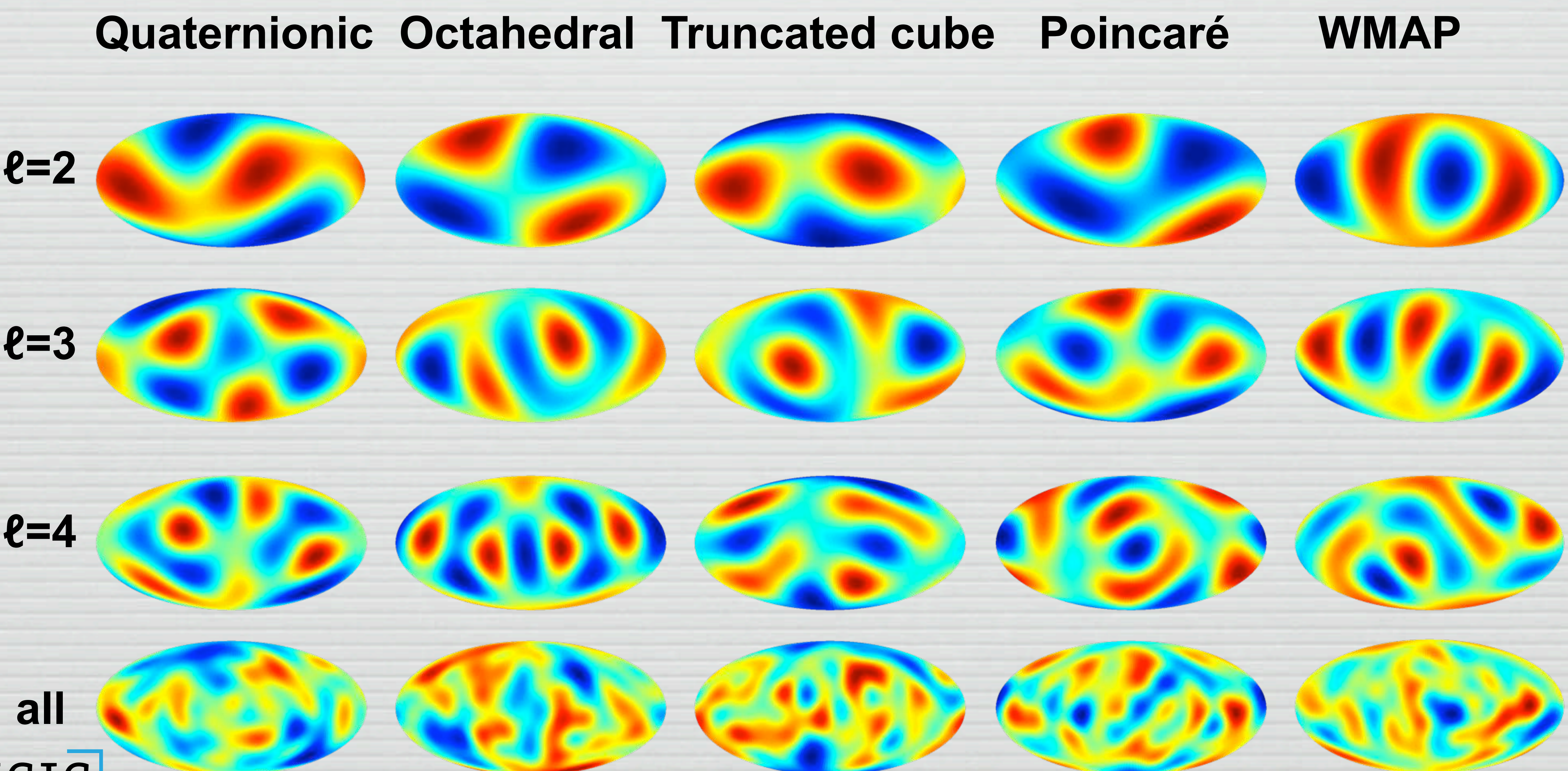


Poincaré/icosahedral



ICIC

Lowest multipoles



Bayesian topology

$$P(a|C) = \frac{1}{|2\pi C|^{1/2}} \exp\left(-\frac{1}{2} a^T C^{-1} a\right)$$

- Full correlation matrix:
- $C = \langle a_{\ell m}\, a_{\ell' m'} \rangle = C_{\ell\ell' mm'} = C$(cosmology, topology)
- $$C_{\ell\ell'}^{mm'} \propto \int d^3k \Delta_\ell(k,\Delta\eta)\Delta_{\ell'}(k,\Delta\eta)P(k) \to \sum_{\mathbf{n}} \Delta_\ell(k_n,\Delta\eta)\Delta_{\ell'}(k_n,\Delta\eta)P(k_n)Y_{\ell m}(\hat{\mathbf{n}})Y^*_{\ell' m'}(\hat{\mathbf{n}}),$$
- $a = a_{\ell m}$
  - (Noise irrelevant on scales of interest)
  - Suppressed power ⇒ stronger correlations

ICIC

Pixel correlations

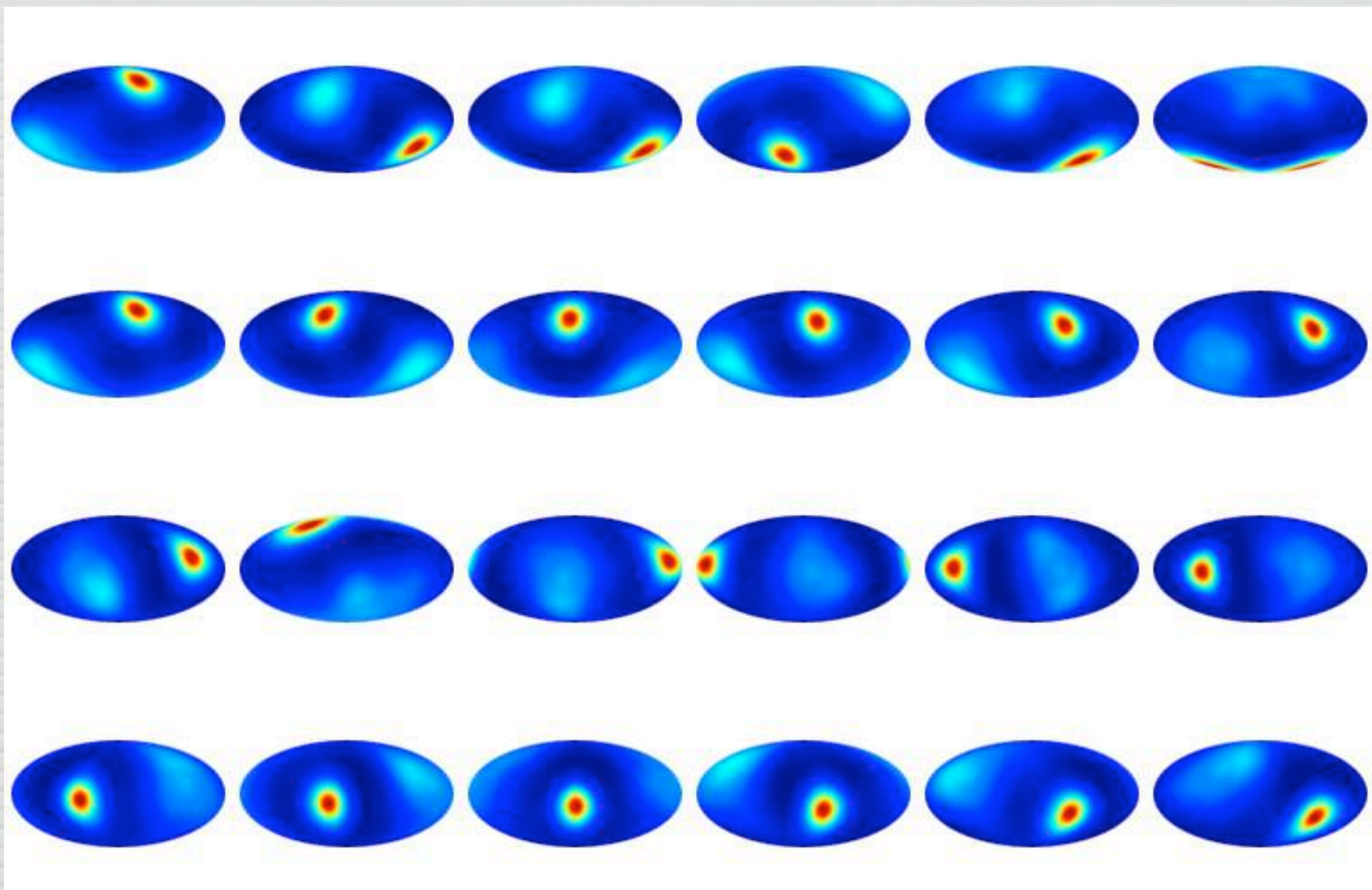
- Octahedral:

$$C_{pp'} = \left\langle \frac{\Delta T}{T}(\hat{x}_p)\frac{\Delta T}{T}(\hat{x}_{p'}) \right\rangle = \sum_{\ell\ell' mm'} C_{\ell\ell' mm'} B_\ell B_{\ell'} Y_{\ell m}(\hat{x}_p) Y_{\ell' m'}(\hat{x}_{p'})$$

$$\rightarrow \sum_\ell \frac{2\ell+1}{4\pi} C_\ell B_\ell^2 P_\ell(\hat{x}_p \cdot \hat{x}_{p'})$$

  - h=0.64, Ω_k = - 0.017

Rows of the correlation matrix:



Pixel correlations

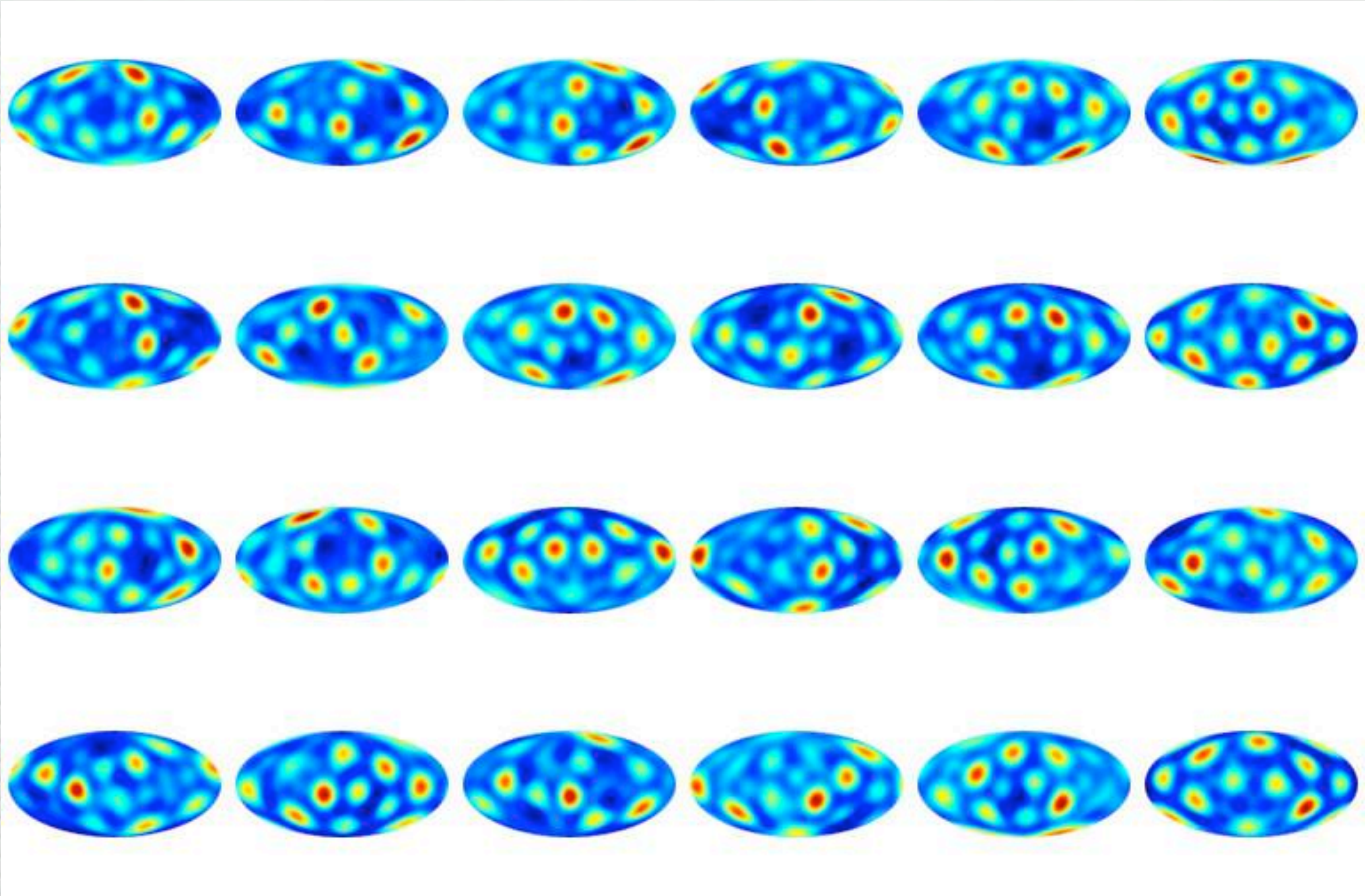
- Poincaré:

$$C_{pp'} = \left\langle \frac{\Delta T}{T}(\hat{x}_p)\frac{\Delta T}{T}(\hat{x}_{p'}) \right\rangle = \sum_{\ell\ell' mm'} C_{\ell\ell' mm'} B_\ell B_{\ell'} Y_{\ell m}(\hat{x}_p) Y_{\ell' m'}(\hat{x}_{p'})$$

$$\rightarrow \sum_\ell \frac{2\ell+1}{4\pi} C_\ell B_\ell^2 P_\ell(\hat{x}_p \cdot \hat{x}_{p'})$$

  - h=0.52, Ω_k = - 0.063

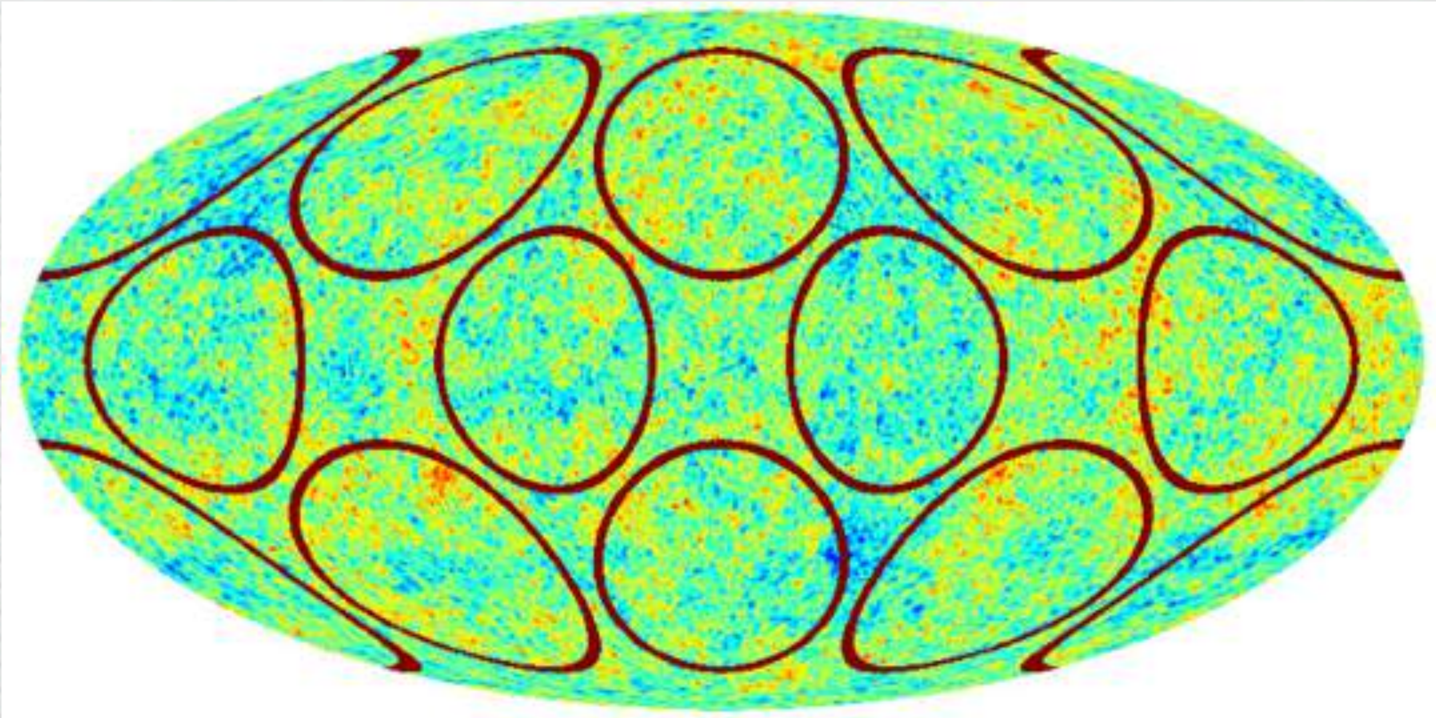
Rows of the correlation matrix:



ICIC

Topology from Planck

- "Matched circles" in a simulated Universe:

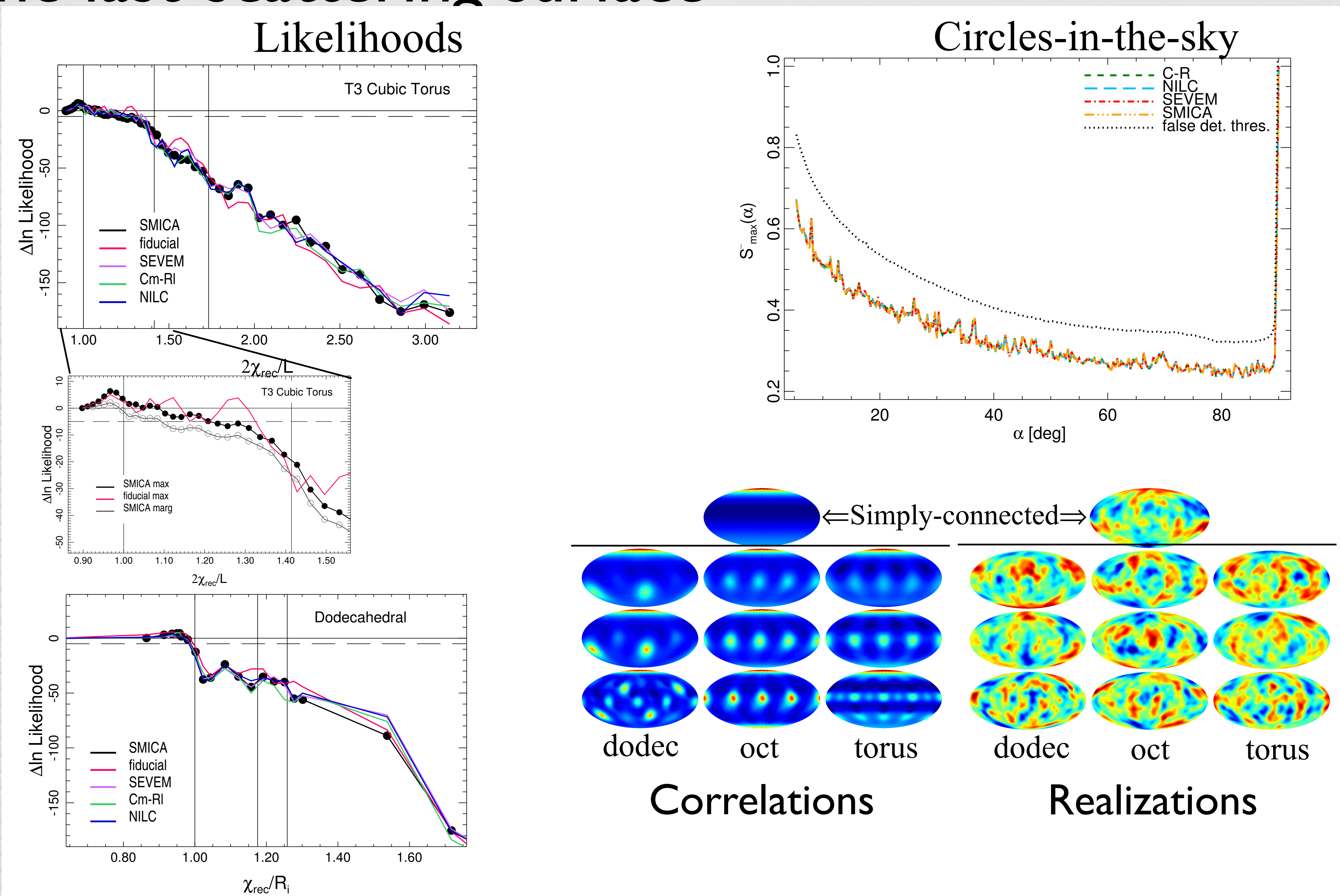


- Alas, not found… we can limit the size of the "fundamental cube" to be greater than the size of the surface we observe with the CMB:
  - side $L \gtrsim 26$ Gpc

ICIC

Topology: results

- No strong evidence for topology on the scale of the last-scattering surface



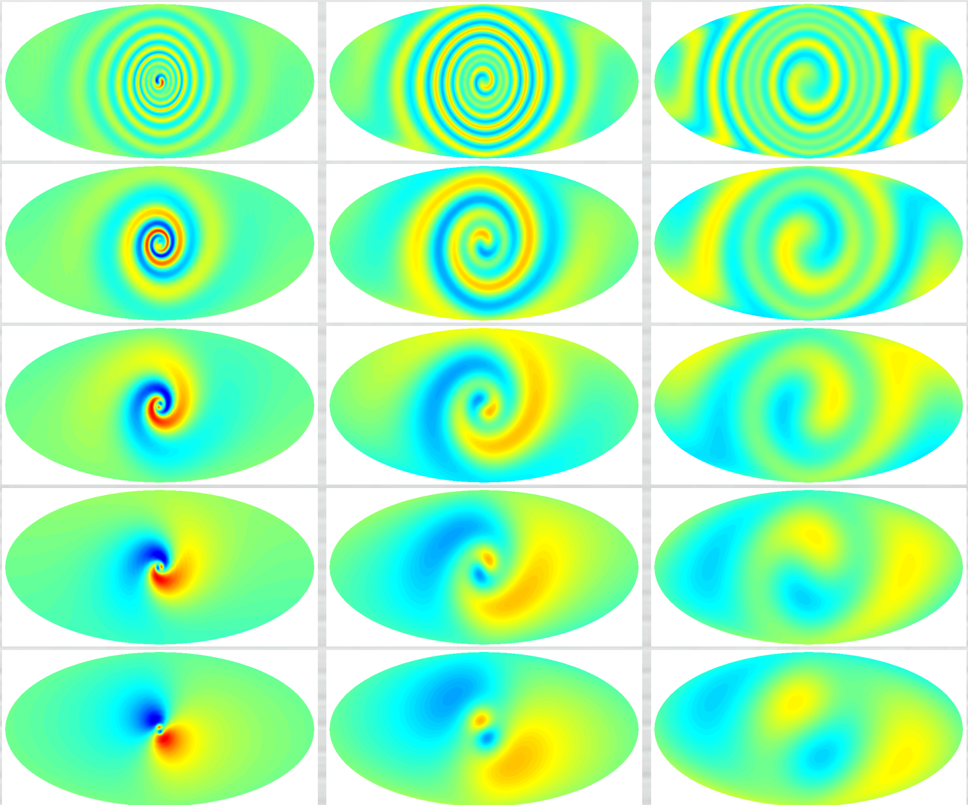
ICIC

Bianchi Models

- Homogeneous, anisotropic spaces
- VII_h: global shear and rotation
  - parameter h relates vorticity ω_i to shear σ_{ij}, Ω_{tot}

$$\left(\frac{\omega}{H}\right)_0 = \frac{(1+h)^{1/2}(1+9h)^{1/2}}{6h}\frac{1-\Omega_{\rm tot}}{\Omega_{\rm tot}}\sqrt{\left(\frac{\sigma_{12}}{H}\right)_0^2+\left(\frac{\sigma_{13}}{H}\right)_0^2}$$

- Focusing induces specific pattern of temperature anisotropy on large scales
- Full likelihood calculation (Gaussian added to deterministic template)
  - consistent cosmology very low likelihood



ICIC

Bianchi Models

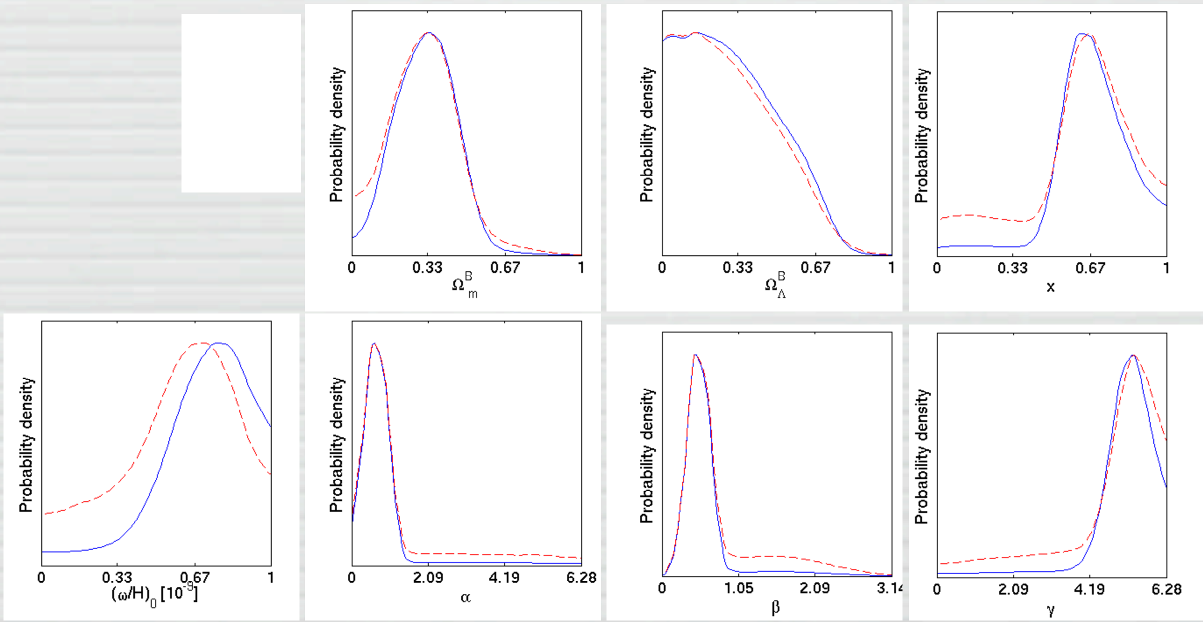
Flat-decoupled

Bianchi Parameter	SMICA MAP	SMICA Mean	SEVEM MAP	SEVEM Mean
Ω_m^B	0.38	0.32 ± 0.12	0.35	0.31 ± 0.15
Ω_Λ^B	0.20	0.31 ± 0.20	0.22	0.30 ± 0.20
x	0.63	0.67 ± 0.16	0.66	0.62 ± 0.23
$(\omega/H)_0$	8.8×10^{-10}	$(7.1 \pm 1.9) \times 10^{-10}$	9.4×10^{-10}	$(5.9 \pm 2.4) \times 10^{-10}$
α	38.8°	51.3° ± 47.9°	40.5°	77.4° ± 80.3°
β	28.2°	33.7° ± 19.7°	28.4°	45.6° ± 32.7°
γ	309.2°	292.2° ± 51.9°	317.0°	271.5° ± 80.7°

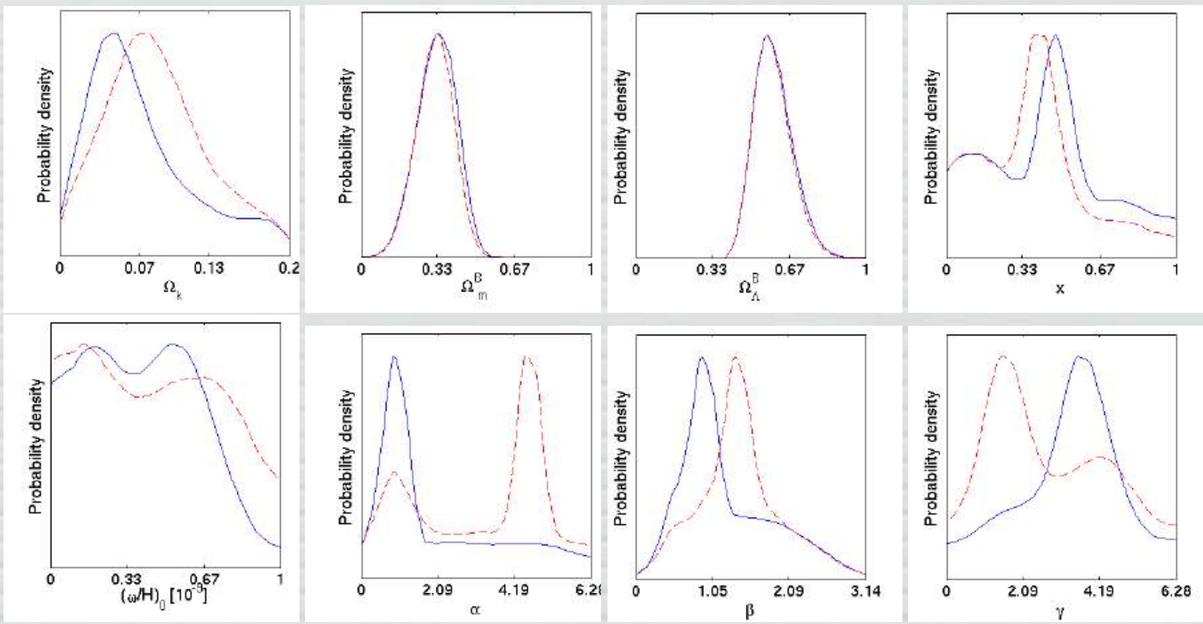
Open-coupled

Bianchi Parameter	SMICA MAP	SMICA Mean	SEVEM MAP	SEVEM Mean
Ω_k	0.05	0.07 ± 0.05	0.09	0.08 ± 0.04
Ω_m^B	0.41	0.33 ± 0.07	0.41	0.32 ± 0.07
Ω_Λ^B	0.55	0.60 ± 0.07	0.50	0.59 ± 0.07
x	0.46	0.44 ± 0.24	0.38	0.39 ± 0.22
$(\omega/H)_0$	5.9×10^{-10}	$(4.0 \pm 2.4) \times 10^{-10}$	9.3×10^{-10}	$(4.5 \pm 2.8) \times 10^{-10}$
α	57.4°	122.5° ± 96.0°	264.1°	188.6° ± 98.7°
β	54.1°	70.8° ± 35.5°	79.6°	81.1° ± 31.7°
γ	202.6°	193.5° ± 77.4°	90.6°	160.4° ± 91.1°

Flat-coupled: $\omega_0/H_0 < 8.1 \times 10^{-10}$ (95%)

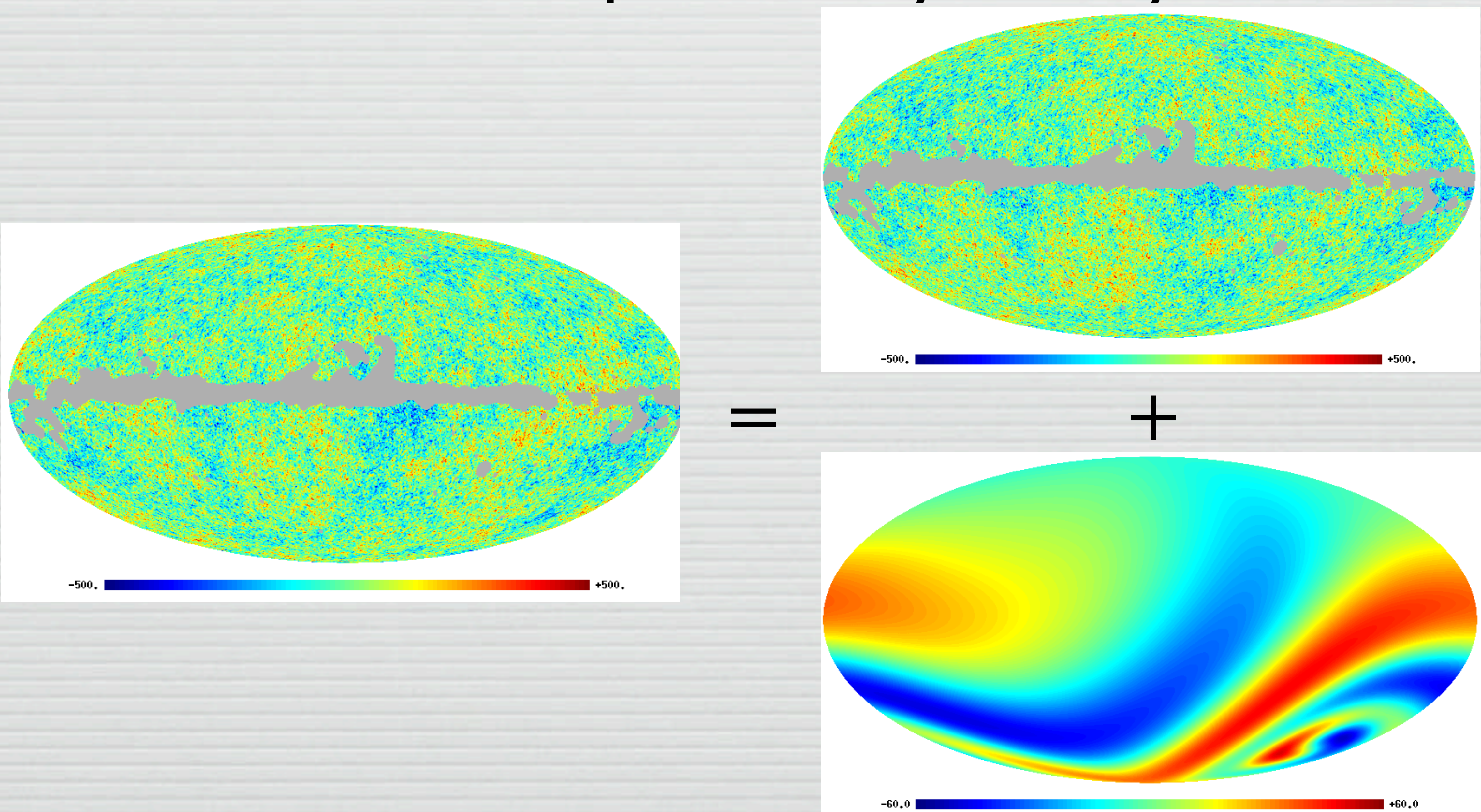


(a) Flat-decoupled-Bianchi model.



(b) Open-coupled-Bianchi model.

- The Bianchi VII$_h$ uncoupled "model" accounts for much of the hemispherical asymmetry



Non-gaussianity

- Another way to go beyond (and check) the simple assumptions
- In general, can't write down "the" distribution for a parameter with some specific set of higher-order moments. (cf. yesterday's discussion of Gaussian as MaxEnt)
- Use frequentist estimators...
- In the absence of a specific model, want to determine phenomenological parameters describing departure from an isotropic multivariate gaussian distribution.
  - e.g., moments — but not unique (there is no distribution that has mean, variance, skewness, but no higher moments)
    - for (suitably defined) small non-gaussianity, third-order moments should dominate
    - full determination of 3-pt function is computationally infeasible (and we lack sufficient S/N)
    - parameterize non-gaussianity

ICIC

non-Gaussianity: f_{NL}

- Heuristically $\phi = \phi_G + f_{\rm NL}(\phi_G^2 - \langle\phi_G^2\rangle)$
  for a Gaussian ϕ_G (e.g., multi-field inflation)
  - This is the (spatially) local model for non-Gaussianity
  - Induces specific 3-d correlations

$$\langle\phi\phi\phi\rangle \sim 3f_{\rm NL}\left(\langle\phi_G\phi_G\phi_G\phi_G\rangle - \langle\phi_G\phi_G\rangle\langle\phi_G\phi_G\rangle\right) + O(f_{\rm NL}^2)$$
$$\sim 6f_{\rm NL}\langle\phi_G\phi_G\rangle\langle\phi_G\phi_G\rangle + O(f_{\rm NL}^2)$$

    and hence 2-d correlations in the CMB
  - Corresponds to Fourier bispectrum $B(k_1, k_2, k_3)$ which peaks in squeezed case $k_1 \ll k_2 \simeq k_3$
    - modulate small-scale structure by large-scale modes
      - cf. galaxy bias
  - More generally, consider other shapes (e.g., equilateral) motivated by specific theories

ICIC

Estimating non-Gaussianity

$$\langle a_{\ell_1 m_1} a_{\ell_2 m_2} a_{\ell_2 m_2} \rangle = \mathcal{G}^{\ell_1 \ell_2 \ell_3}_{m_1 m_2 m_3} b_{\ell_1 \ell_2 \ell_3} = \mathcal{G}^{\ell_1 \ell_2 \ell_3}_{m_1 m_2 m_3} h^{-2}_{\ell_1 \ell_2 \ell_3} B_{\ell_1 \ell_2 \ell_3}$$

- Expect to be able to estimate the third moment by taking some weights average over cubic products of data
  - (cf. quadratic estimators of power spectra)
    - "optimal" (min-var) weights computationally infeasible (Heavens 1998) — average over *all* triples of data
    - ignoring off-diagonal covariance gives somewhat more tractable case (Creminelli et al. 2006).
    - further simplify for "separable" shapes (Komatsu et al=KSW) and linear combinations thereof (Fergusson & Shellard)
    - generalize to S_ℓ=skew-C_ℓ, retains shape information in one ℓ direction (Heavens & Munshi)

ICIC

Non-Gaussianity from Planck

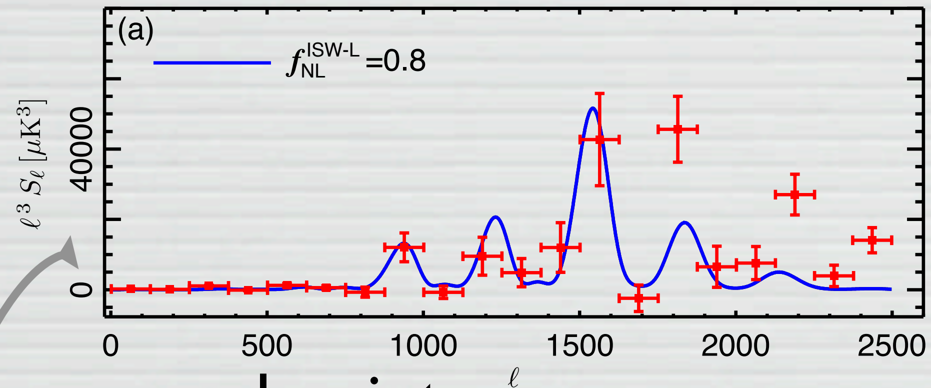
- Planck detects (non-Primordial) non-Gaussianity…

	Independent				ISW-lensing subtracted		
	KSW	Binned	Modal		KSW	Binned	Modal
SMICA							
Local	9.8 ± 5.8	9.2 ± 5.9	8.3 ± 5.9	. . . . .	**2.7 ± 5.8**	2.2 ± 5.9	1.6 ± 6.0
Equilateral	−37 ± 75	−20 ± 73	−20 ± 77	. . . . .	**−42 ± 75**	−25 ± 73	−20 ± 77
Orthogonal	−46 ± 39	−39 ± 41	−36 ± 41	. . . . .	**−25 ± 39**	−17 ± 41	−14 ± 42
NILC							
Local	11.6 ± 5.8	10.5 ± 5.8	9.4 ± 5.9	. . . . .	4.5 ± 5.8	3.6 ± 5.8	2.7 ± 6.0
Equilateral	−41 ± 76	−31 ± 73	−20 ± 76	. . . . .	−48 ± 76	−38 ± 73	−20 ± 78
Orthogonal	−74 ± 40	−62 ± 41	−60 ± 40	. . . . .	−53 ± 40	−41 ± 41	−37 ± 43
SEVEM							
Local	10.5 ± 5.9	10.1 ± 6.2	9.4 ± 6.0	. . . . .	3.4 ± 5.9	3.2 ± 6.2	2.6 ± 6.0
Equilateral	−32 ± 76	−21 ± 73	−13 ± 77	. . . . .	−36 ± 76	−25 ± 73	−13 ± 78
Orthogonal	−34 ± 40	−30 ± 42	−24 ± 42	. . . . .	−14 ± 40	−9 ± 42	−2 ± 42
C-R							
Local	12.4 ± 6.0	11.3 ± 5.9	10.9 ± 5.9	. . . . .	6.4 ± 6.0	5.5 ± 5.9	5.1 ± 5.9
Equilateral	−60 ± 79	−52 ± 74	−33 ± 78	. . . . .	−62 ± 79	−55 ± 74	−32 ± 78
Orthogonal	−76 ± 42	−60 ± 42	−63 ± 42	. . . . .	−57 ± 42	−41 ± 42	−42 ± 42

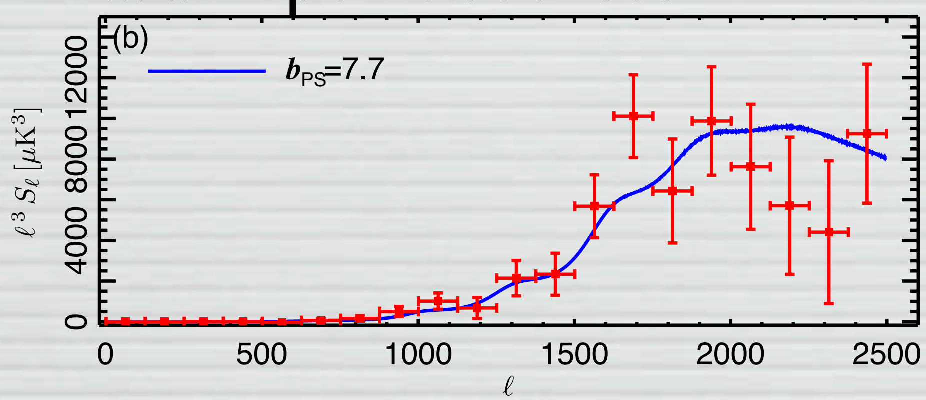
ICIC

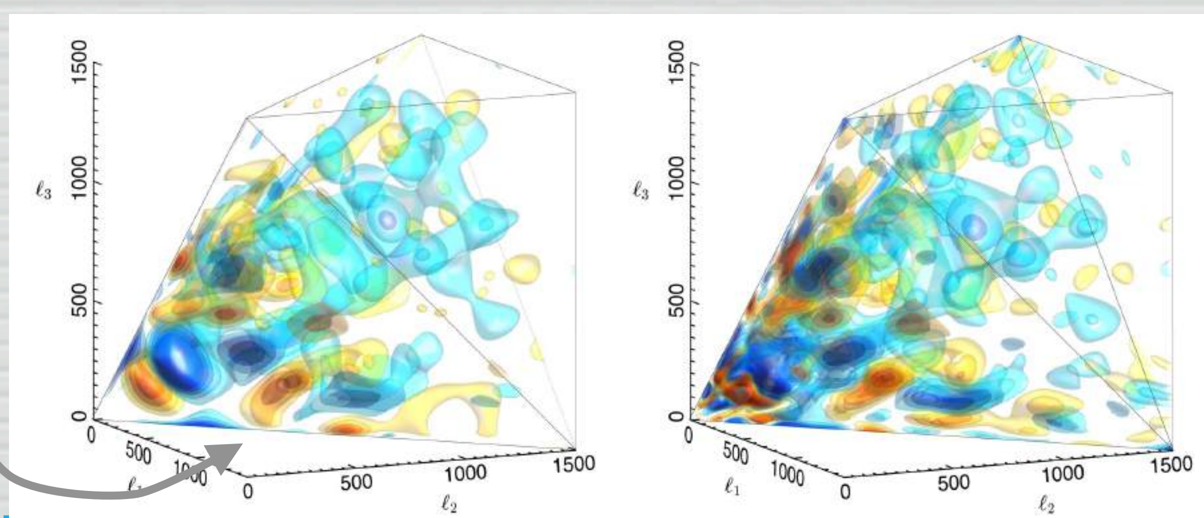
Non-Gaussianity from Planck

We see ISW lensing...

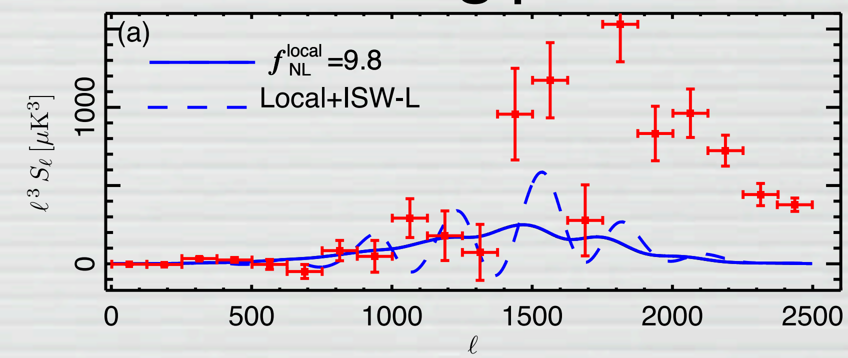


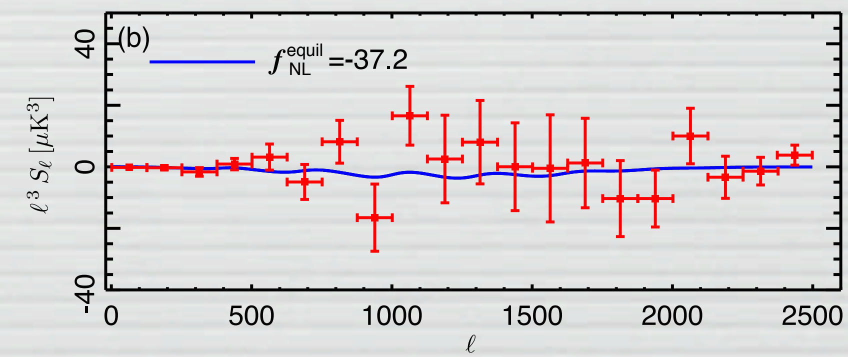
... and point sources

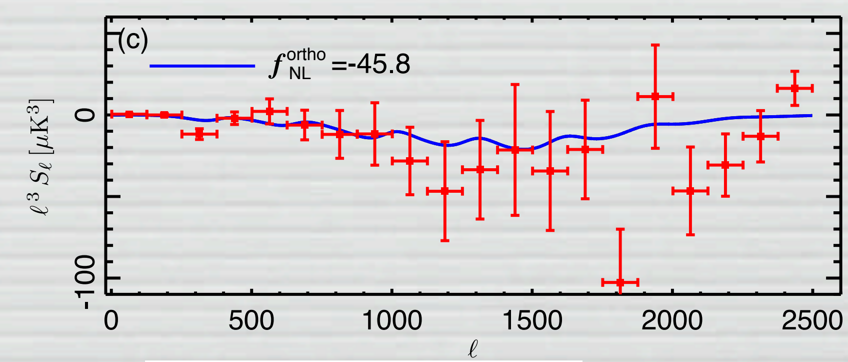


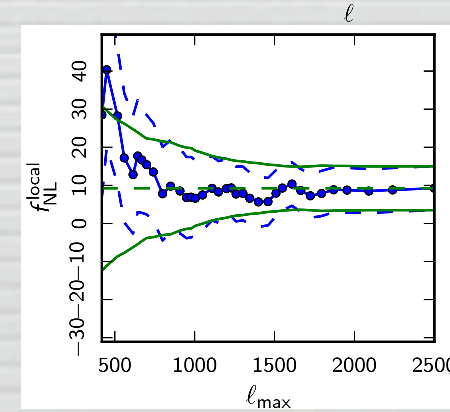


... but nothing primordial









The CMB: A Case Study

- Hierarchical Bayesian formalism
  - raw-data ⇒ maps ⇒ spectra ⇒ parameters
  - radical data compression
  - need to keep track of likelihood function details
- Checking assumptions
  - “anomalies”?
    - No obvious solution by changing the large-scale structure of spacetime (topology, Bianchi)
  - non-Gaussianity
    - lensing, point sources, correlations detected in Planck
    - no evidence yet for primordial non-Gaussianity

ICIC