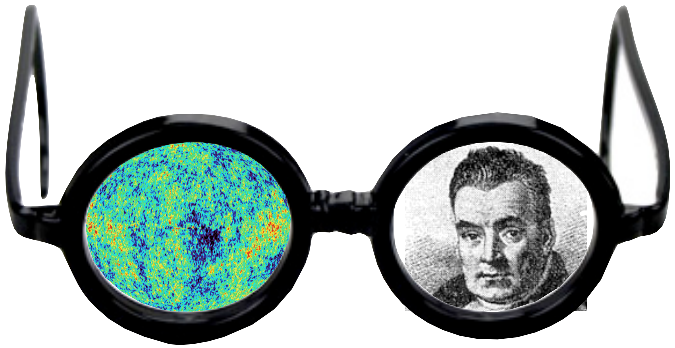
Progress on statistical issues in searches
SLAC, June 2012

Dogs, non-dogs and statistics: (Bayesian) searches in cosmology



Roberto Trotta
Astrophysics, Imperial College London
www.robertotrotta.com

Imperial College London

Contents

Imperial College London

- The Cosmological Concordance Model
- Searches of non-Gaussianity
- Searches of non-trivial topology
- “Generic” departures from LCDM: looking for non-dogs
- Principled Bayesian model selection: pros and cons
- Conclusions and open questions

Imperial Centre for Inference and Cosmology (ICIC)

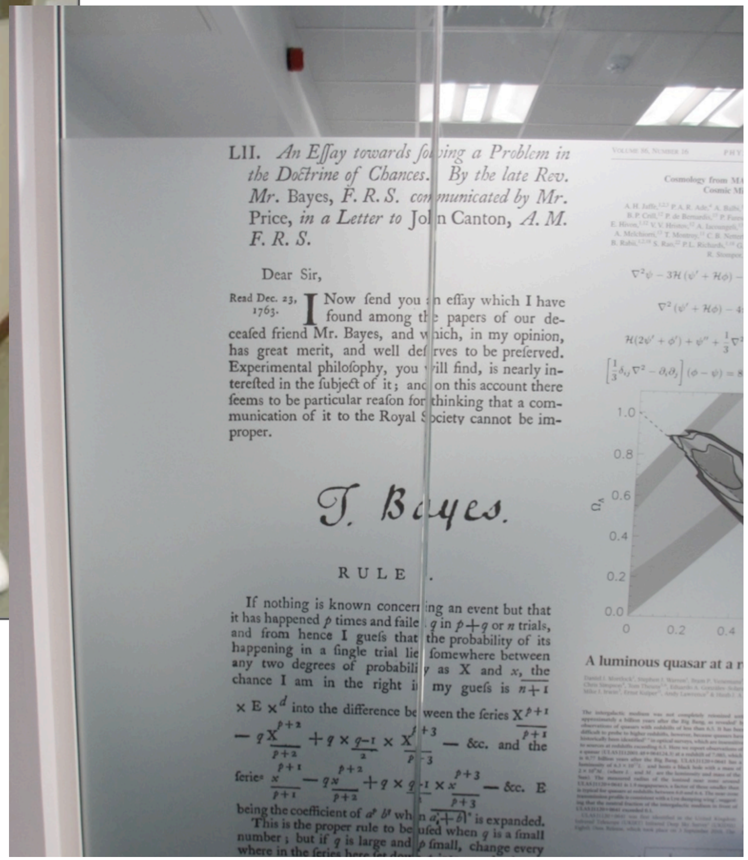
ICIC director:
Alan Heavens

Imperial Centre for Inference and Cosmology (ICIC)

ICIC director:
Alan Heavens

Imperial Centre for Inference and Cosmology (ICIC)

ICIC director:
Alan Heavens



The cosmological concordance model

The ΛCDM cosmological concordance model is built on three pillars:

1. **INFLATION:**
   A burst of exponential expansion in the first ~10^{-32} s after the Big Bang, probably powered by a yet unknown scalar field.
2. **DARK MATTER:**
   The growth of structure in the Universe and the observed gravitational effects require a massive, neutral, non-baryonic yet unknown particle making up ~25% of the energy density - See Jan Conrad's talk.
3. **DARK ENERGY:**
   The accelerated cosmic expansion (together with the flat Universe implied by the Cosmic Microwave Background) requires a smooth yet unknown field with negative equation of state, making up ~70% of the energy density.

The next 5 to 10 years are poised to bring major observational breakthroughs in each of those topics!

The cosmological concordance model

Imperial College London

The ΛCDM cosmological concordance model is built on three pillars:

1. **INFLATION:**
   A burst of exponential expansion in the first ~10^{-32} s after the Big Bang, probably powered by a yet unknown scalar field.
2. **DARK MATTER:**
   The growth of structure in the Universe and the observed gravitational effects require a massive, neutral, non-baryonic yet unknown particle making up ~25% of the energy density - See Jan Conrad's talk.
3. **DARK ENERGY:**
   The accelerated cosmic expansion (together with the flat Universe implied by the Cosmic Microwave Background) requires a smooth yet unknown field with negative equation of state, making up ~70% of the energy density.

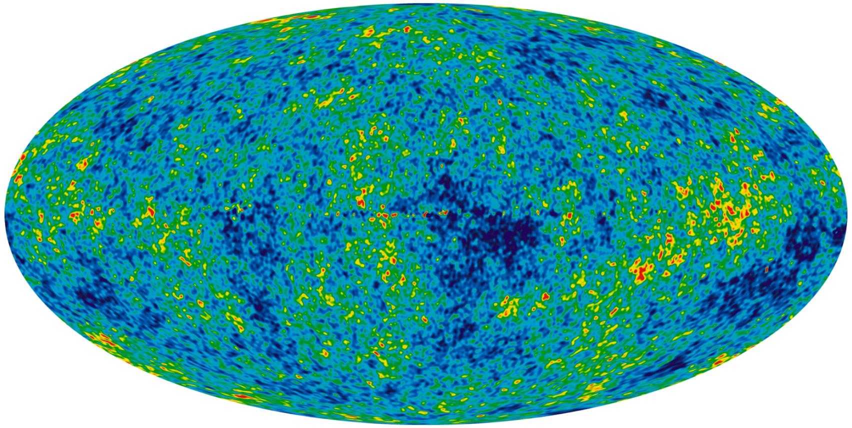
The next 5 to 10 years are poised to bring major observational breakthroughs in each of those topics!

Fundamental assumptions

- **Isotropy and homogeneity:** on sufficiently large scales, the average properties of the Universe do not depend on the direction and observer's location
  Violated if: Anisotropic Universe, cosmic bubbles, non-trivial topology (circles in the sky)
- **(Almost) Gaussianity of CMB fluctuations:** The initial condition for cosmic microwave background temperature fluctuations are close to a Gaussian random field (single field inflation)
  Violated if: multi-field inflation, curvaton, hybrid inflation, DBI inflation, ...
- **Adiabaticity:** The initial conditions were adiabatic (all matter species shared the same fluctuations up to a constant - single field inflation)
  Violated if: multiple scalar fields inflation (isocurvature modes)
- **Anomalies:** "Let's point our telescope there and see what we find!"

Cosmic Microwave Background analysis

Imperial College London



WMAP7 internal linear combination map

The observed anisotropies are a superposition of:

1. Initial conditions (inflation/early Universe physics)
2. Temperature/potential fluctuations at decoupling
3. Line-of-sight effects (ISW, SZ, lensing)

Temperature fluctuations:

$$\frac{\delta T}{T}(\vec{n}) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m}(\vec{n})$$

2-point correlation function

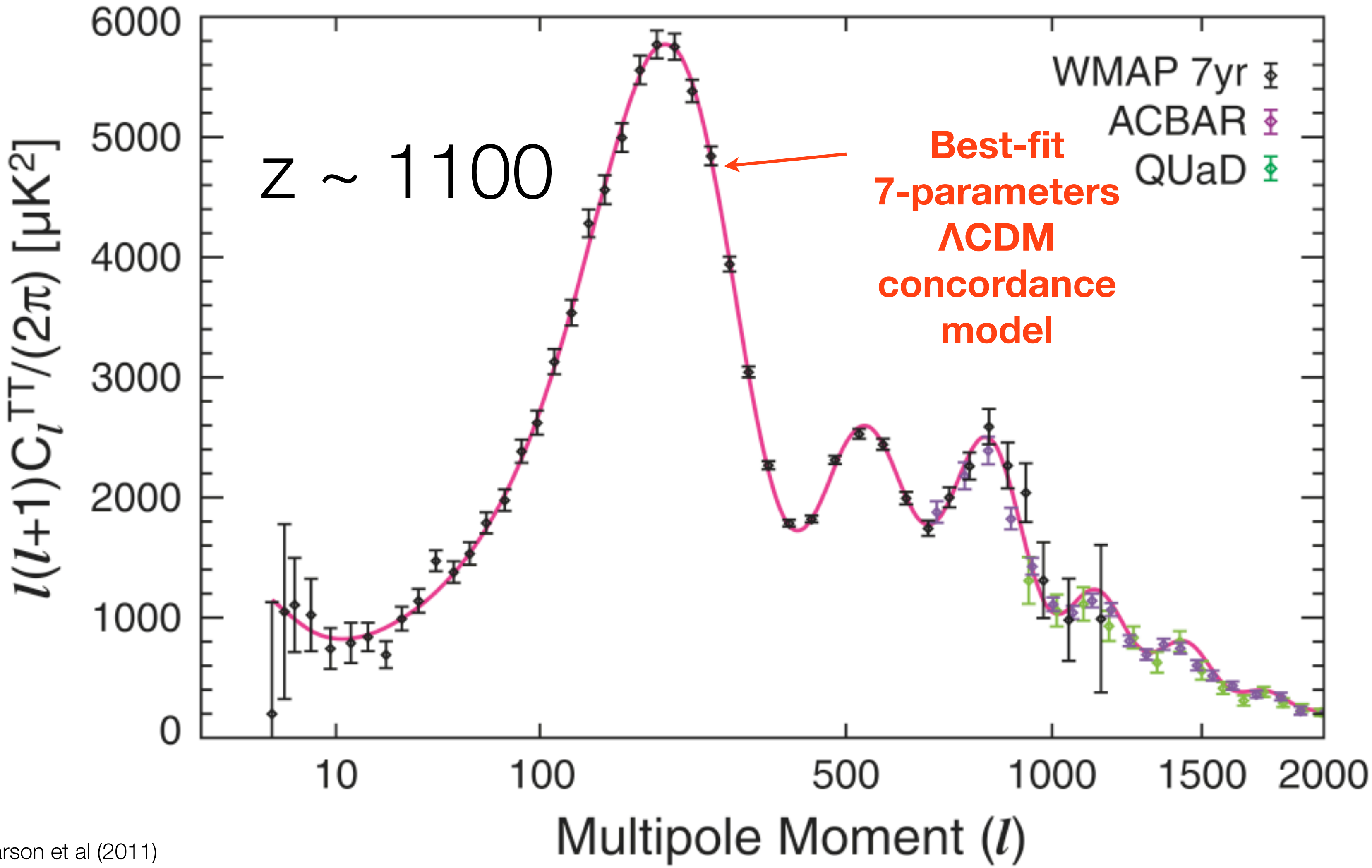
$$\xi(\theta) = \langle \frac{\delta T}{T}(\vec{n}) \frac{\delta T}{T}(\vec{n}') \rangle$$
$$= \frac{1}{4\pi} \sum_{\ell} (2\ell + 1) C_\ell P_\ell(\vec{n} \cdot \vec{n}')$$

Angular power spectrum (assumes isotropy)

$$C_\ell = \langle |a_{\ell m}|^2 \rangle$$

The power spectrum contains the full statistical information
IF fluctuations are Gaussian

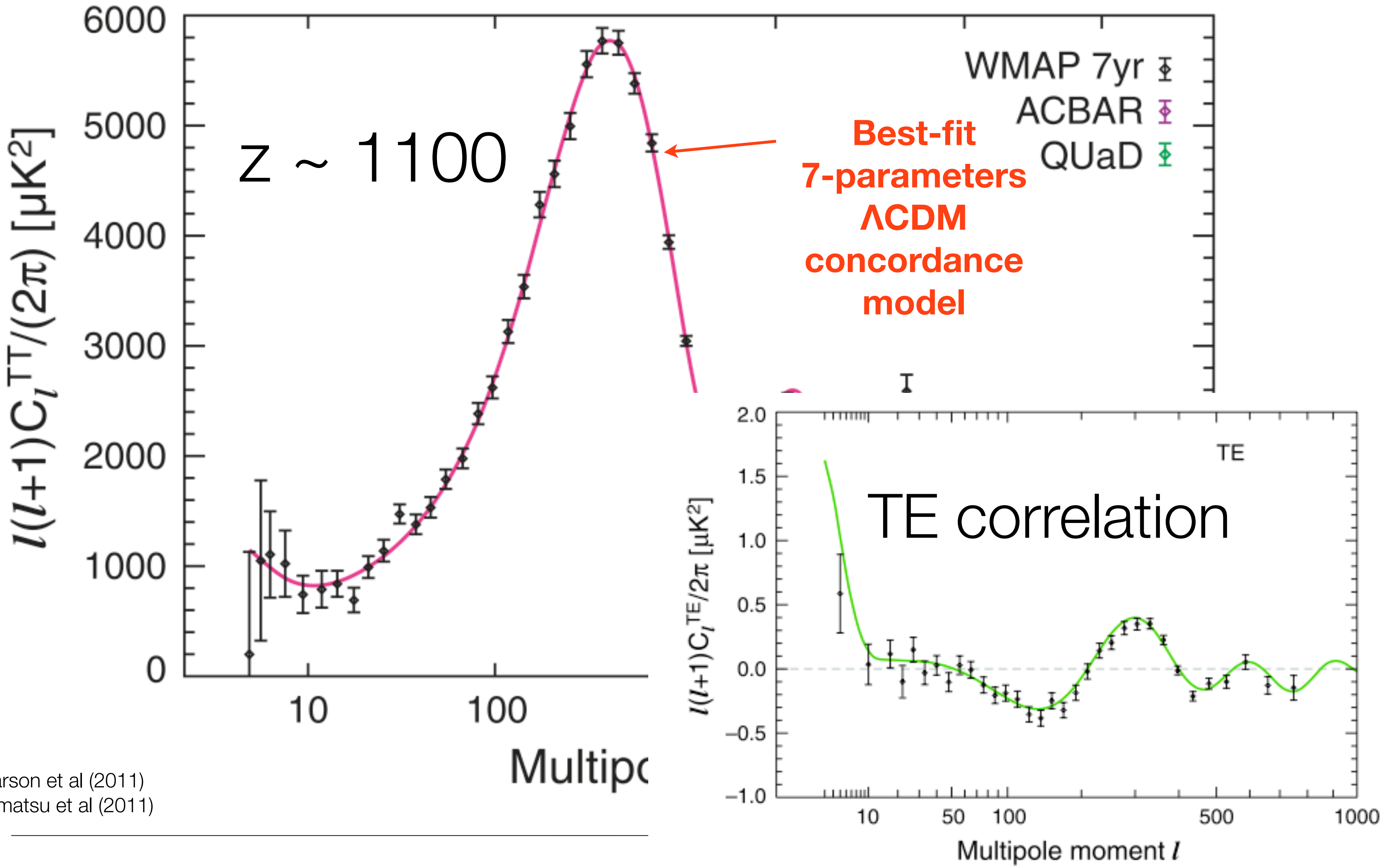
WMAP 7 temperature power spectrum



Larson et al (2011)
Komatsu et al (2011)

WMAP 7 temperature power spectrum

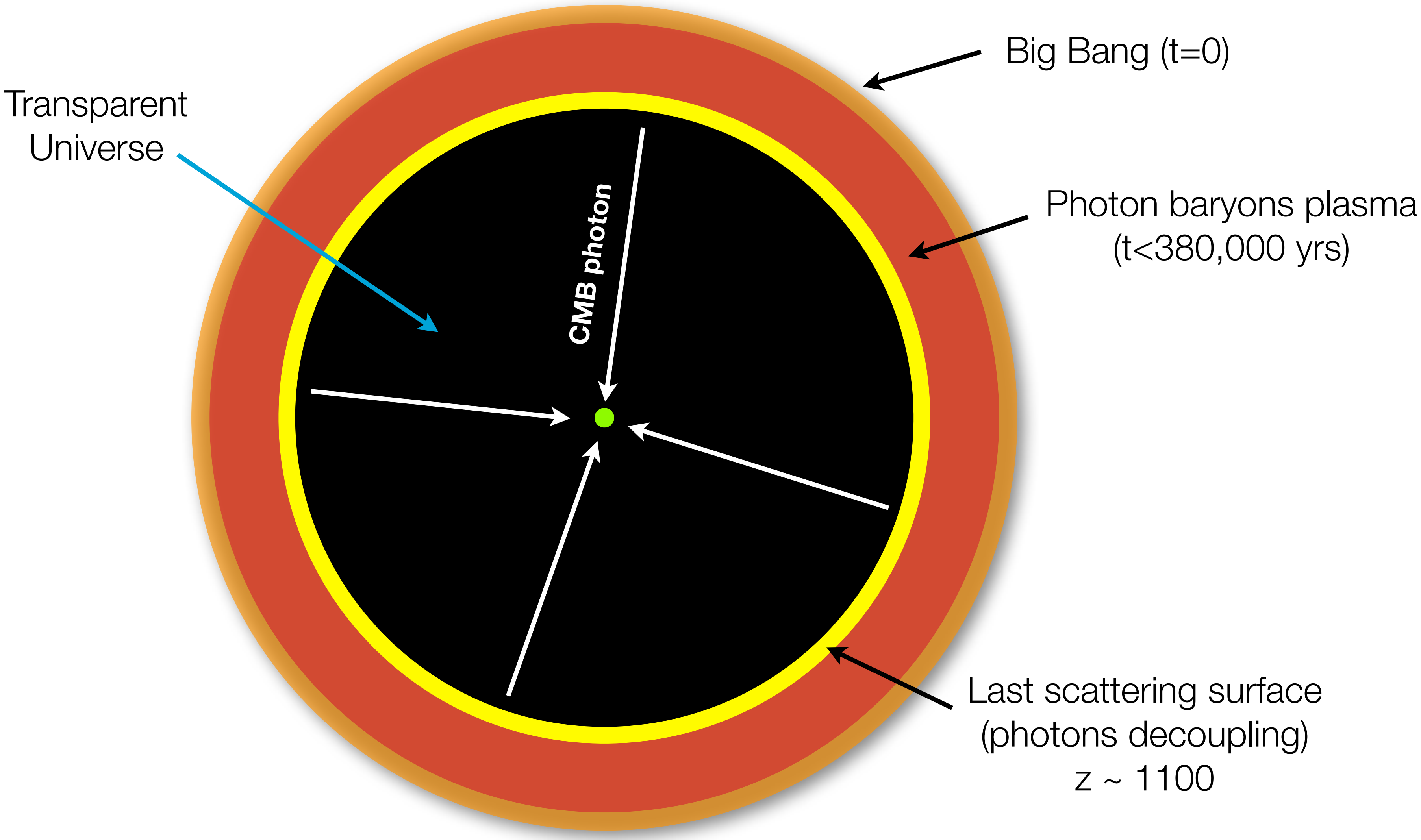
Imperial College London



Larson et al (2011)
Komatsu et al (2011)

Origin of the CMB

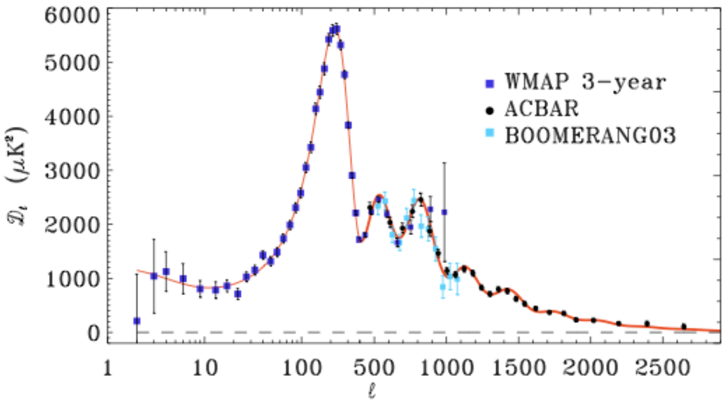
Imperial College London

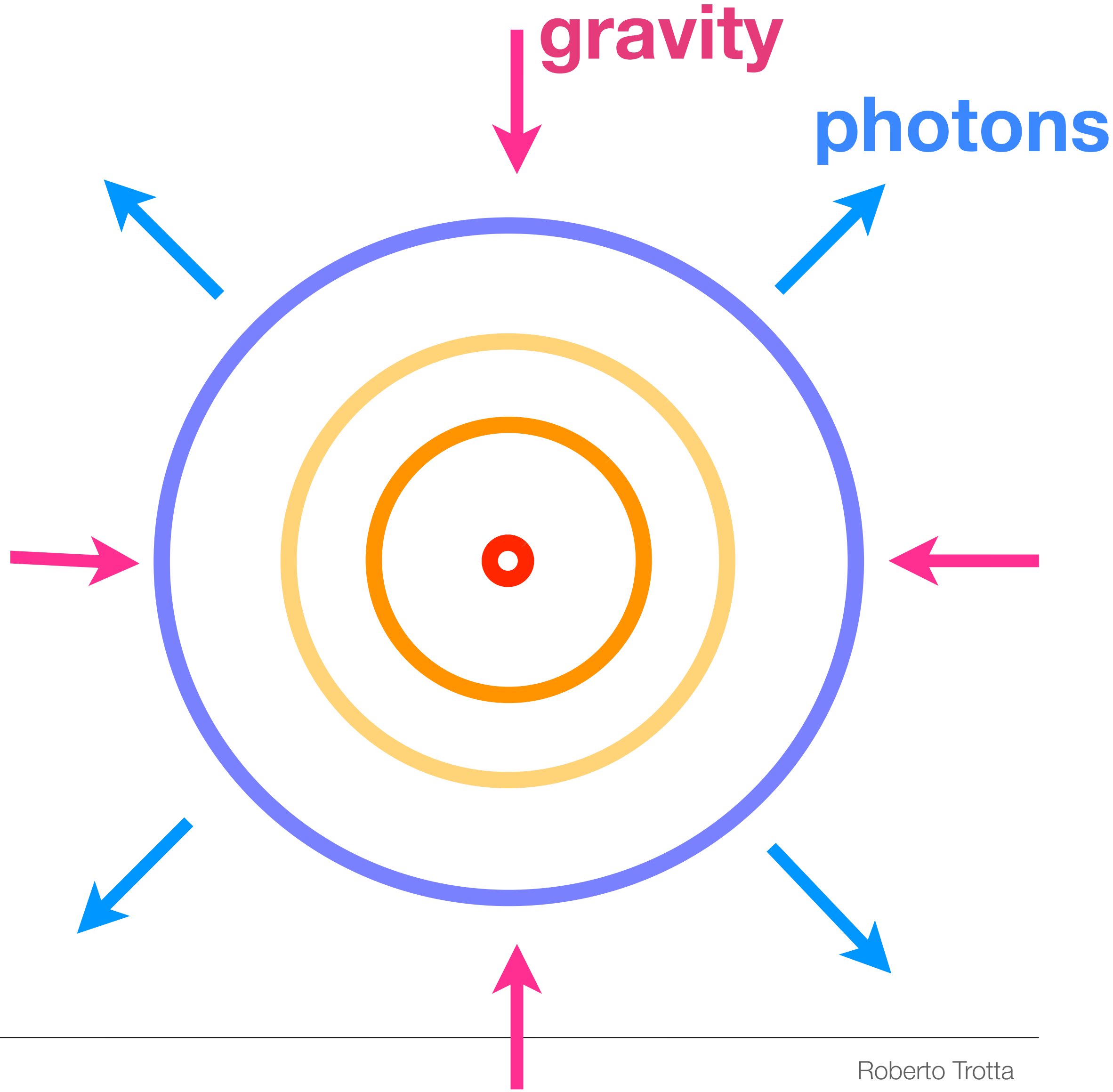


Roberto Trotta

Cosmic sound

Imperial College London

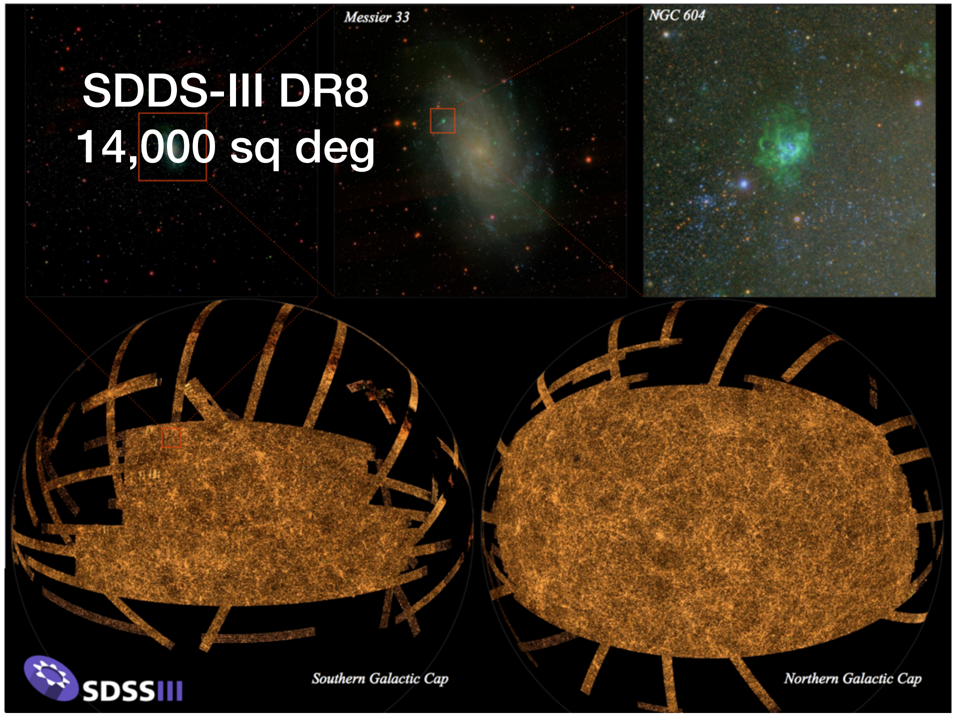




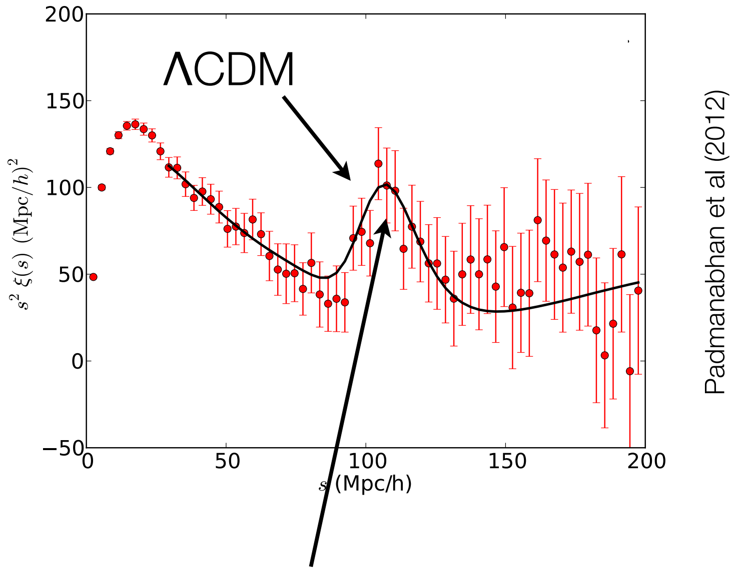
Roberto Trotta

BAO: correlation between galaxies' position

Primordial sound waves introduce extra correlation between galaxies on scales ~ 150 Mpc:
this corresponds to (on average) 1 extra galaxy at this preferential separation



Baryonic acoustic oscillations (z~0.35)



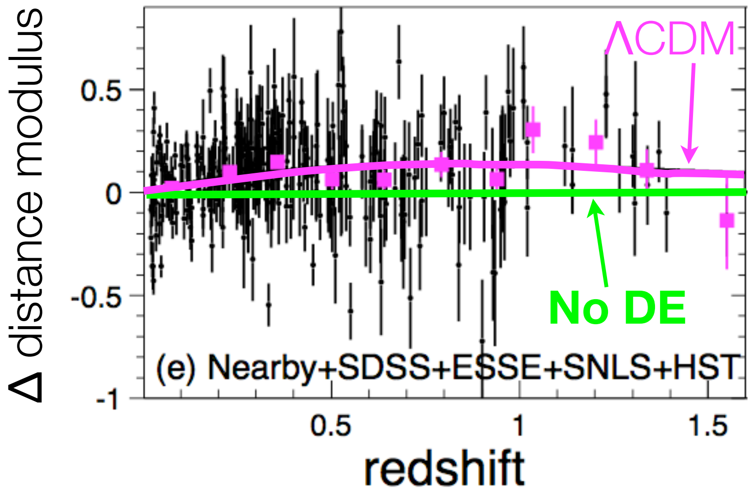
Baryonic **A**coustic **O**scillations from
~50,000 LRGs
1.9% distance accuracy to z=0.35
~4 sigma significance after reconstruction

Low redshift cosmological probes

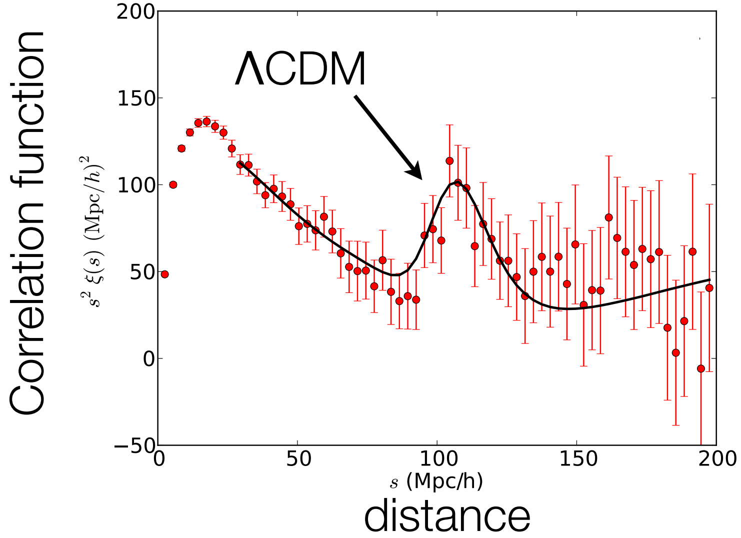
Imperial College London

Supernovae type Ia (z < 1.5)

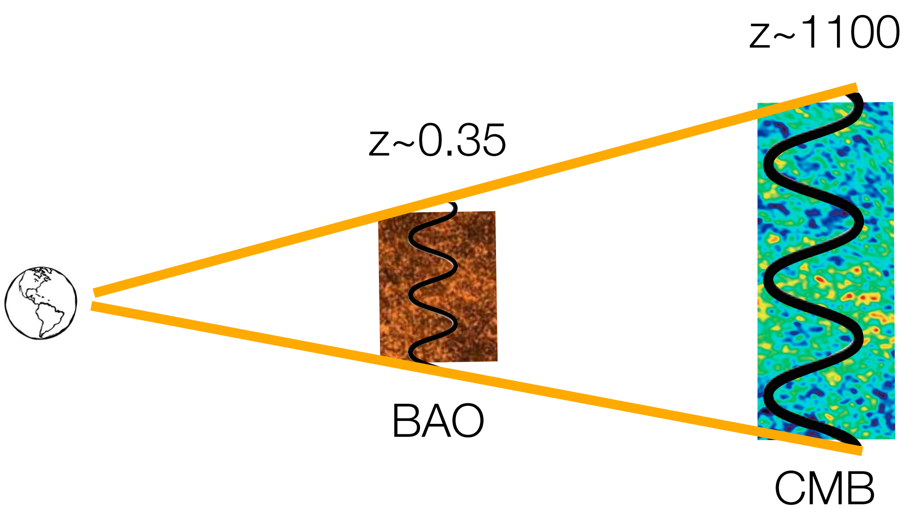
Kessler et al (SDSS collaboration) (2010)

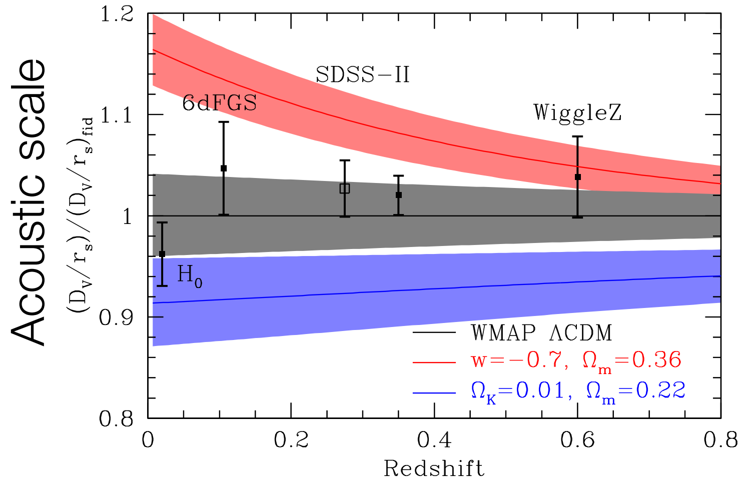


Baryonic acoustic oscillations (z~0.35)



Padmanabhan et al (2012)





Mehta et al (2012)

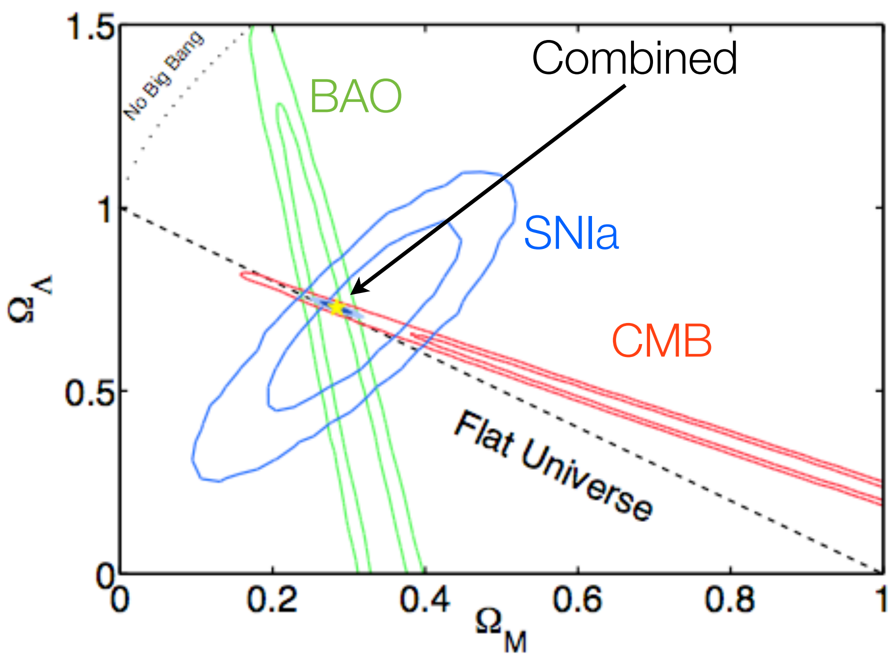
Roberto Trotta

Putting it all together...

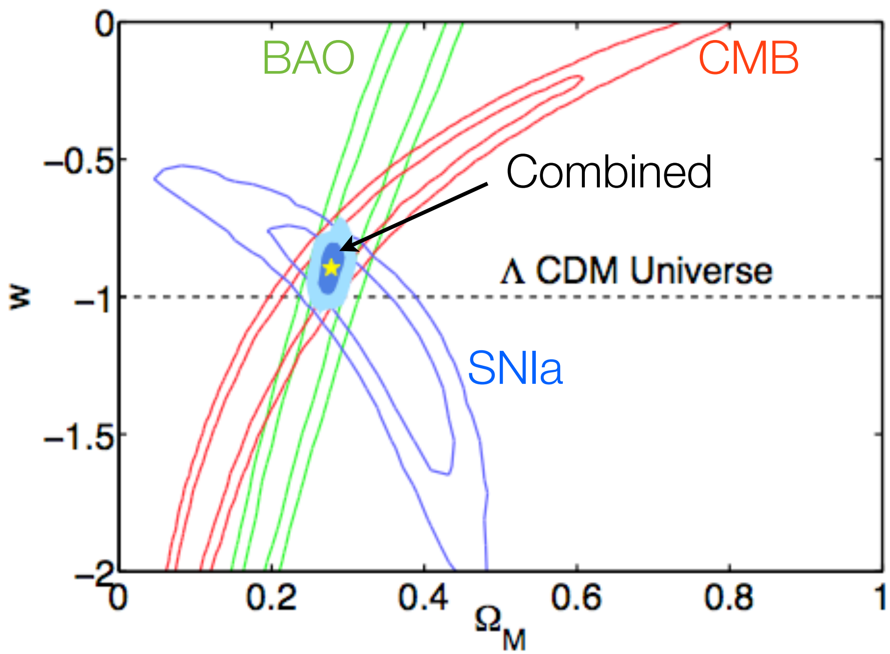
Imperial College London

Combined constraints on total matter ($\Omega_M=\Omega_B+\Omega_{CDM}$) and dark energy ($\Omega_\Lambda$) content (dark energy equation of state parameter w = pressure/energy density):

Assuming Λ (w=-1)



Assuming flatness ($\Omega_\Lambda+\Omega_M=1$)

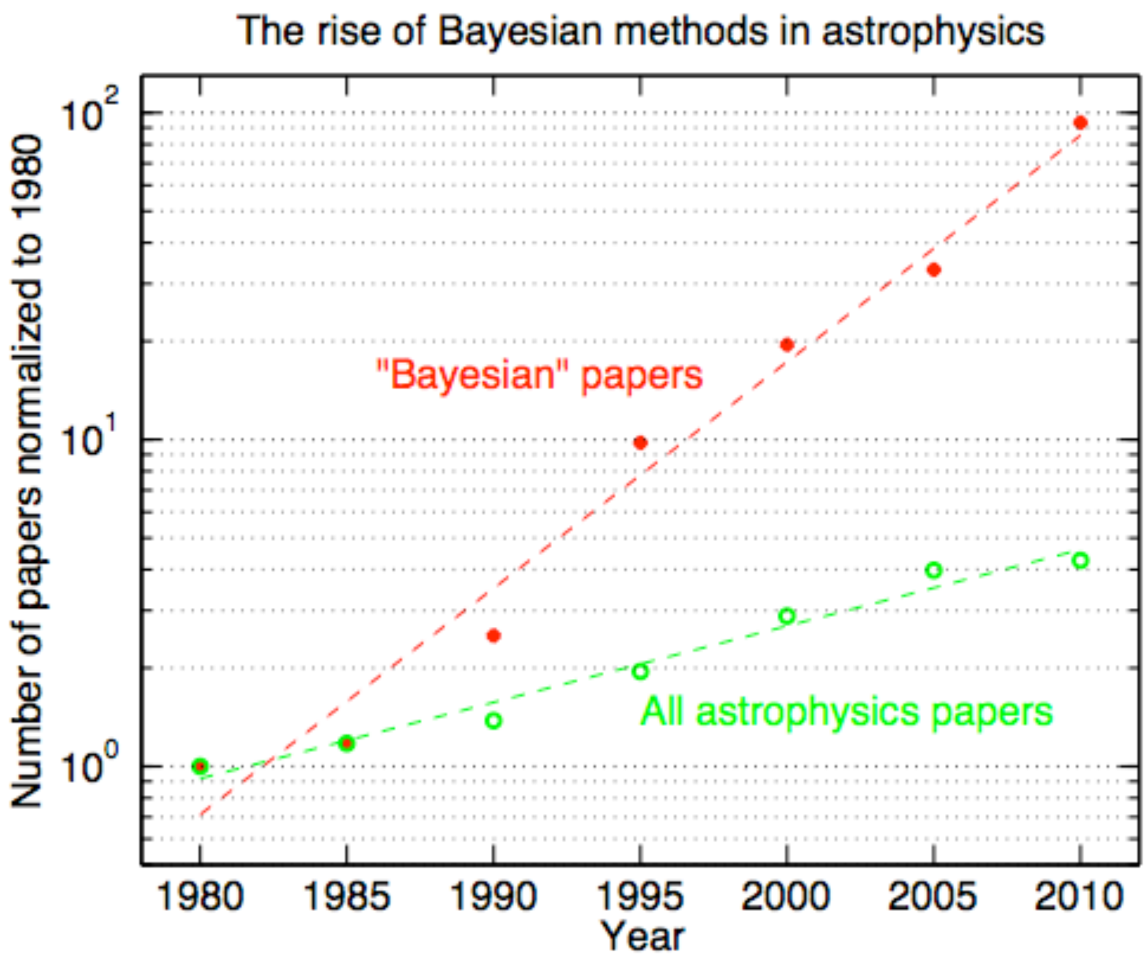


March, RT et al (2012)

Roberto Trotta

Bayes in the sky

Imperial College London



Non-Gaussianity

Non-Gaussianity from inflation

Imperial College London

- Single field inflation predicts that *primary* CMB fluctuations are almost perfectly Gaussian distributed
- Departures from Gaussianity would be a smoking gun signature of non-standard physics at the time of inflation (e.g., multiple fields)

primordial curvature perturbation

Gaussian field (linear theory)

Non-Gaussian parameter

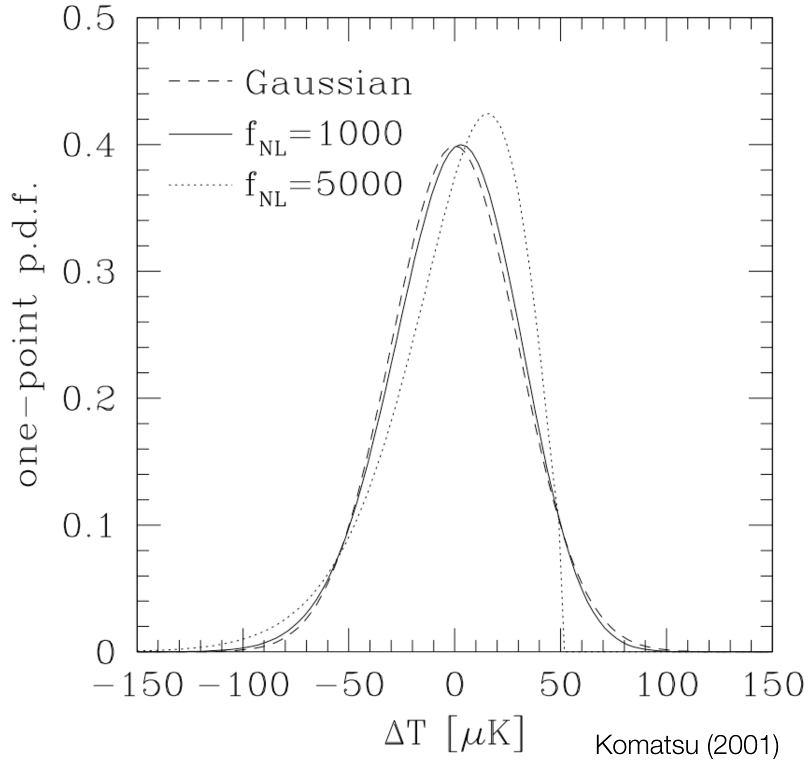
2nd order perturbation theory

$$\Phi = \Phi_L + f_{NL}\Phi_L^2 + \cdots$$

Levels of non-Gaussianity

1. Primordial: $f_{NL} \sim 10^{-2}$ to a few
2. Line-of-sight (ISW-lensing): $f_{NL} \sim 7\text{-}9$
3. Extragalactic point sources: $f_{NL} \sim 2$

Impact of f_{NL} non-Gaussianity



Komatsu (2001)

Roberto Trotta

Searches for non-Gaussianity

Imperial College London

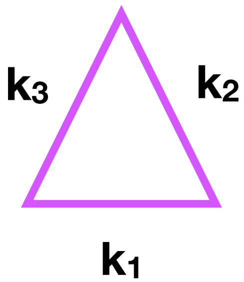
- Departures from Gaussianity are looked for using several techniques/estimators: bispectrum, wavelets, skweness, kurtosis, genus statistics, Minkowski functionals, needlets ...
- The first order non-Gaussian signature is in the three-point correlation function:

$$\langle \Phi(\mathbf{k_1}), \Phi(\mathbf{k_2}), \Phi(\mathbf{k_3}) \rangle = \overbrace{\boxed{F(k_1, k_2, k_3)}}^{\text{Bispectrum}} (2\pi)^3 \delta(\mathbf{k_1} - \mathbf{k_2} - \mathbf{k_3})$$

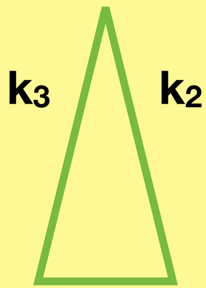
This assumes **isotropy** (rotation invariant) and **homogeneity** (translation invariant)

$$F(\alpha k_1, \alpha k_2, \alpha k_3) = f(\alpha) F(k_1, k_2, k_3)$$

Equilateral



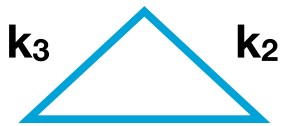
Squeezed



$k_1 \ll k_2, k_3$

~ 10^{-2} for single field inflation

Flattened



$k_1 \approx k_2 + k_3$

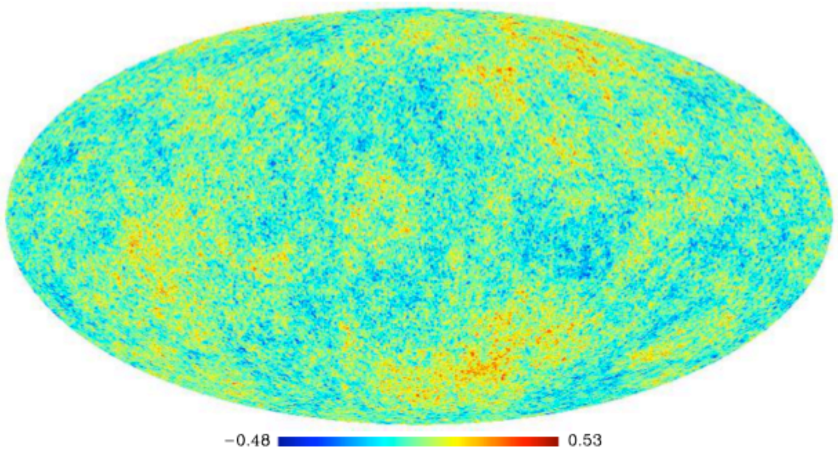
(Highly) Non-Gaussian CMB maps

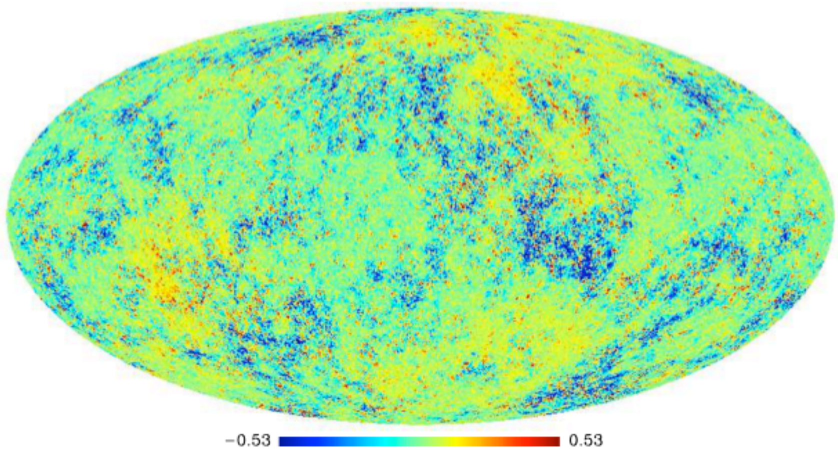
Imperial College London

$$\frac{\Delta T}{T} = g\Phi$$

Gaussian field
$f_{NL} = 0$

Non-Gaussian field
$f_{NL} = 5000$





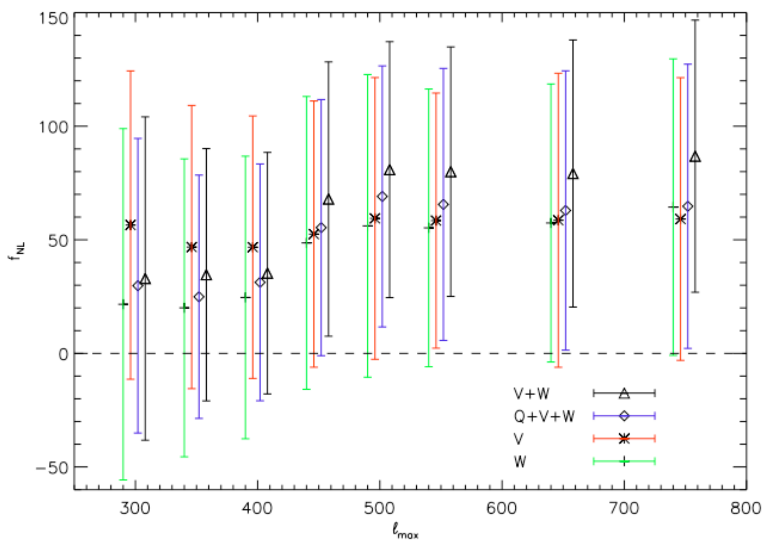
Bernui et al (2012)

Constraints on non-Gaussianity

Imperial College London

- In practice, searches for non-Gaussianity look for evidence of $f_{NL} \neq 0$ (or equivalent) for different families of triangles
- This relies on developing unbiased maximum-likelihood estimators for f_{NL} (and its uncertainty, often based on Gaussian approximation)
- A fully Bayesian treatment is presently lacking (computationally intractable)

Yadav & Wandelt (2008) claimed
f_{NL} = 0 rejected at 99.5% CL



Today (WMAP7), **null results** for various triangle configurations

$$
\begin{aligned}
f_{NL}^{\rm local} &= 32 \pm 21 \\
f_{NL}^{\rm equil} &= 26 \pm 140 \\
f_{NL}^{\rm orthog} &= -202 \pm 104
\end{aligned}
$$

Next year, **Planck** will improve limits by a factor ~ 5

Roberto Trotta

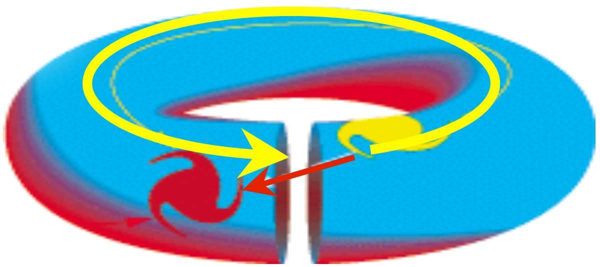
Non-trivial topology

The shape of the Universe

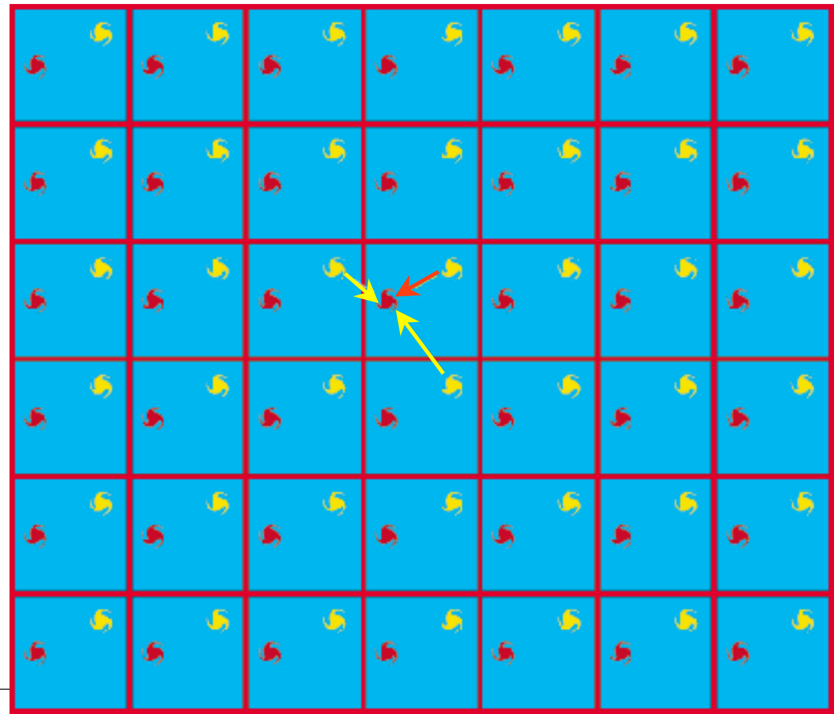
Imperial College London

- If space has a non-trivial topology, light from distant objects can reach us along multiple paths
- This means that we could in principle see "copies" in the sky of distant objects (with different orientation and redshift, in principle)

A 2D Universe with a torus shape

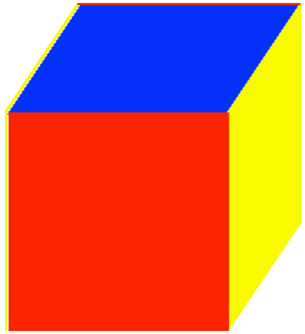


As seen in the fundamental cell

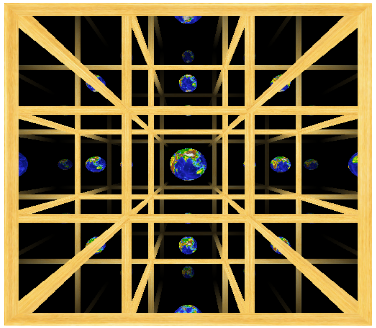


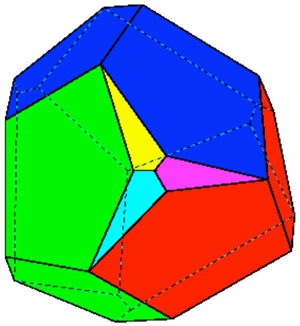
Roberto Trotta

Fundamental cell | **As seen by an observer in 3D space**

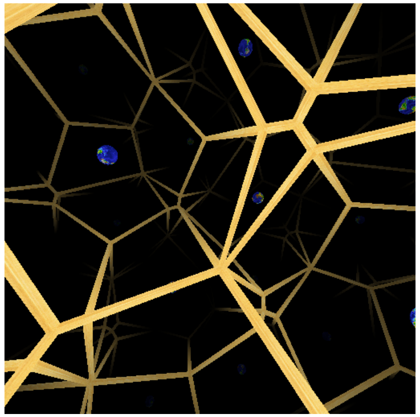


Flat space
(3-torus)

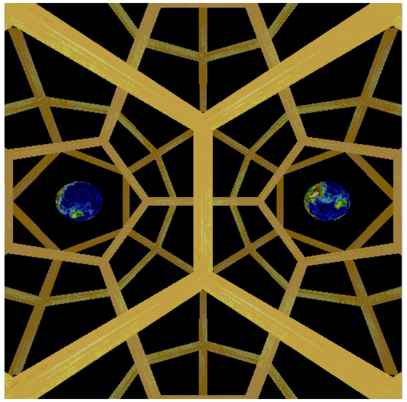




Hyperbolic space



Spherical space
(dodecahedron)

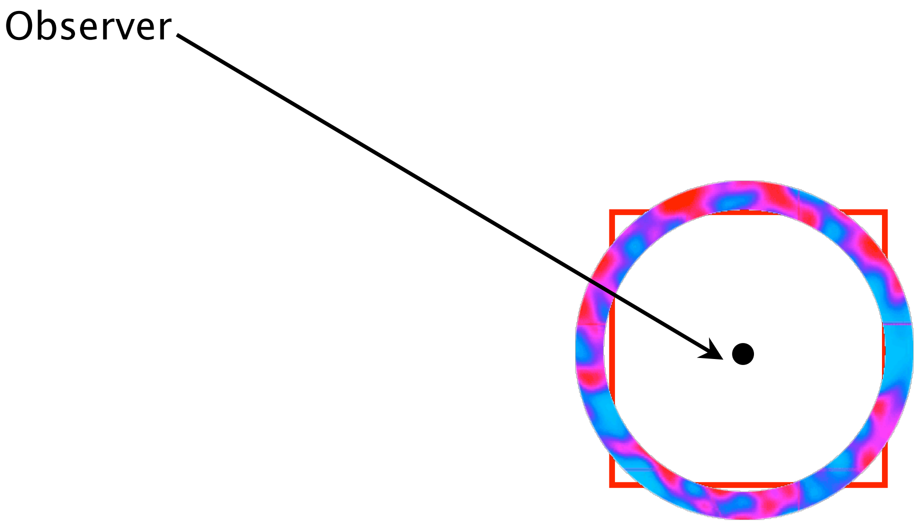


Circles in the sky

Imperial College London

- As the most distant observable in the Universe, the CMB last scattering surface ($z \sim 1100$) provides a powerful test of cosmic topology (Cornish et al, 1998).
- Signature of non-trivial topology are **matching circles in the sky.** Also, it might explain **lack of power at large angular scales** (Luminet et al, 2003)

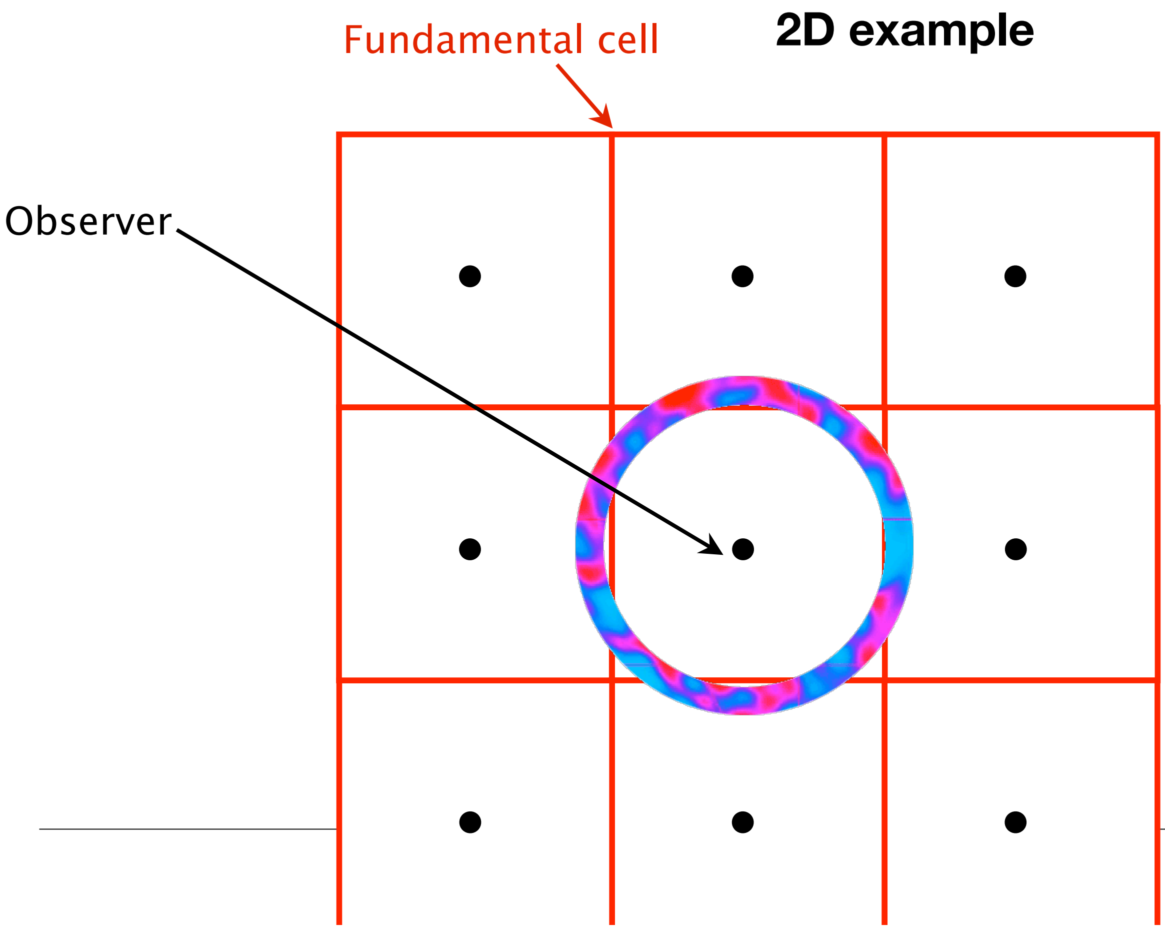
2D example



Circles in the sky

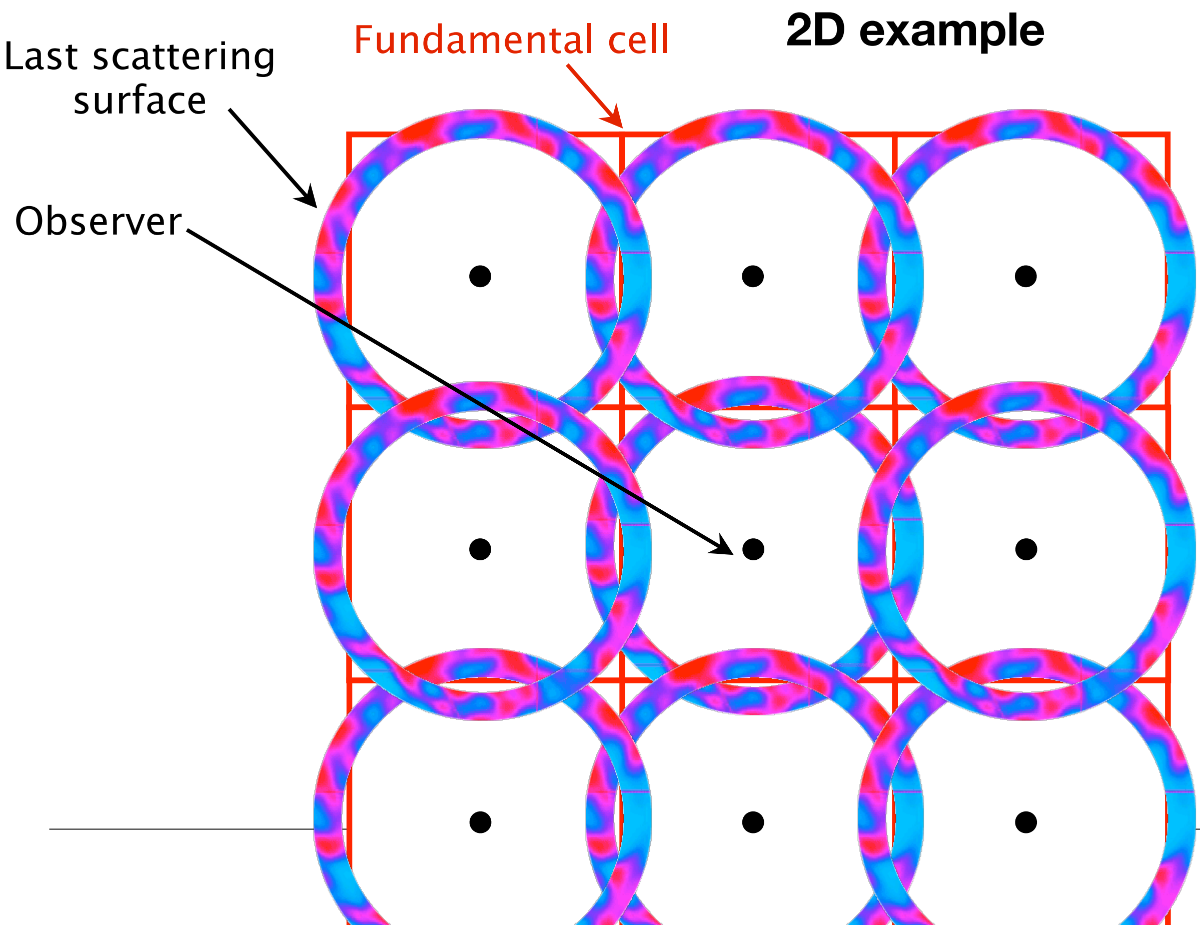
Imperial College London

- As the most distant observable in the Universe, the CMB last scattering surface (z~1100) provides a powerful test of cosmic topology (Cornish et al, 1998).
- Signature of non-trivial topology are **matching circles in the sky.** Also, it might explain **lack of power at large angular scales** (Luminet et al, 2003)



Circles in the sky

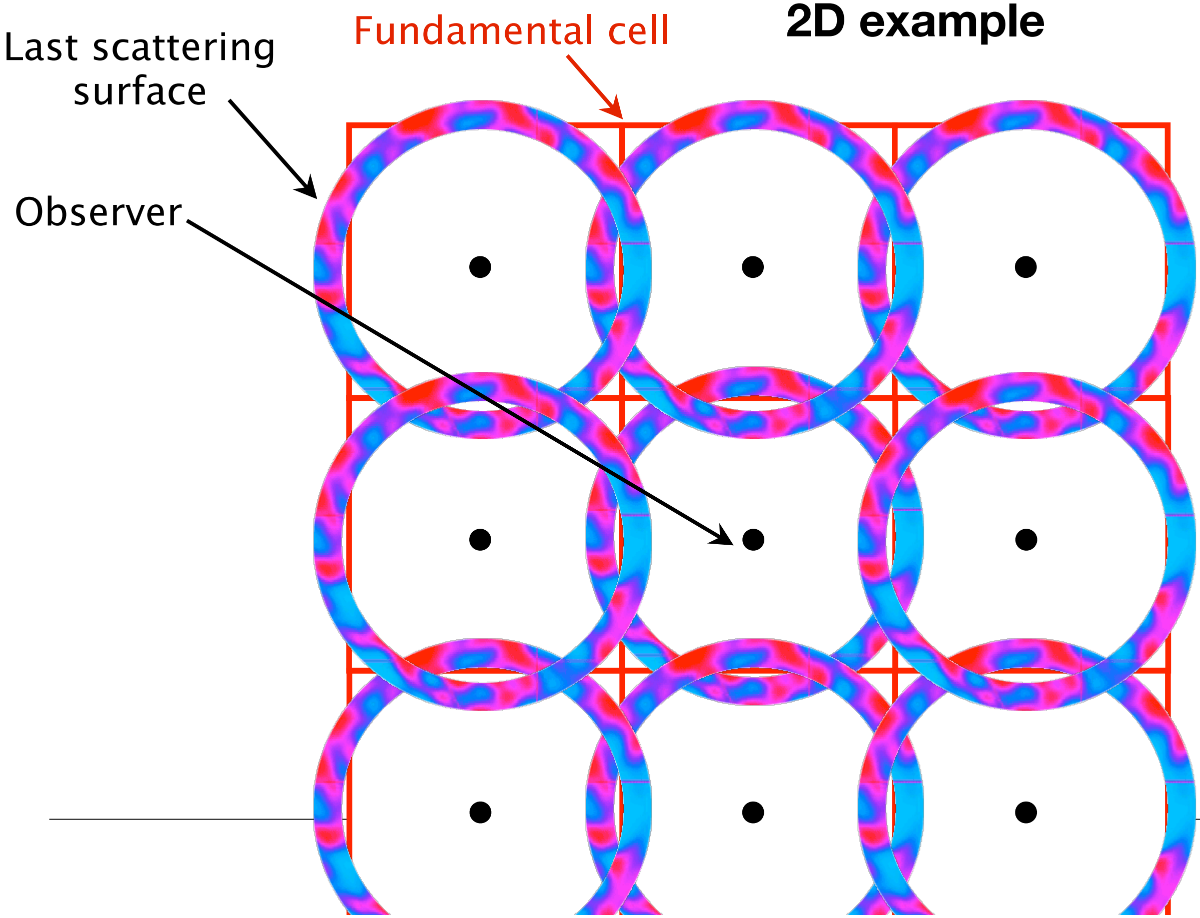
- As the most distant observable in the Universe, the CMB last scattering surface (z~1100) provides a powerful test of cosmic topology (Cornish et al, 1998).
- Signature of non-trivial topology are **matching circles in the sky.** Also, it might explain **lack of power at large angular scales** (Luminet et al, 2003)



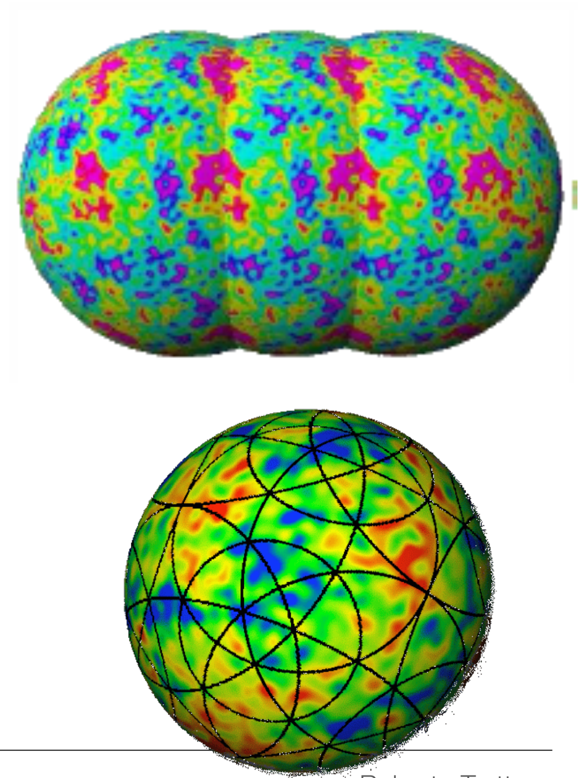
Circles in the sky

- As the most distant observable in the Universe, the CMB last scattering surface (z~1100) provides a powerful test of cosmic topology (Cornish et al, 1998).
- Signature of non-trivial topology are **matching circles in the sky.** Also, it might explain **lack of power at large angular scales** (Luminet et al, 2003)

2D example

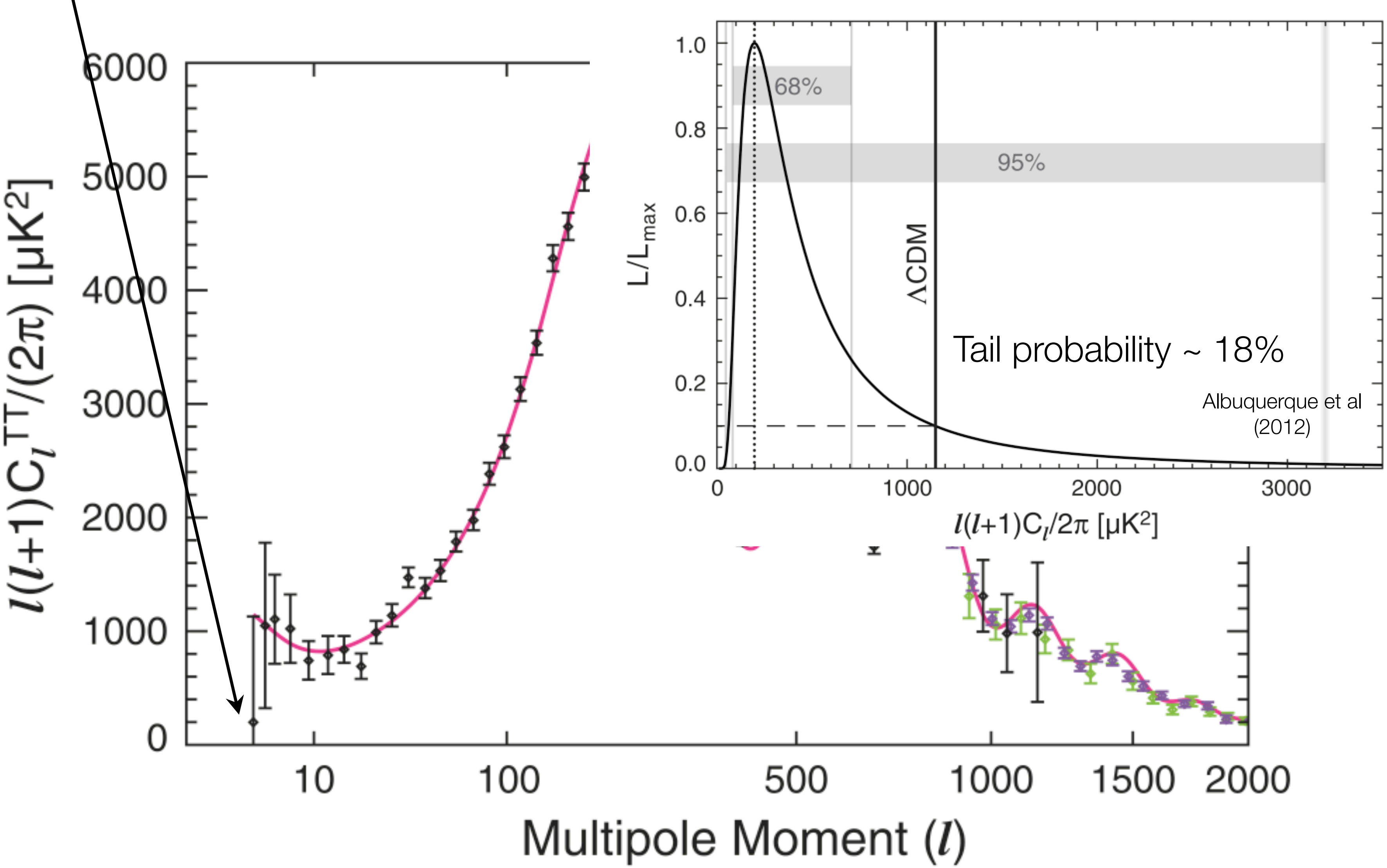


3D: matching circles



Small quadrupole anomaly: fundamental cell too small to support long-wavelenght modes ?

Posterior pdf for quadrupole power



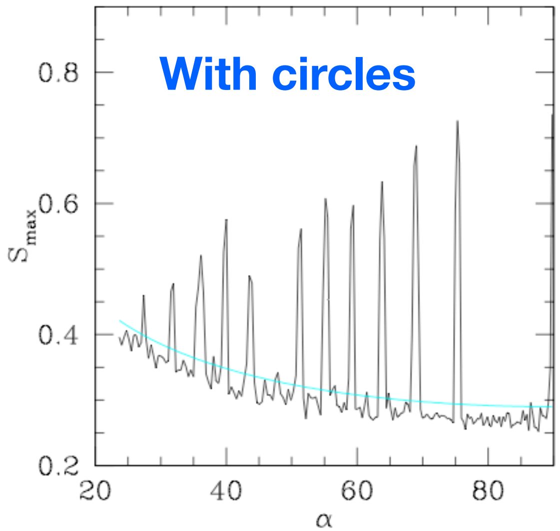
Searching for matching circles

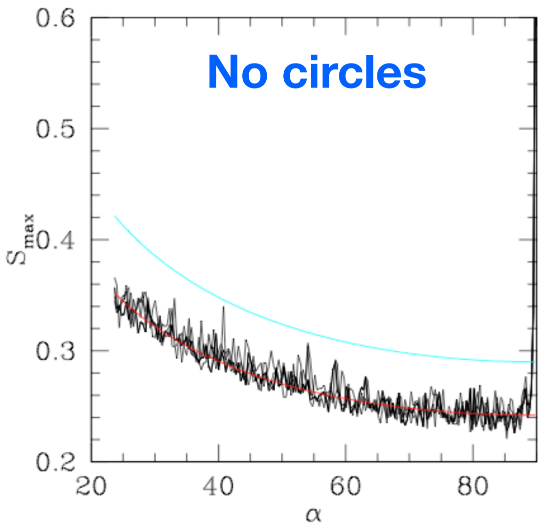
- Cornish et al (1998) introduced a "circle comparison statistics" between circles centered at points A,B on the sky, with angular radius α and relative phase ϕ_0:

$$S(\alpha, \phi_0) = \frac{\langle 2T_A(\phi)T_B(\phi+\phi_0)\rangle}{\langle T_A(\phi)^2 + T_B(\phi)^2\rangle}$$

< . > denotes averaging over ϕ along the circle circumference.

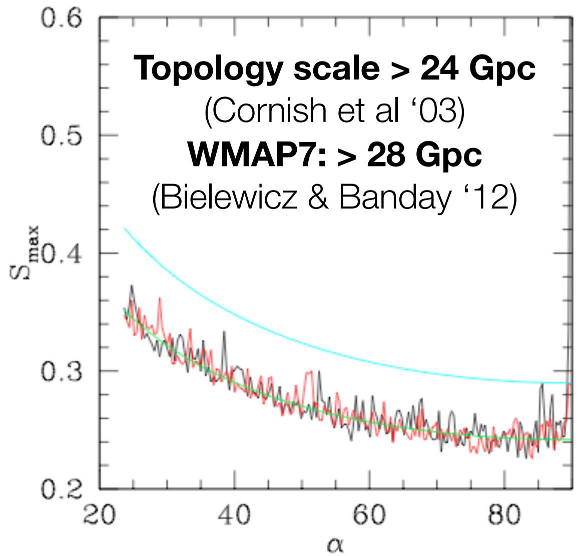
Test statistics for simulated maps as a function of angle α (back-to-back circles)





Cornish et al (2003)

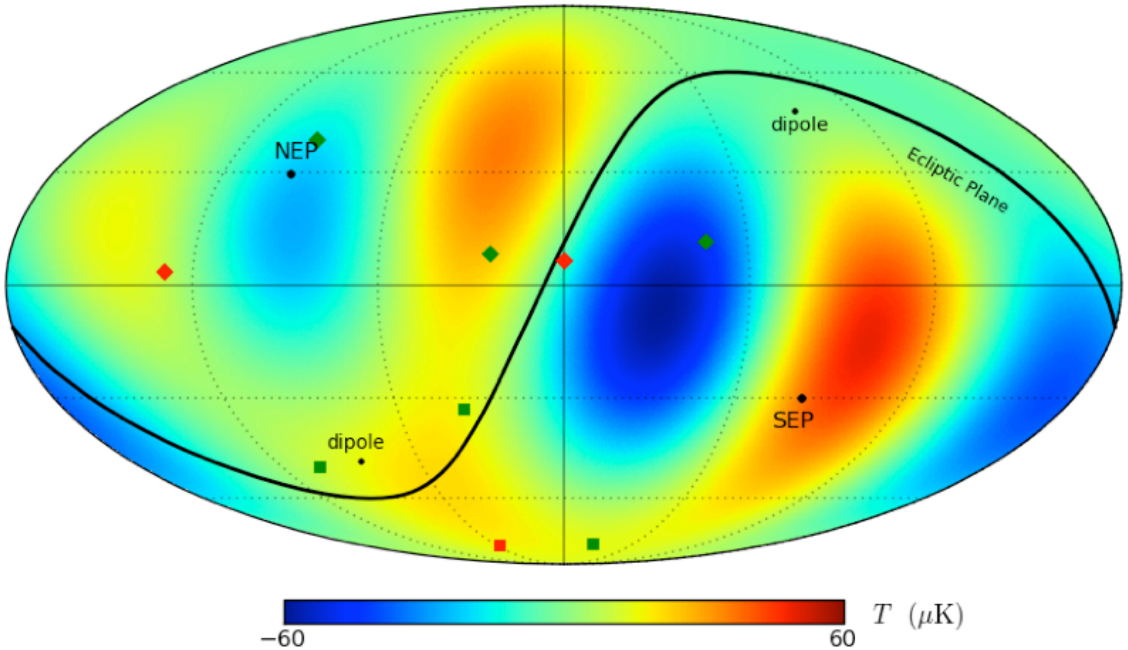
WMAP1 data



Large scale anomalies?

Imperial College London

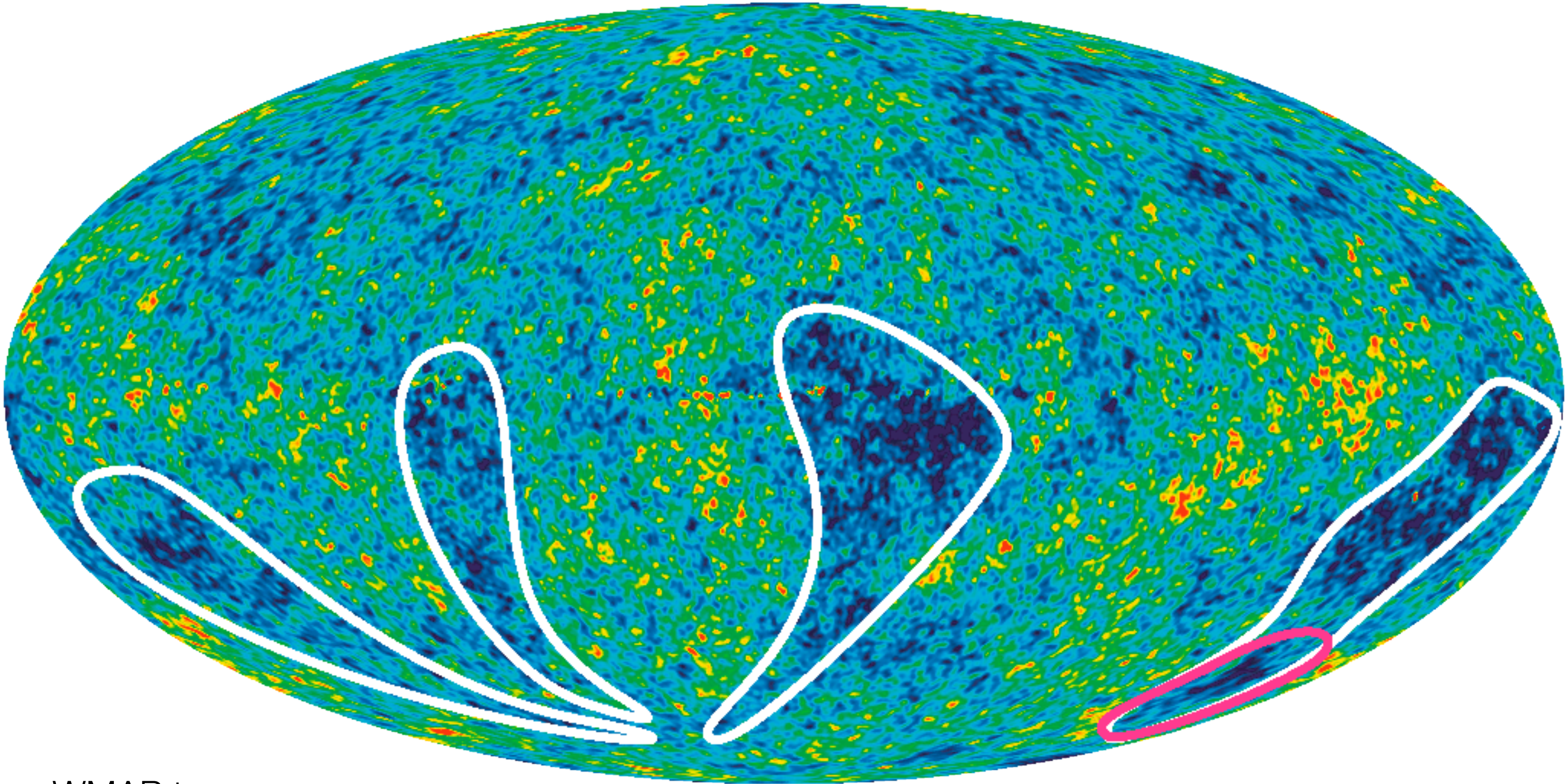
- Quadrupole (ell = 2) is low
- Four area vectors of Quadrupole and Octupole are mutually close: p-value = 0.004
- Quadrupole and Octupole aligned with ecliptic: p-value = 0.041
- Normals to area vector planes aligned with dipole: p-value = 0.003
- Hot/Cold spots divided by the ecliptic: p-value ~ 0.05
- Two-point correlation function vanishes above 60°: p-value ~ 0.0002



Copi et al (2010)

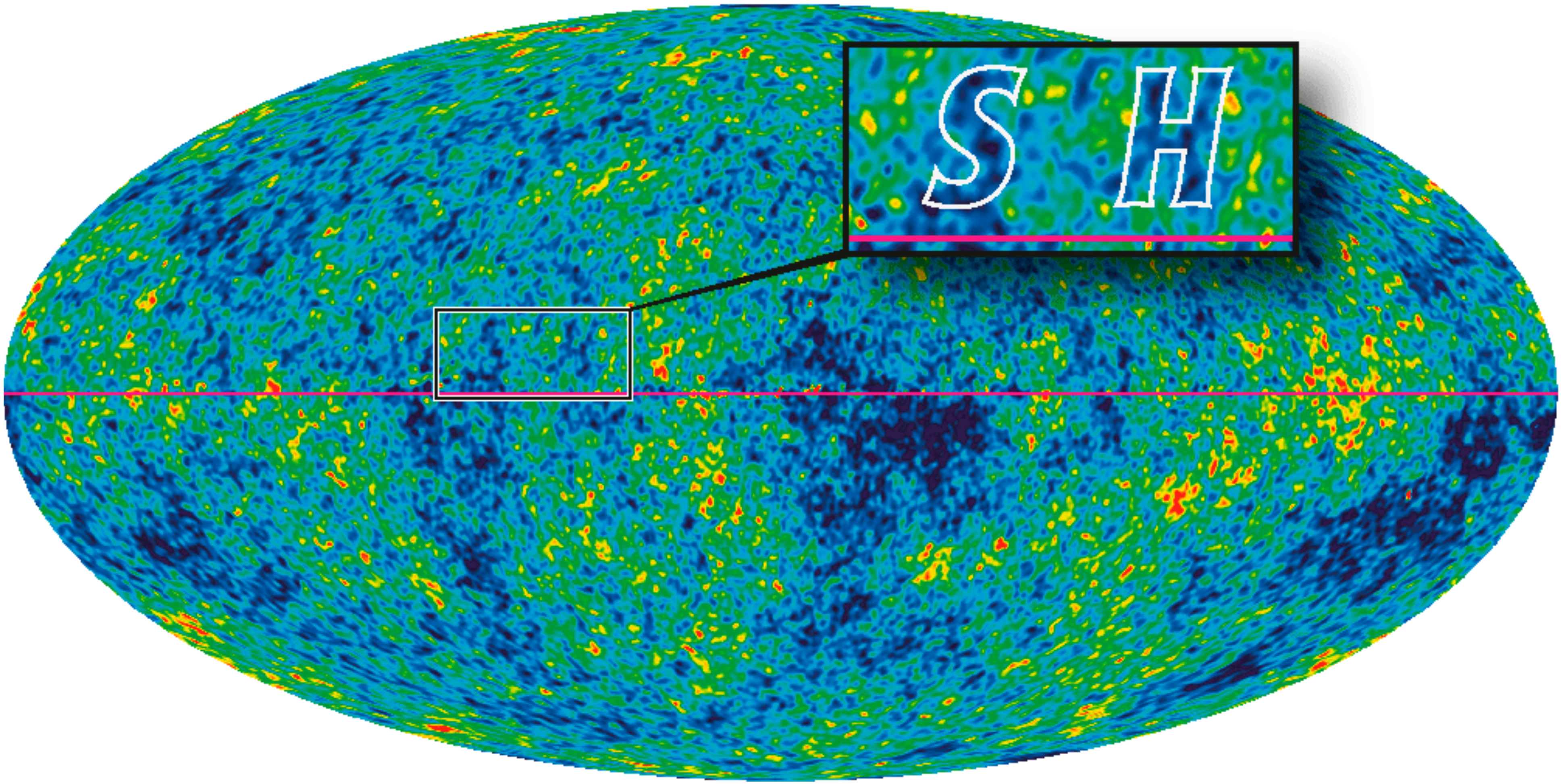
Roberto Trotta

Anomalous cold regions in the southern hemisphere



WMAP team

Look elsewhere effect: the "SH" initials of Stephen Hawking are shown in the ILC sky map.



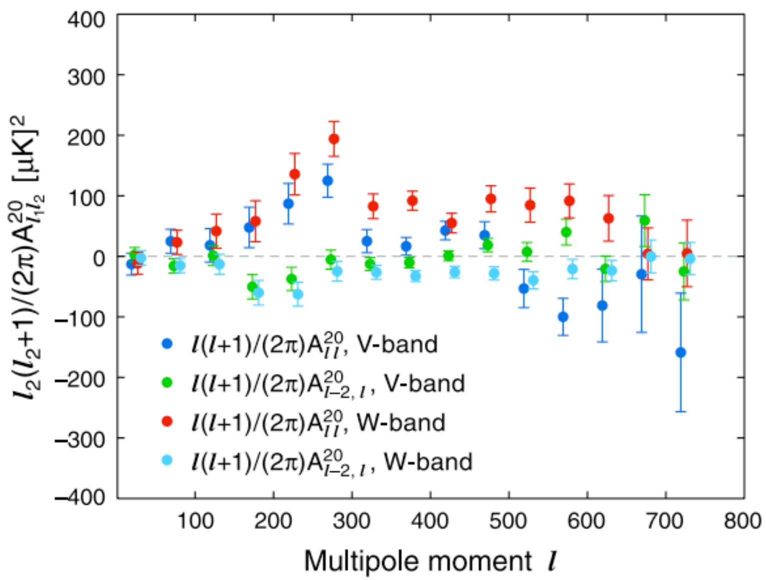
WMAP team

WMAP7 conclusions on anomalies

Imperial College London

- Cold spots statistically consistent with random
- Amplitude of the quadrupole within the expected 95% confidence range
- No significant anomaly with a lack of large angular scale CMB power
- The quadrupole and octupole components “remarkably” aligned: “It may be due, in part, to chance alignments between the primary and secondary anisotropy”
- “While this alignment appears to be remarkable, there was no model that predicted it, nor has there been a model that provides a compelling retrodiction”
- Hemispherical or dipole power asymmetry: evidence not statistically significant.

- Analysis of bipolar power spectra confirms a strong quadrupolar power asymmetry, but the effect is probably not cosmological (residual beam asymmetry?)



Bennett et al (2011)

Generic departures from LCDM

The “null hypothesis”

Imperial College London

- The Concordance Model with “minimal” assumptions (single field inflation, FRW metric, General Relativity, dark matter, cosmological constant) provides a number of “generic predictions” that act as a starting point for searches for deviations:

Quantity	Prediction	Status
Initial conditions	Adiabatic. No isocurvature modes	**OK (within ~ 10%)**
Spectral distribution of fluctuations	Featureless power law	**OK (within ~ 5%)**
Gaussianity	$f_{NL} \sim 0$	**OK (f_{NL} < ~ 40)**
Dark energy	Cosmological constant. $w(z) = -1$	**OK (within ~ 10%)**
Curvature of spatial sections	$\Omega_K = 0$	**OK (within ~ 0.1%)**
Primordial gravity waves	r = T/S unknown (probably small)	**Unknown (r < 0.2)**
Consistency relation D_L, D_A	$D_L/(D_A(1+z)^2) = 1$	**OK (within ~ 20%)**
Number of neutrino species	3	**2-sigma deviation**
Statistical isotropy	YES	**Unclear (anomalies?)**
Non-trivial topology	NO	**None fund (< 29 Gpc scale)**
Neutrino mass	m > 0.056 eV (normal hierarchy)	**Upper limits, m < 0.4 eV**

Principled Bayesian model selection

The 3 levels of inference

Imperial College London

LEVEL 1
I have selected a model M and prior P(θ|M)

Parameter inference
What are the favourie values of the parameters?
(assumes M is true)

$$P(\theta|d,M) = \frac{P(d|\theta,M)P(\theta|M)}{P(d|M)}$$

LEVEL 2
Actually, there are several possible models: M_0, M_1,...

Model comparison
What is the relative plausibility of M_0, M_1,... in light of the data?

$$\mathrm{odds} = \frac{\mathrm{P}(\mathrm{M}_0|\mathrm{d})}{\mathrm{P}(\mathrm{M}_1|\mathrm{d})}$$

LEVEL 3
None of the models is clearly the best

Model averaging
What is the inference on the parameters accounting for model uncertainty?

$$P(\theta|d) = \sum_i P(M_i|d)P(\theta|d,M_i)$$

Roberto Trotta

Examples of model comparison questions

Imperial College London

ASTROPARTICLE

Gravitational waves detection

Do cosmic rays correlate with AGNs?

Which SUSY model is ‘best’?

Is there evidence for DM modulation?

Is there a DM signal in gamma ray/ neutrino data?

COSMOLOGY

Is the Universe flat?

Does dark energy evolve?

Are there anomalies in the CMB?

Which inflationary model is ‘best’?

Is there evidence for modified gravity?

Are the initial conditions adiabatic?

Many scientific questions are of the model comparison type

ASTROPHYSICS

Exoplanets detection

Is there a line in this spectrum?

Is there a source in this image?

Level 2 inference: model comparison

Imperial College London

$$P(\theta|d, M) = \frac{P(d|\theta, M)P(\theta|M)}{P(d|M)}$$

Bayesian evidence or model likelihood

The evidence:

$$P(d|M) = \int_\Omega d\theta P(d|\theta, M)P(\theta|M)$$

Posterior probability for the model M:

$$P(M|d) = \frac{P(d|M)P(M)}{P(d)}$$

When comparing two models:

$$\frac{P(M_0|d)}{P(M_1|d)} = \frac{P(d|M_0)}{P(d|M_1)}\frac{P(M_0)}{P(M_1)}$$

The Bayes factor:

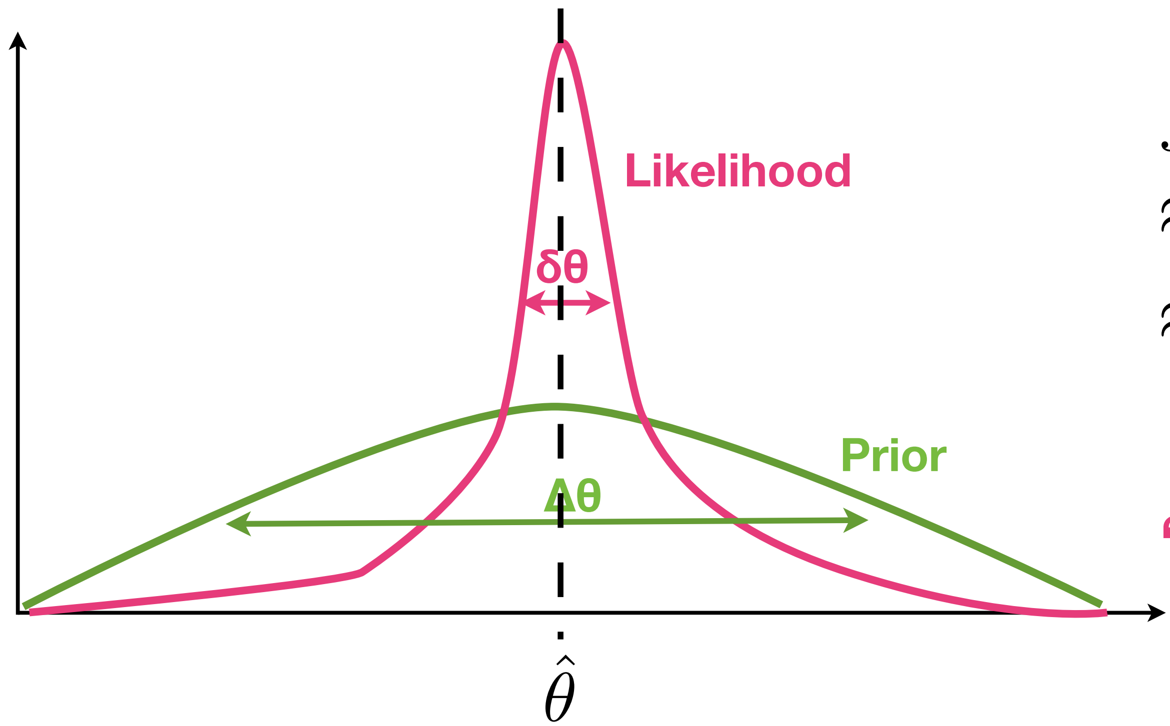
$$B_{01} \equiv \frac{P(d|M_0)}{P(d|M_1)}$$

Posterior odds = Bayes factor × prior odds

An in-built Occam's razor

Imperial College London

- The Bayesian evidence balances *quality of fit* vs *extra model complexity*.
- It rewards highly predictive models, penalizing "wasted" parameter space.
- **The prior here is important:** it quantifies the *predictive power* of the model.



$$P(d|M) = \int d\theta L(\theta) P(\theta|M) \approx L(\hat{\theta}) \delta\theta P(\hat{\theta}) \approx \frac{\delta\theta}{\Delta\theta} L(\hat{\theta})$$

"Occam's factor" **Quality of fit**

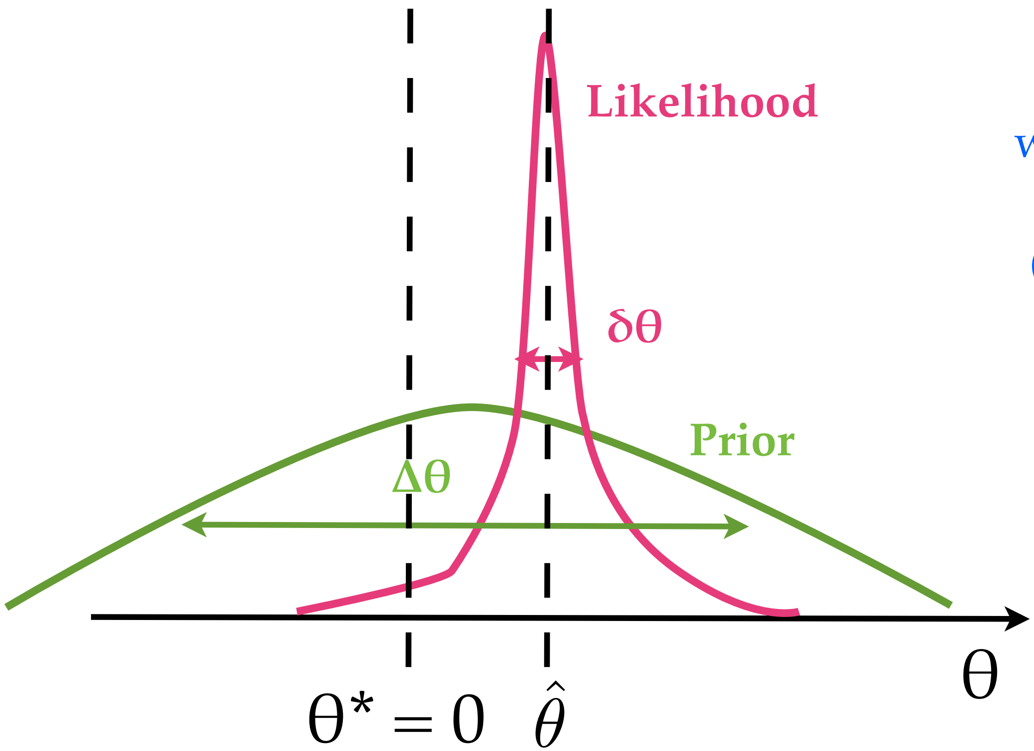
Roberto Trotta

Nested models

M$_0$: $\boldsymbol{\theta} = 0$
M$_1$: $\boldsymbol{\theta} \neq 0$ with prior p($\boldsymbol{\theta}$)

$$\lambda \equiv \frac{\hat{\theta} - \theta^\star}{\delta\theta}$$

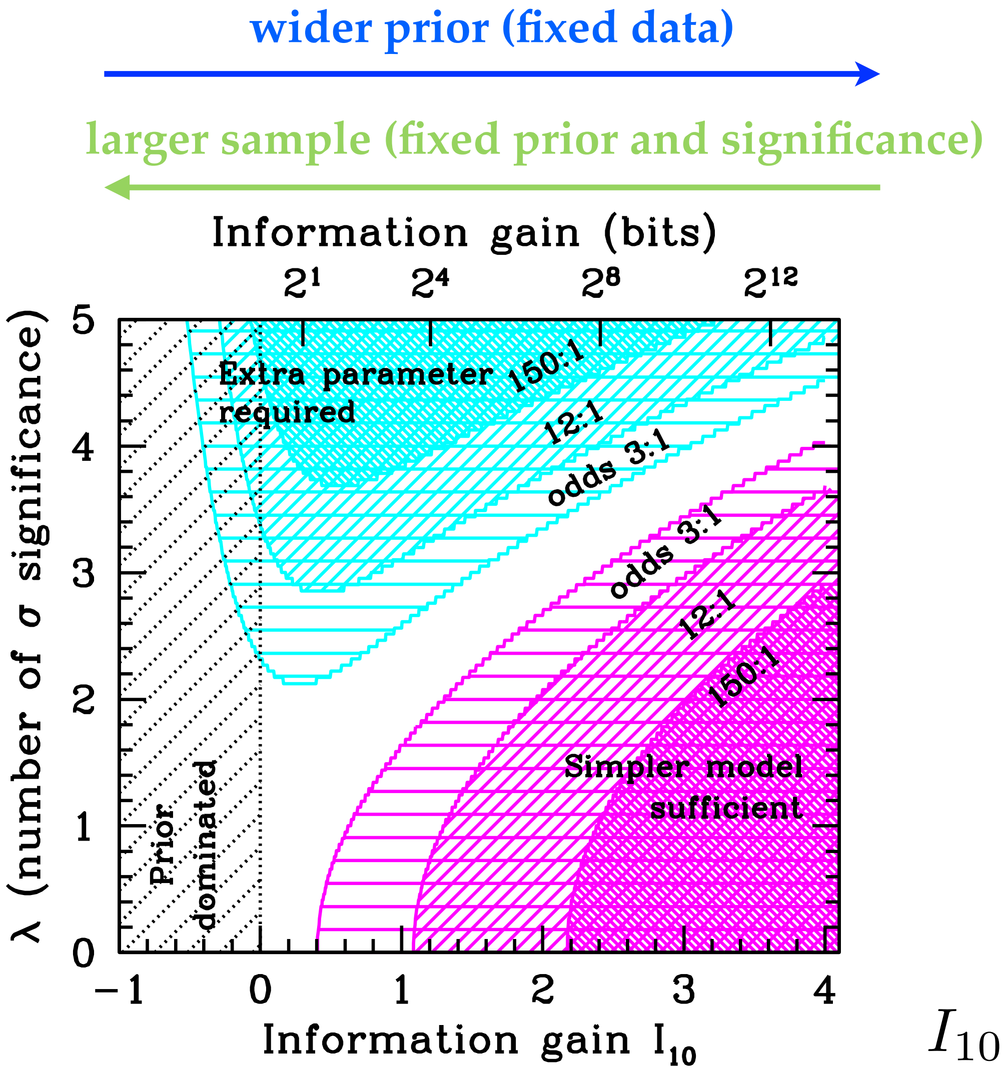
$$\ln B_{01} \approx \ln \frac{\Delta\theta}{\delta\theta} - \frac{\lambda^2}{2}$$



wasted parameter space (favours simpler model)

mismatch of prediction with observed data (favours more complex model)

Model selection for nested models



In Bayesian model comparison, **the prior scale never goes away.**

Also, the **alternative hypothesis** needs to be formulated from the outset (Jaynes: *"there is no point in rejecting a model unless one has a better alternative"*)

One should look at the scale of the prior and hope that the result is **robust** for "reasonable" prior choices

$$I_{10} \equiv \log_{10} \frac{\Delta\theta}{\delta\theta}$$

Trotta (2008)

Scale for the strength of evidence

Imperial College London

The Jeffreys' scale to assess the strength of evidence is empirically calibrated

$\lvert \ln B \rvert$	relative odds	favoured model's probability	Interpretation
< 1.0	$< 3:1$	< 0.750	not worth mentioning
< 2.5	$< 12:1$	0.923	weak
< 5.0	$< 150:1$	0.993	moderate
> 5.0	$> 150:1$	> 0.993	strong

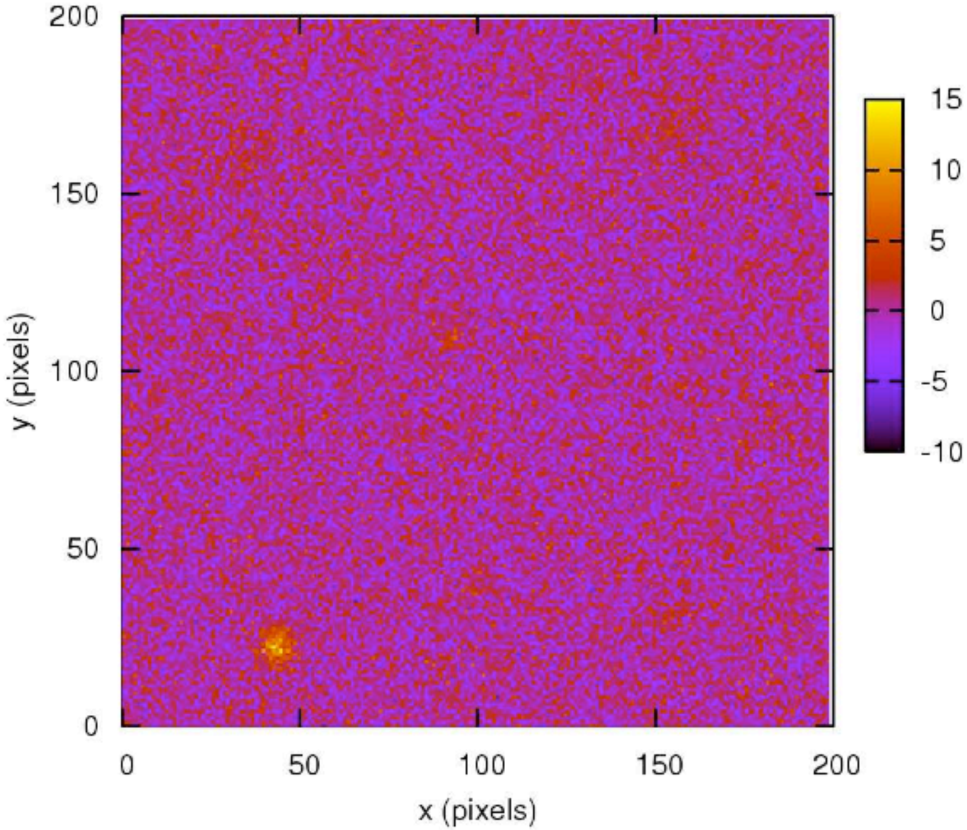
Trotta '08

Astro example: how many sources?

Imperial College London

Feroz and Hobson (2007)

Signal + Noise



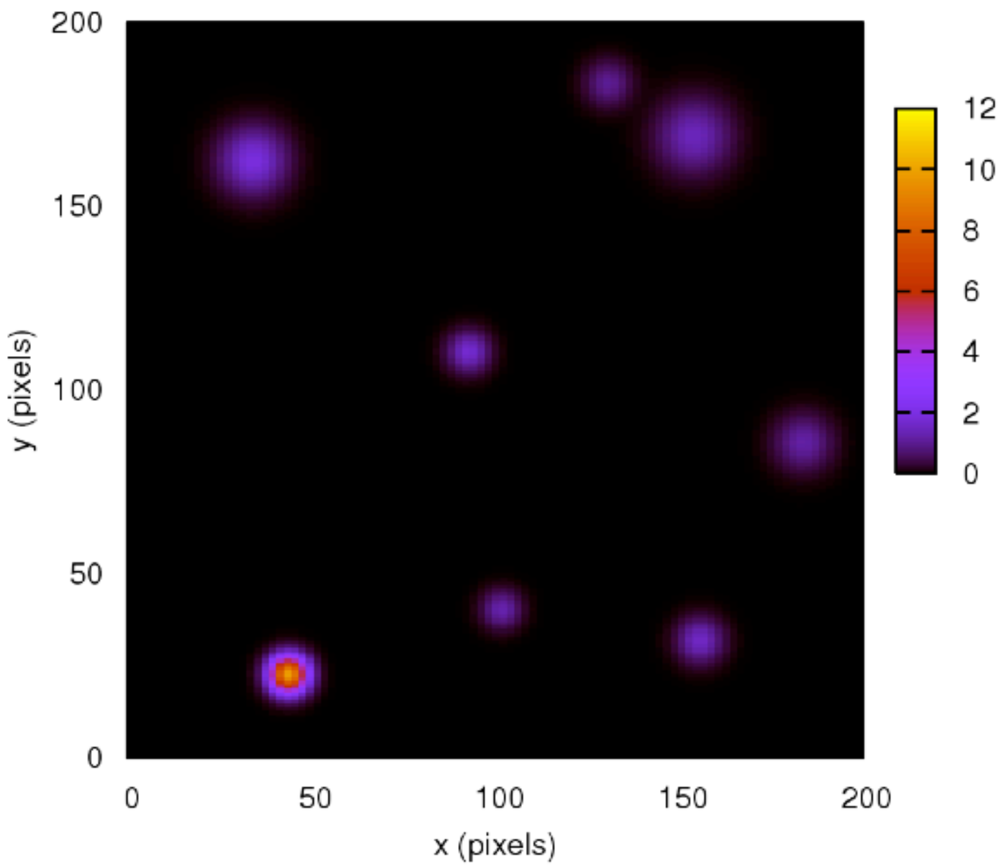
Roberto Trotta

Astro example: how many sources?

Imperial College London

Feroz and Hobson (2007)

Signal: 8 sources

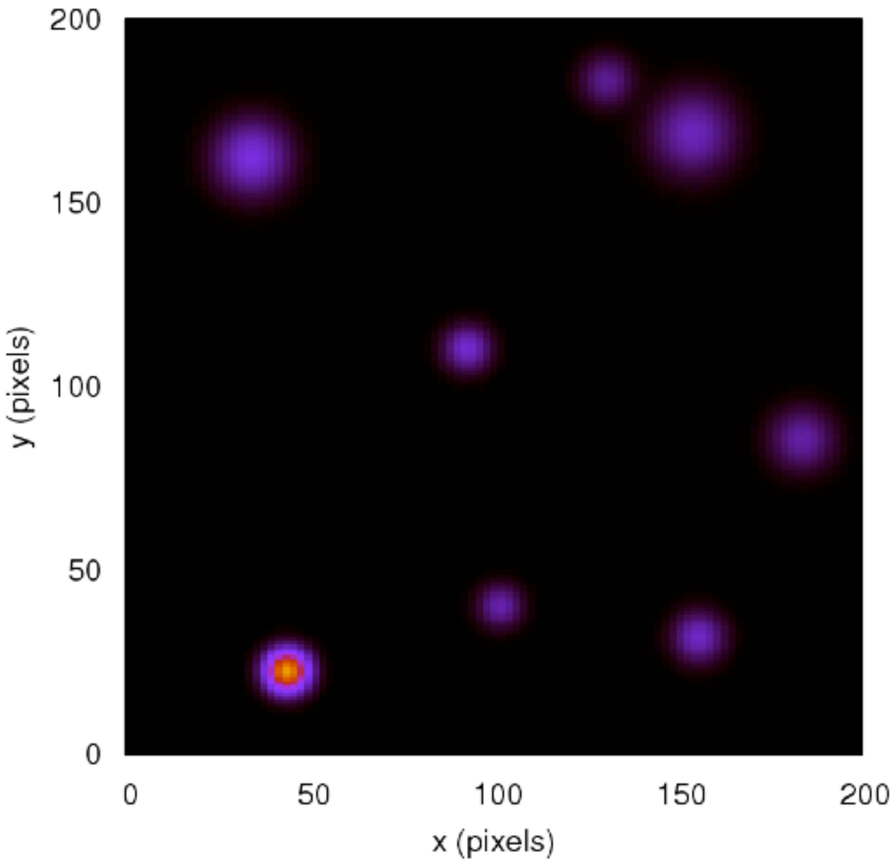


Roberto Trotta

Astro example: how many sources?

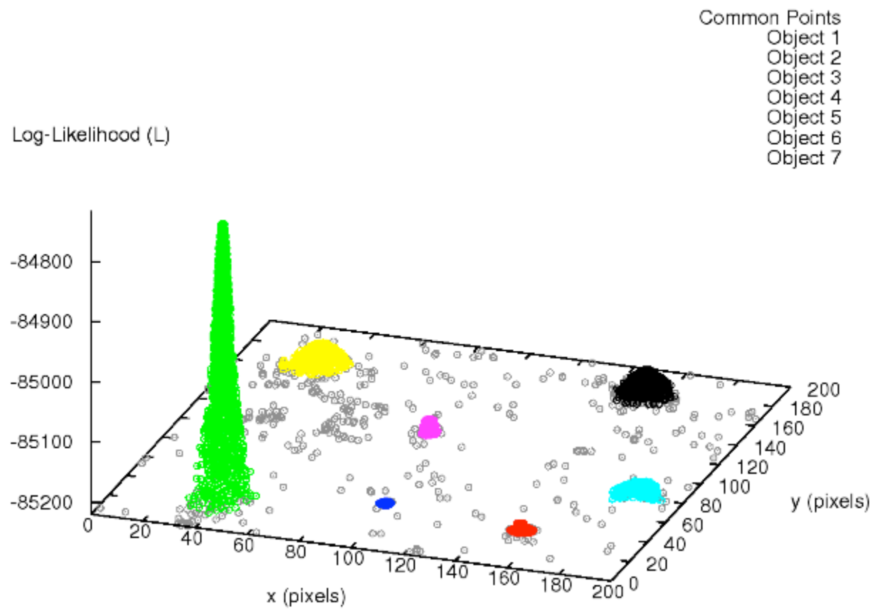
Imperial College London

Feroz and Hobson (2007)



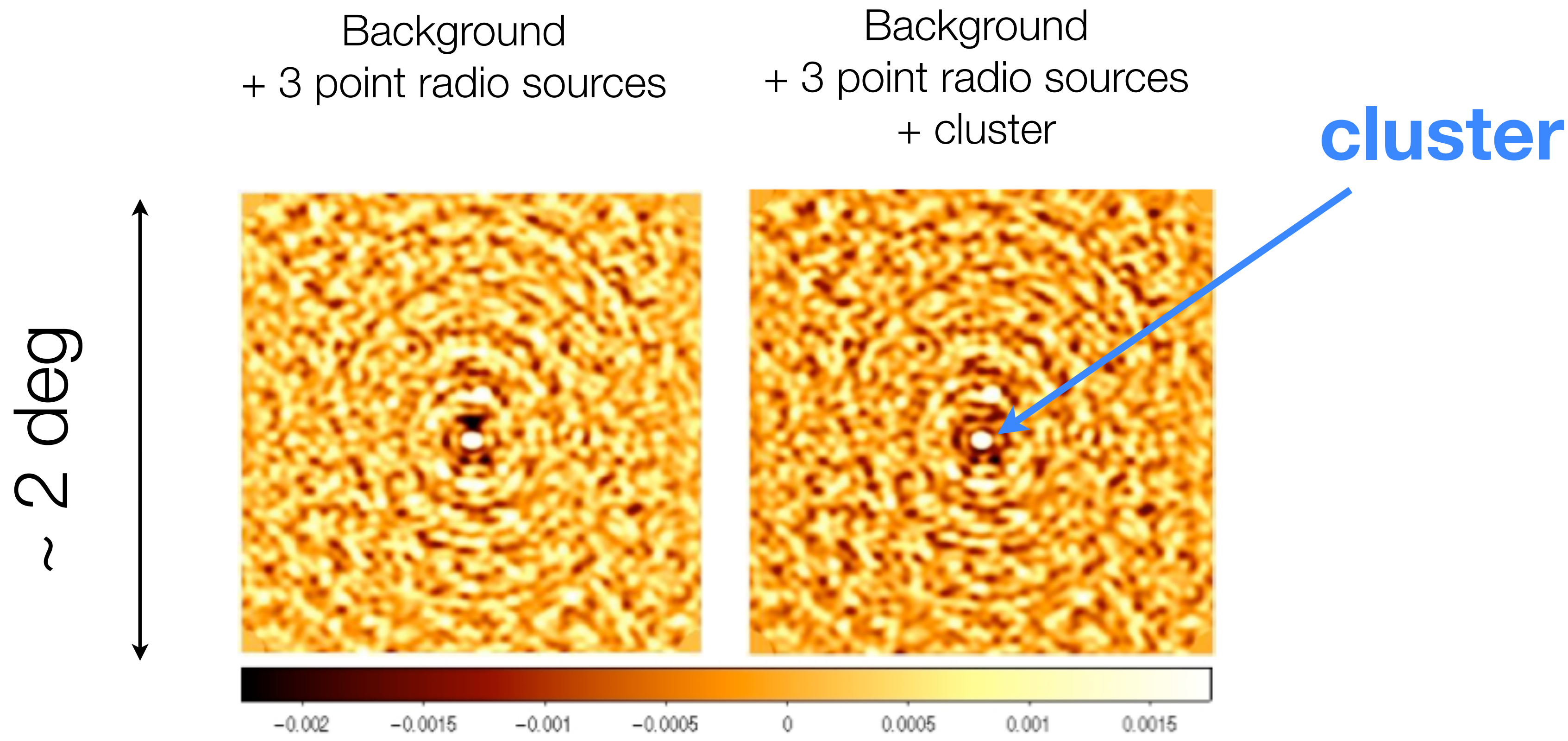
Bayesian reconstruction

7 out of 8 objects correctly identified.
Mistake happens because 2 objects very close.



Roberto Trotta

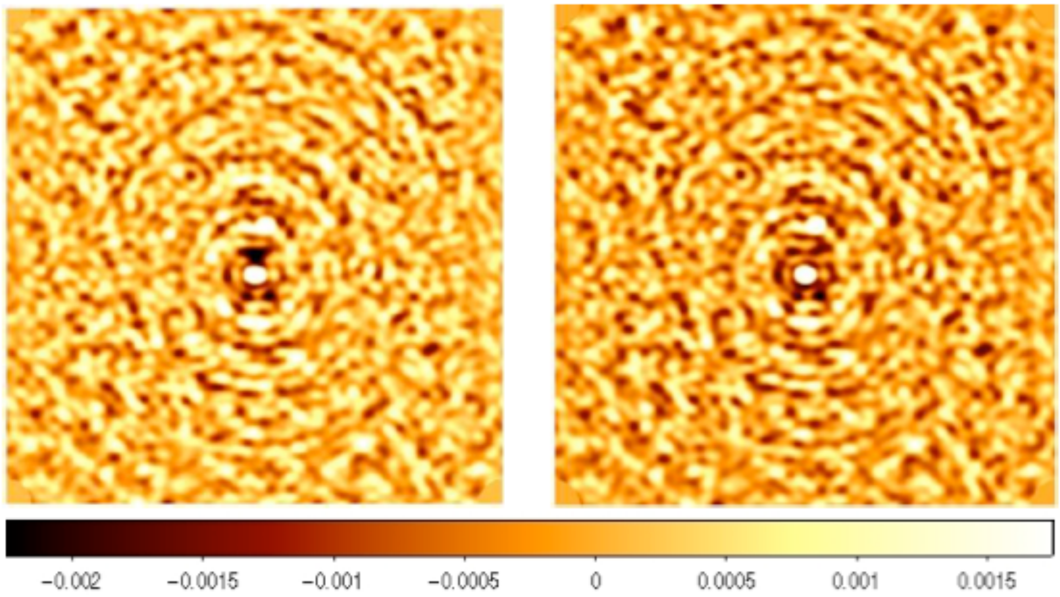
Cluster detection from Sunyaev-Zeldovich effect in cosmic microwave background maps



Feroz et al 2009

Background
+ 3 point radio sources

Background
+ 3 point radio sources
+ cluster



Posterior odds:

R = P(cluster | data)/P(no cluster | data)

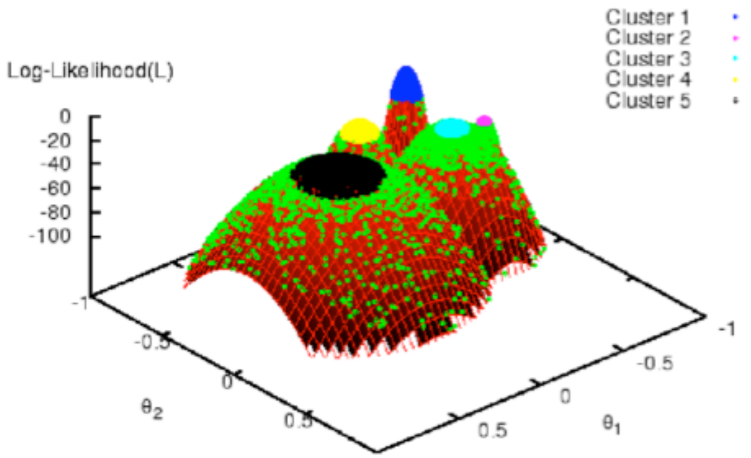
$R = 0.35 \pm 0.05$ $R \sim 10^{33}$

Cluster parameters also recovered (position, temperature, profile, etc)

Computation of the evidence with Multinest

Imperial College London

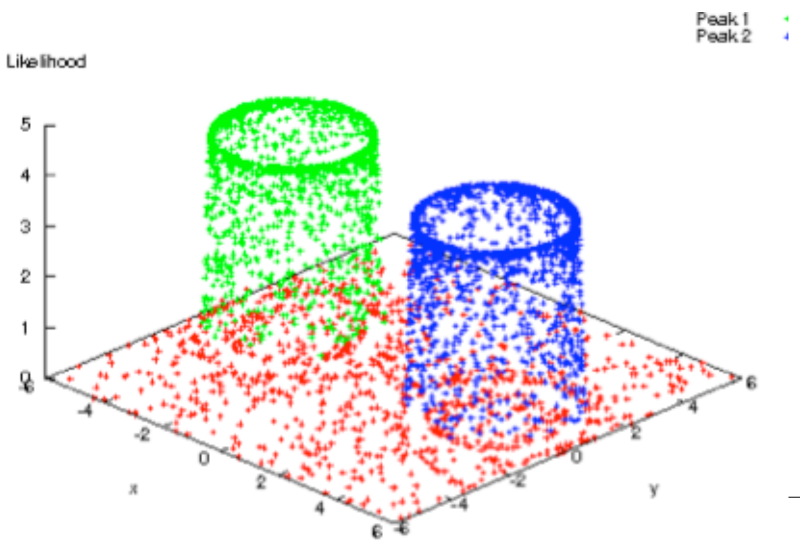
Feroz and Hobson (2007)



Gaussian mixture model:

True evidence: log(E) = -5.27
Multinest:
Reconstruction: log(E) = -5.33 ± 0.11
Likelihood evaluations ~ 10^4
Thermodynamic integration:
Reconstruction: log(E) = -5.24 ± 0.12
Likelihood evaluations ~ 10^6

Courtesy Mike Hobson



D	N_{like}	efficiency	likes per dimension
2	7000	70%	83
5	18000	51%	7
10	53000	34%	3
20	255000	15%	1.8
30	753000	8%	1.6

Roberto Trotta

Cosmological model comparison

Imperial College London

Competing model	ΔN_{par}	$\ln B$	Ref	Data	Outcome
Initial conditions					
Isocurvature modes					
CDM isocurvature	+1	-7.6	[58]	WMAP3+, LSS	Strong evidence for adiabaticity
+ arbitrary correlations	+4	-1.0	[46]	WMAP1+, LSS, SN Ia	Undecided
Neutrino entropy	+1	$[-2.5, -6.5]^p$	[60]	WMAP3+, LSS	Moderate to strong evidence for adiabaticity
+ arbitrary correlations	+4	-1.0	[46]	WMAP1+, LSS, SN Ia	Undecided
Neutrino velocity	+1	$[-2.5, -6.5]^p$	[60]	WMAP3+, LSS	Moderate to strong evidence for adiabaticity
+ arbitrary correlations	+4	-1.0	[46]	WMAP1+, LSS, SN Ia	Undecided
Primordial power spectrum					
No tilt ($n_s = 1$)	-1	$+0.4$	[47]	WMAP1+, LSS	Undecided
		$[-1.1, -0.6]^p$	[51]	WMAP1+, LSS	Undecided
		-0.7	[58]	WMAP1+, LSS	Undecided
		-0.9	[70]	WMAP1+	Undecided
		$[-0.7, -1.7]^{p,d}$	[186]	WMAP3+	$n_s = 1$ weakly disfavoured
		-2.0	[185]	WMAP3+, LSS	$n_s = 1$ weakly disfavoured
		-2.6	[70]	WMAP3+	$n_s = 1$ moderately disfavoured
		-2.9	[58]	WMAP3+, LSS	$n_s = 1$ moderately disfavoured
		$< -3.9^c$	[65]	WMAP3+, LSS	Moderate evidence at best against $n_s \neq 1$
Running	+1	$[-0.6, 1.0]^{p,d}$	[186]	WMAP3+, LSS	No evidence for running
		$< 0.2^c$	[166]	WMAP3+, LSS	Running not required
Running of running	+2	$< 0.4^c$	[166]	WMAP3+, LSS	Not required
Large scales cut–off	+2	$[1.3, 2.2]^{p,d}$	[186]	WMAP3+, LSS	Weak support for a cut–off
Matter–energy content					
Non–flat Universe	+1	-3.8	[70]	WMAP3+, HST	Flat Universe moderately favoured
		-3.4	[58]	WMAP3+, LSS, HST	Flat Universe moderately favoured
Coupled neutrinos	+1	-0.7	[193]	WMAP3+, LSS	No evidence for non–SM neutrinos
Dark energy sector					
$w(z) = w_{\rm eff} \neq -1$	+1	$[-1.3, -2.7]^p$	[187]	SN Ia	Weak to moderate support for Λ
		-3.0	[50]	SN Ia	Moderate support for Λ
		-1.1	[51]	WMAP1+, LSS, SN Ia	Weak support for Λ
		$[-0.2, -1]^p$	[188]	SN Ia, BAO, WMAP3	Undecided
		$[-1.6, -2.3]^d$	[189]	SN Ia, GRB	Weak support for Λ
$w(z) = w_0 + w_1 z$	+2	$[-1.5, -3.4]^p$	[187]	SN Ia	Weak to moderate support for Λ
		-6.0	[50]	SN Ia	Strong support for Λ
		-1.8	[188]	SN Ia, BAO, WMAP3	Weak support for Λ
$w(z) = w_0 + w_a(1-a)$	+2	-1.1	[188]	SN Ia, BAO, WMAP3	Weak support for Λ
		$[-1.2, -2.6]^d$	[189]	SN Ia, GRB	Weak to moderate support for Λ
Reionization history					
No reionization ($\tau = 0$)	-1	-2.6	[70]	WMAP3+, HST	$\tau \neq 0$ moderately favoured
No reionization and no tilt	-2	-10.3	[70]	WMAP3+, HST	Strongly disfavoured

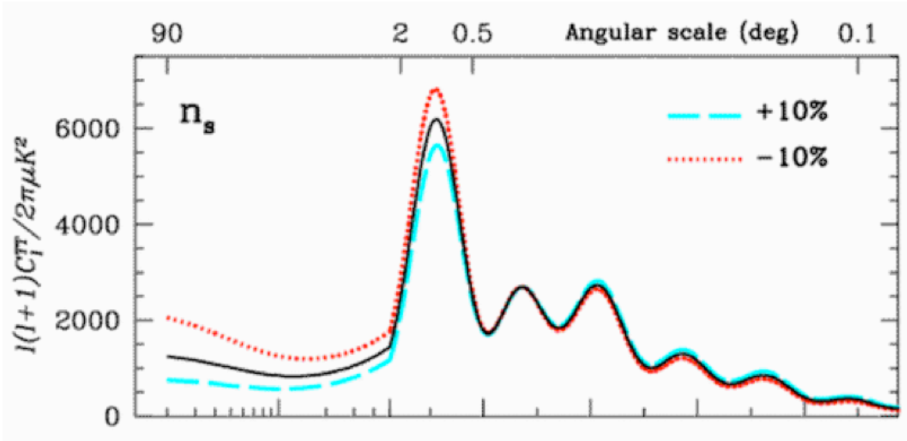
Trotta '08

lnB < 0: ΛCDM remains the "best" model from a Bayesian perspective!

Cosmological model selection

Imperial College London

- Is the spectrum of primordial fluctuations scale-invariant (n = 1)?
- Model comparison:
  n = 1 vs n ≠ 1 (with inflation-motivated prior)
- Results:
  n ≠ 1 favoured with odds of 17:1
  (Trotta 2007)
  n ≠ 1 favoured with odds of 15:1
  (Kunz, Trotta & Parkinson 2007)
  n ≠ 1 favoured with odds of 7:1
  (Parkinson 2007 et al 2006)

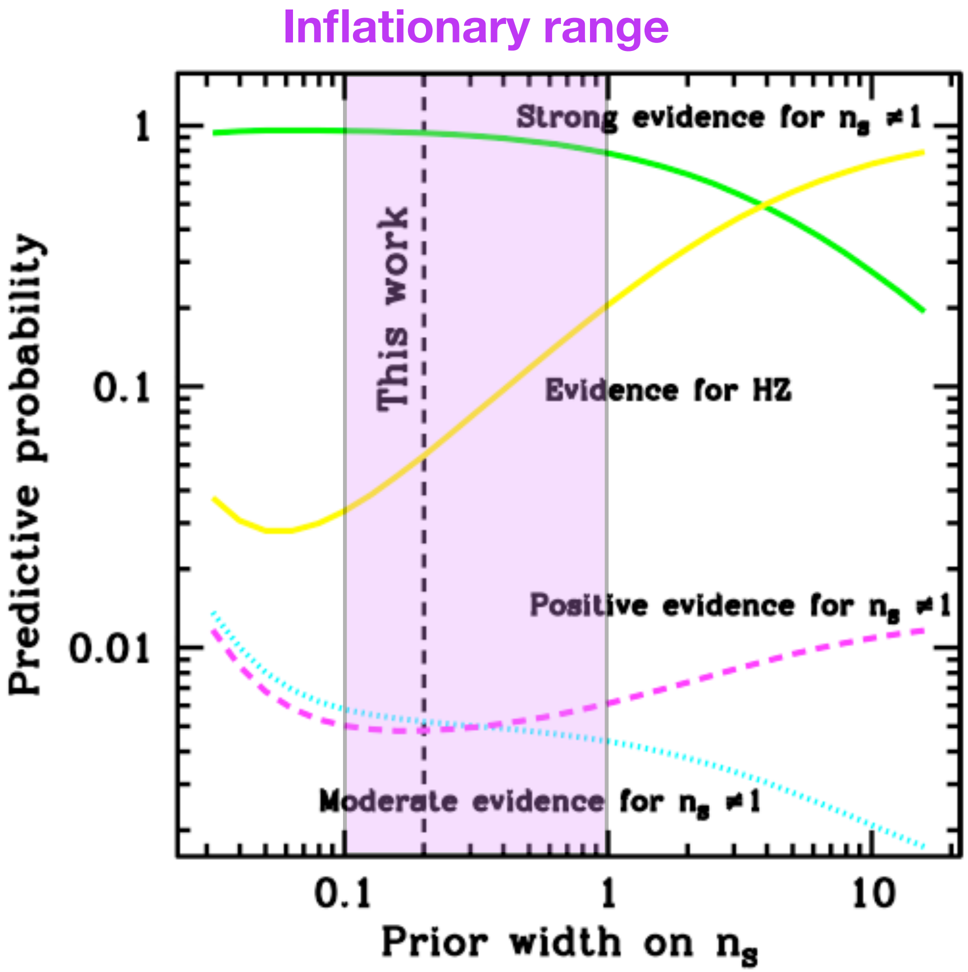


Roberto Trotta

Example of reasonable sensitivity analysis

Imperial College London

- The favoured model (non-scale invariant CMB spectrum) is robust for physically reasonable changes (motivated by inflation) in the prior width

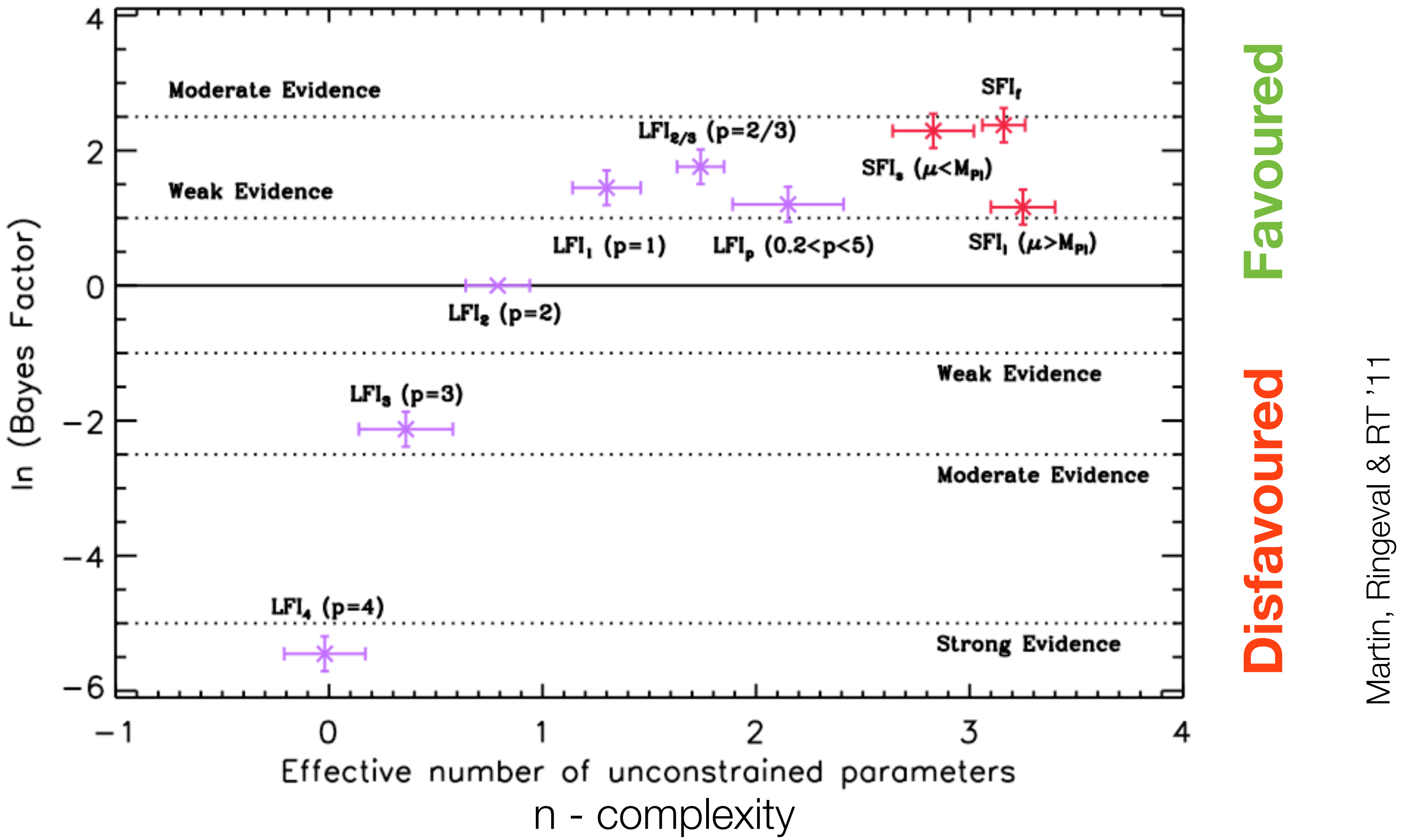


Trotta (2007)

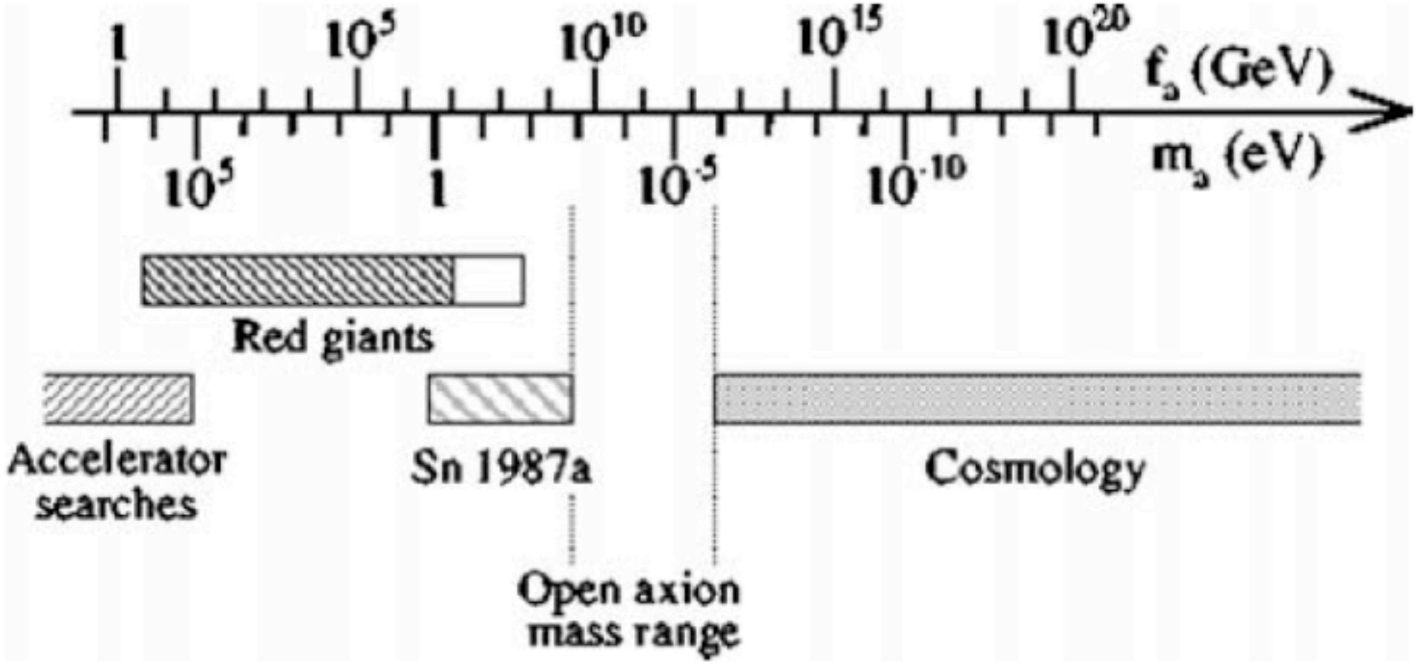
Small field vs large field inflation

Imperial College London

The probability of small field models rises from an initial 50% to
P(small field | all data) = 0.77 ± 0.03



Axion discovery space



For most exploratory experiments I can think of, these metrics just don't exist in a relevant way.

(Bob Cousin's talk at Banff 2010)

Information criteria

Imperial College London

- Several information criteria exist for approximate model comparison
  k = number of fitted parameters, N = number of data points,
  -2 ln(L_{max}) = best-fit chi-squared
- **Akaike Information Criterium (AIC):** $\text{AIC} \equiv -2\ln\mathcal{L}_{\text{max}} + 2k$
- **Bayesian Information Criterium (BIC):** $\text{BIC} \equiv -2\ln\mathcal{L}_{\text{max}} + k\ln N$
- **Deviance Information Criterium (DIC):** $\text{DIC} \equiv -2\widehat{D_{\text{KL}}} + 2\mathcal{C}_b.$

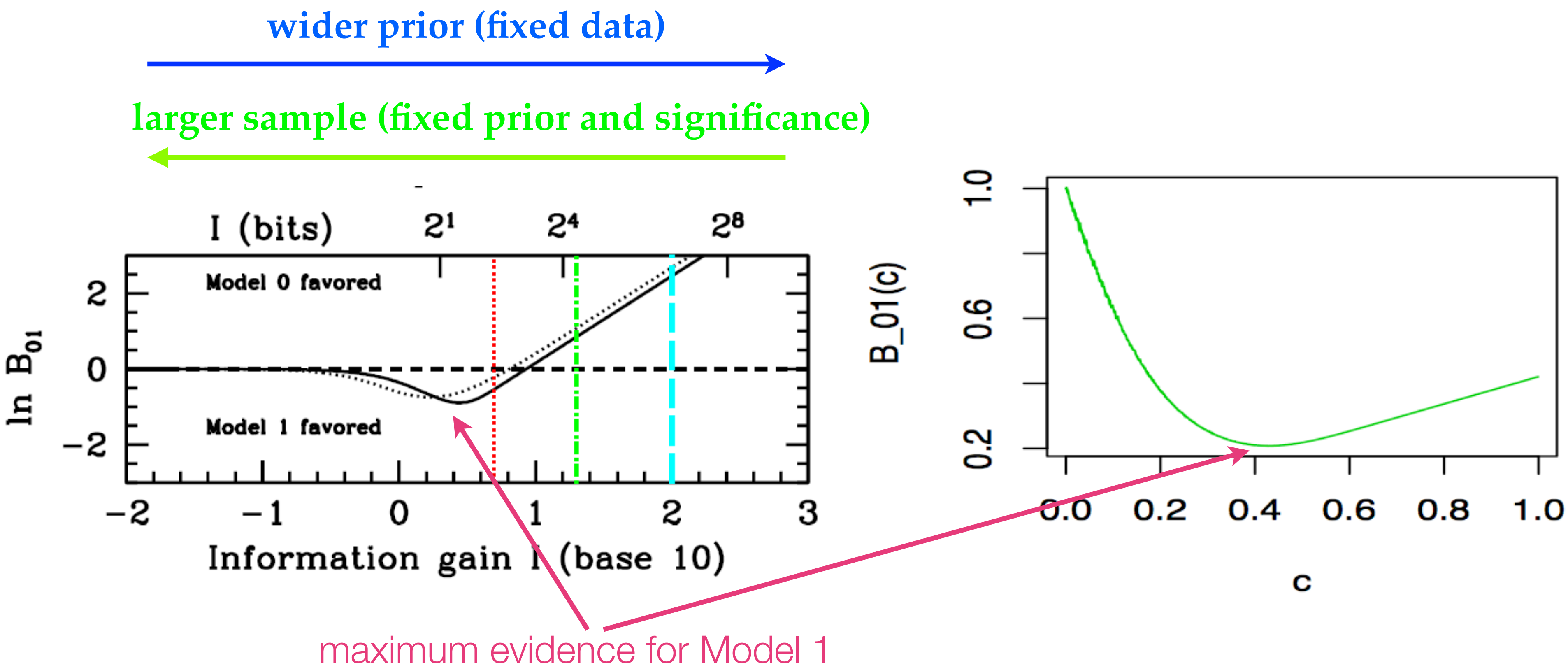
Pitfalls of Bayesian model selection

- In cosmology/High Energy Physics, there are many situations with nested models with extra unknown parameters for the fundamental theory.
- Little or nothing is known about the metric to be imposed on such a parameter space
- “The concept of total ignorance about θ does not have any precise meaning” (Bob Cousins)
- Often, deviations are looked for using arbitrarily parameterized alternative models (not tied to any specific physics), e.g. Gaussian Processes.
- Occam’s razor factor may be arbitrary. HOWEVER: if the range of your prior is arbitrary (by many orders of magnitude) then arguably the physics behind it is not strongly predictive...
- In some cases, the upper bound formalism might be useful (Jim Berger and collaborators)

“Prior-free” evidence bounds

Imperial College London

- What if we do not know how to set the prior? For nested models, we can still choose a prior that will maximise the support for the more complex model:



Maximum evidence for a detection

Imperial College London

- **The absolute upper bound:** put all prior mass for the alternative onto the observed maximum likelihood value. Then

$$B < \exp(-\chi^2/2)$$

- **More reasonable class of priors:** symmetric and unimodal around Ψ=0, then (α = significance level)

$$B < \frac{-1}{\exp(1)\alpha \ln \alpha}$$

If the upper bound is small, no other choice of prior will make the extra parameter significant.

Sellke, Bayarri & Berger, *The American Statistician*, 55, 1 (2001)

How to interpret the "number of sigma's"

Imperial College London

p	sigma	Absolute bound on lnB (B)	"Reasonable" bound on lnB (B)
0.05	2.0	2.0 (7:1) weak	0.9 (3:1) undecided
0.003	3.0	4.5 (90:1) moderate	3.0 (21:1) moderate
0.0003	3.6	6.48 (650:1) strong	5.0 (150:1) strong

A conversion table

p–value	$\bar{B}$	$\ln \bar{B}$	sigma	category
0.05	2.5	0.9	2.0	
0.04	2.9	1.0	2.1	'weak' at best
0.01	8.0	2.1	2.6	
0.006	12	2.5	2.7	'moderate' at best
0.003	21	3.0	3.0	
0.001	53	4.0	3.3	
0.0003	150	5.0	3.6	'strong' at best
6×10^{-7}	43000	11	5.0	

Rule of thumb:

a n-sigma result should be interpreted as a n-1 sigma result

A conversion table

p–value	$\bar{B}$	$\ln \bar{B}$	sigma	category
0.05	2.5	0.9	2.0	
0.04	2.9	1.0	2.1	‘weak’ at best
0.01	8.0	2.1	2.6	
0.006	12	2.5	2.7	‘moderate’ at best
0.003	21	3.0	3.0	
0.001	53	4.0	3.3	
0.0003	150	5.0	3.6	‘strong’ at best
6×10^{-7}	43000	11	5.0	

Rule of thumb:

a n-sigma result should be interpreted as a n-1 sigma result

A conversion table

Imperial College London

p–value	$\bar{B}$	$\ln \bar{B}$	sigma	category
0.05	2.5	0.9	2.0	
0.04	2.9	1.0	2.1	'weak' at best
0.01	8.0	2.1	2.6	
0.006	12	2.5	2.7	'moderate' at best
0.003	21	3.0	3.0	
0.001	53	4.0	3.3	
0.0003	150	5.0	3.6	'strong' at best
6×10^{-7}	43000	11	5.0	

Rule of thumb:

a n-sigma result should be interpreted as
a n-1 sigma result

Roberto Trotta

A conversion table

Imperial College London

p–value	$\bar{B}$	$\ln \bar{B}$	sigma	category
0.05	2.5	0.9	2.0	
0.04	2.9	1.0	2.1	'weak' at best
0.01	8.0	2.1	2.6	
0.006	12	2.5	2.7	'moderate' at best
0.003	21	3.0	3.0	
0.001	53	4.0	3.3	
0.0003	150	5.0	3.6	'strong' at best
6×10^{-7}	43000	11	5.0	

Rule of thumb:

a n-sigma result should be interpreted as a n-1 sigma result

A conversion table

Imperial College London

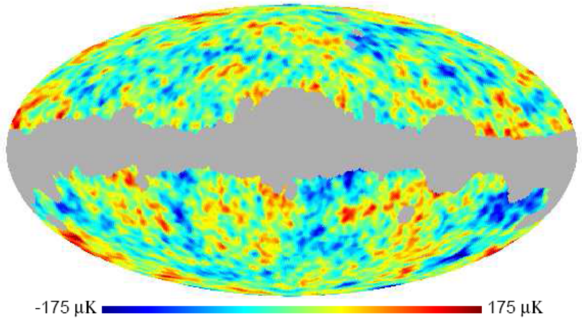
p–value	$\bar{B}$	$\ln \bar{B}$	sigma	category
0.05	2.5	0.9	2.0	
0.04	2.9	1.0	2.1	'weak' at best
0.01	8.0	2.1	2.6	
0.006	12	2.5	2.7	'moderate' at best
0.003	21	3.0	3.0	
0.001	53	4.0	3.3	
0.0003	150	5.0	3.6	'strong' at best
6×10^{-7}	43000	11	5.0	

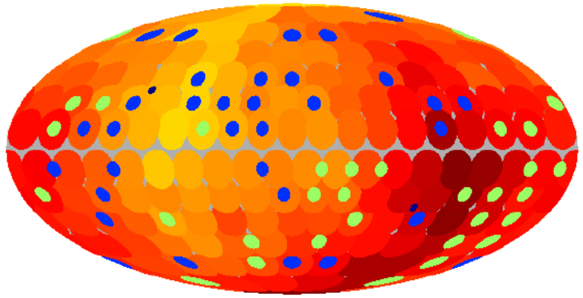
Rule of thumb:

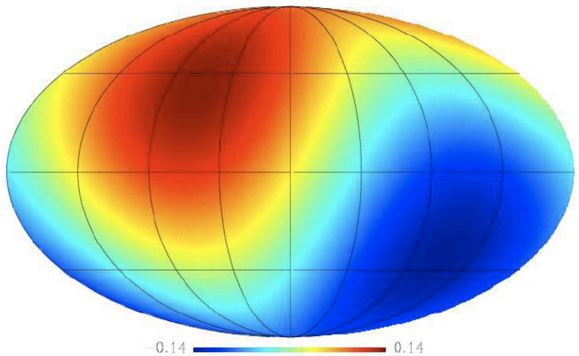
a n-sigma result should be interpreted as a n-1 sigma result

Application: dipole modulation

- Eriksen et al (2004) found hints for a dipolar modulation in WMAP1 ILC map
- Adding a phenomenological dipole pattern improves the chi-square by 9 units (for 3 extra parameters)
- Is this significant evidence?
- Not really: upper bound on B is odds of 9:1
  The absolute upper bound is about the same
  (Gordon and Trotta 2007)





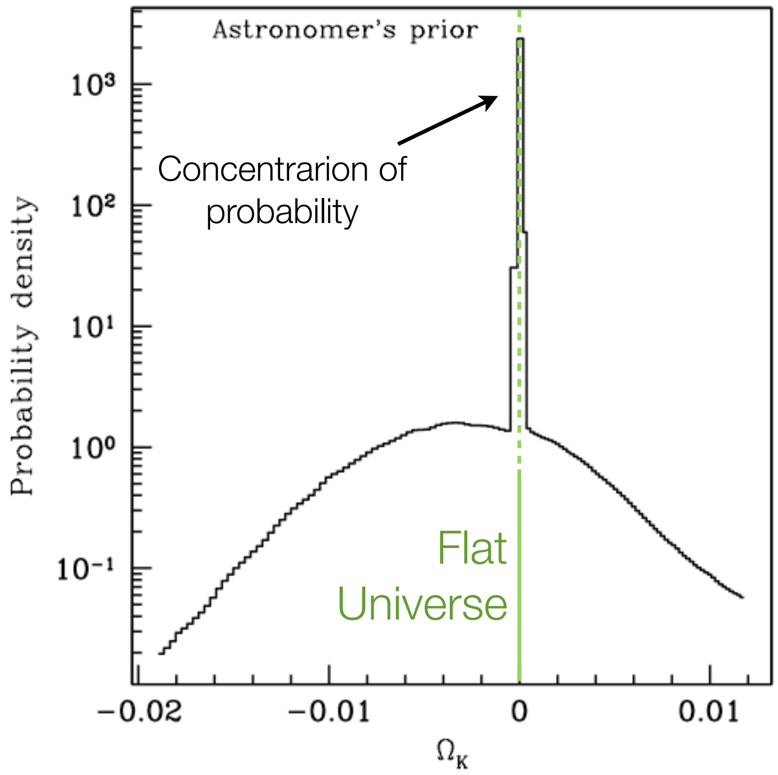


Level 3 inference: model averaging

Imperial College London

$$P(\theta|d) = \sum_i P(M_i|d)P(\theta|d, M_i)$$

- **Aim:** model-independent constraints that account for model uncertainty
- **Model posterior:** flat models are preferred by Bayesian model selection → probability gets concentrated onto those models
- **Consequence:** constraints on the curvature, number of Hubble spheres and size of the Universe can be **stronger** after Baysian model averaging!
- **Number of Hubble spheres** $N_U > 251$ (99%) ~8 times stronger
  Radius of curvature > 42 Gpc (99%)
  1.5 times stronger



Vardanyan, RT & Silk '11

Conclusions

SEARCHES

- **Searches** in astrophysics and cosmology are often poorly defined
- Generically, the Concordance Model gives a **“null hypothesis”** of sorts: Gaussian fluctuations, isotropy, scale invariance, trivial topology, cosmological constant, etc.
- Motivation often comes from **“looking for deviations from the null”**, even in the absence of well-motivated alternative models

Conclusions

Imperial College London

DETECTION

- When an anomaly is first “detected” it is almost invariably at the **sensitivity limit** (large foregrounds/backgrounds, poorly understood systematics, instrument insufficiently characterized, low S/N, small numbers statistics, etc)
- Anomalies are deemed “worth publishing” at the (shockingly) low **2-something-sigma-level**
- The scale of deviations from null predictions is **typically unknown.** For a principled Bayesian model comparison, the prior selection is challenging/impossible.
- (Too) Many anomalies/unexpected deviations go away with better data/modeling/insight. Is this evidence that the community jumps too early and too easily on **statistical flukes?**

Conclusions

Imperial College London

TECHNICALITIES

- A rigorous **hypothesis testing** is usually not performed. Instead, counting of "how-many-sigma-away" from the null the ML lies. Conditions of validity of Wilks/Chernoff theorems usually ignored.
- **No repeated experiment possible** for large scale phenomena (e.g., anomalies of the CMB, topology)
- **Bayesian model selection:** in principle desirable (perhaps), but often impractical. No meaningful measure on the space of unknown alternative models exists in many cases of interest.

Open questions

Imperial College London

- Is Bayesian model selection the correct framework?

  "Bayesians address the question everyone is interested in by using assumptions no–one believes, while frequentists use impeccable logic to deal with an issue of no interest to anyone" Louis Lyons

  - Popperian view according to which models start off being infinitely improbable (and stay like that no matter the confirmative evidence) is untenable. Falsification only half of the story!

  - Criticisms from e.g. G. Efstathiou (arXiv:0802.3185) and Bob Cousins (Phys.Rev.Lett.101,029101, 2008)

  - Bayesian model selection matches goodness-of-fit tests for specific choices of priors

- How do we deal with Lindley's paradox? (Lindley, 1954)

  It is easy to construct cases where frequentist hypothesis testing and Bayesian model comparison disagree. How should we interpret the result?

- What do we do when Frequentist (profile likelihood) and Bayesian (marginal posterior) inferences disagree even at the level of parameter inference? What is the scientific outcome/conclusion of the measurement?

- How do we assess the completness of the set of known models in a Bayesian context?

  Bayesians maintain that it is useless to reject a model unless a better alternative is available. However, an absolute scale of model adequacy seems useful. Can Bayesian model comparison be extended to the space of unknown models? (e.g., March, RT et al, 2010)

More questions...

- Is Bayesian model averaging useful?

  Are model-averaged constraints useful, and if so in which context? Is the propagation of the Occam's razor effect onto parameter inference problematic?

- Is there such a thing as a "correct" prior?

  In fundamental physics, models and parameters (and their priors) are supposed to represent (albeit in an idealized way) the real world, i.e., they are not simply useful representation of the data (as they are in other statistical problems).
  One could imagine that there exist a "correct" (i.e., tied to physics) prior for parameters of our cosmological model, which could in principle be derived from fundamental theories such as string theory.

Conclusion: Searches in astrophysics

Thank you!