Curvature tension

Evidence for a closed universe (?)

Will Handley
wh260@cam.ac.uk

Astrophysics Group, Cavendish Laboratory, University of Cambridge
Kavli Institute for Cosmology, Cambridge
Gonville & Caius college, Cambridge

25th August 2019

Handley arXiv:1908.09139, arXiv:1907.08524
Handley & Lemos arXiv:1902.04029, arXiv:1903.06682

`github.com/williamjameshandley/CosmoChord`
`github.com/williamjameshandley/anesthetic`

Outline

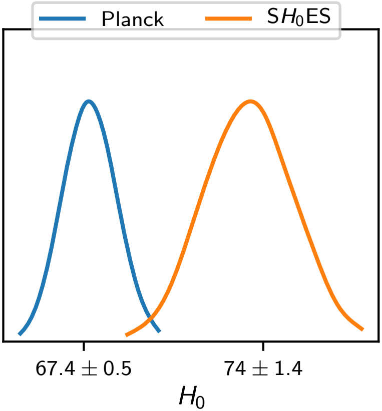
A talk of two halves

First half: Inference, quantifying tensions and observing curvature

Second half: Theory and predictions of primordially curved universes

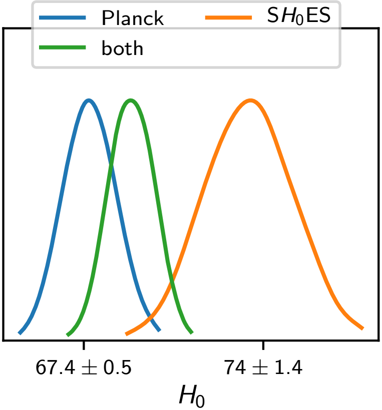
The Hubble H_0 tension

- CMB cosmologists (Planck) infer $H_0 = 67 \pm 0.5$ km s^{-1}Mpc^{-1}
- Supernovae data (SH_0ES) measure $H_0 = 74 \pm 1.4$
- $> 4\sigma$ discrepancy could be due to:
  - Systematic error
  - Problem with standard model of cosmology (ΛCDM)
- Inconsistent datasets shouldn't be combined



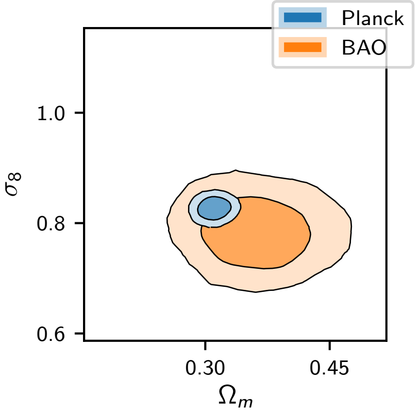
The Hubble H_0 tension

- CMB cosmologists (Planck) infer $H_0 = 67 \pm 0.5$ km s^{-1}Mpc^{-1}
- Supernovae data (SH_0ES) measure $H_0 = 74 \pm 1.4$
- $> 4\sigma$ discrepancy could be due to:
  - Systematic error
  - Problem with standard model of cosmology (ΛCDM)
- Inconsistent datasets shouldn't be combined

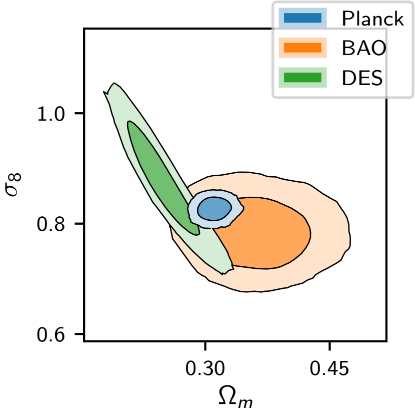


σ_8-Ω_m tension

- Matter density Ω_m and RMS matter fluctuations σ_8 are constrained by
- BAO and Planck look consistent

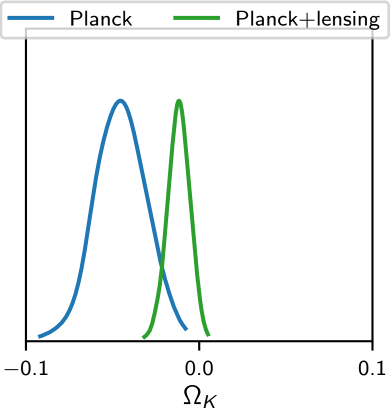


- Matter density Ω_m and RMS matter fluctuations σ_8 are constrained by
- BAO and Planck look consistent
- DES is less clear
- How do you define a tension in terms of "sigma" for this case?



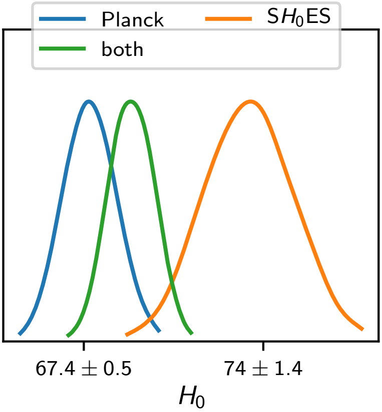
Ω_K tension

- Models with spatial curvature Ω_K.
- Best-kept secret of Planck: only $1/10,000$ MCMC samples $\Omega_K > 0$.
- How consistent do Planck and CMB lensing look?
- arXiv:1908.09139

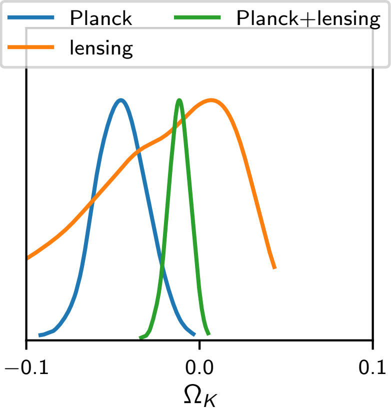


Ω_K tension

- Models with spatial curvature Ω_K.
- Best-kept secret of Planck: only $1/10,000$ MCMC samples $\Omega_K > 0$.
- How consistent do Planck and CMB lensing look?
- arXiv:1908.09139



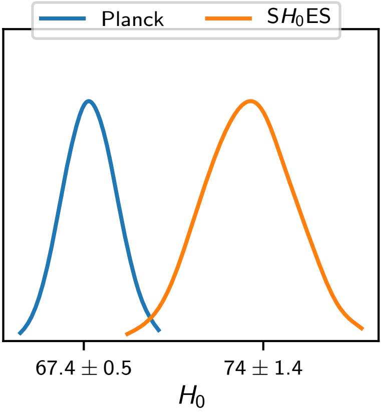
- Models with spatial curvature Ω_K.
- Best-kept secret of Planck: only $1/10,000$ MCMC samples $\Omega_K > 0$.
- How consistent do Planck and CMB lensing look?
- arXiv:1908.09139



For 1D Gaussian distributions, tension is pretty easy to define:

$$X = \frac{|\mu_A - \mu_B|}{\sqrt{\sigma_A^2 + \sigma_B^2}},$$

where μ and σ are the respective parameter means and standard deviations.

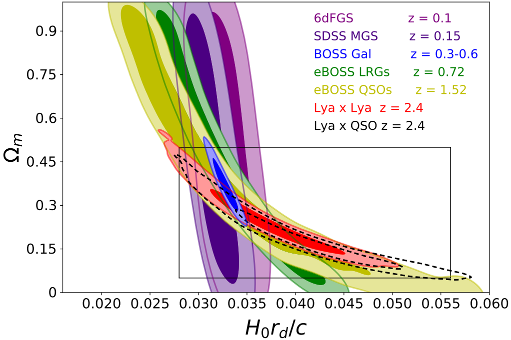


The multivariate d-dimensional equivalent to this tension would be:

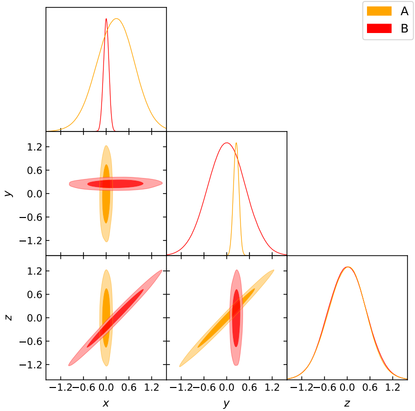
$$X_d^2 = (\mu_A - \mu_B)^T (\Sigma_A + \Sigma_B)^{-1} (\mu_A - \mu_B),$$

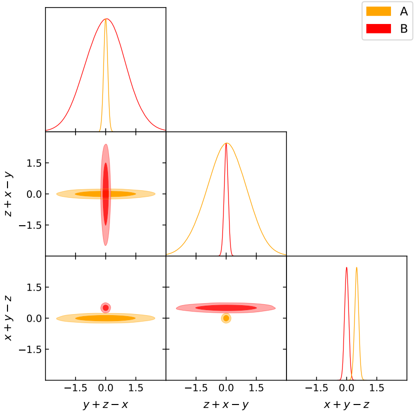
where Σ is in general a covariance matrix.

- Things become less clear when distributions become “banana like” (arXiv:1906.11628), or worse, multimodal.



- Many attempts to generalise the Gaussian case result in a parameterisation-dependent quantity.





- In high dimensions, things can look good when projected into 2D.
- We need a systematic way of seeking out tension, without relying on inspired choices of parameters to reveal them

The DES evidence ratio R

- The Dark Energy Survey (arXiv:1708.01530) quantifies tension between two datasets A and B using the Bayes ratio:

$$R = \frac{\mathcal{Z}_{AB}}{\mathcal{Z}_A \mathcal{Z}_B}$$

  where $\mathcal{Z}$ is the Bayesian evidence.
- Many attractive properties:
  - Symmetry
  - Parameterisation independence
  - Dimensional consistency
  - Use of well-defined Bayesian quantities
- What does it mean?

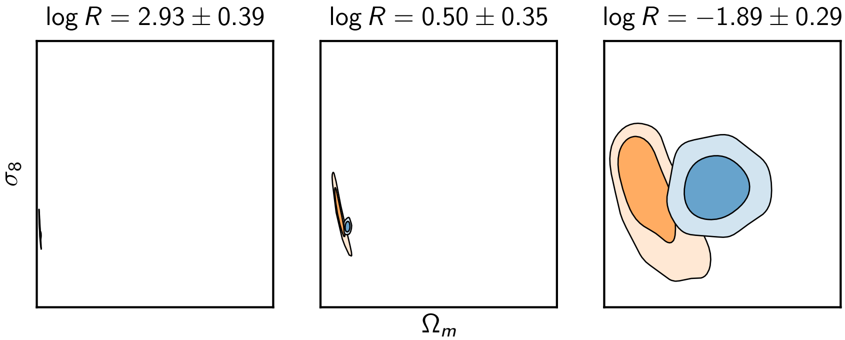
The meaning of the DES evidence ratio R

- The Dark Energy Survey collaboration (arXiv:1708.01530) quantify tension between two datasets A and B using the Bayes ratio:

$$R = \frac{\mathcal{Z}_{AB}}{\mathcal{Z}_A \mathcal{Z}_B} = \frac{P(A \cap B)}{P(A)P(B)} = \frac{P(A|B)}{P(A)} = \frac{P(B|A)}{P(B)}$$

- R gives the relative change in our confidence in data A in light of having seen B (and vice-versa).
- $R > 1$ implies we have more confidence in A having received B.
- Like evidences, it is prior-dependent
- Increasing prior widths $\Rightarrow$ increasing confidence.

The DES evidence ratio R: Prior dependency



- What does it mean if increasing prior widths $\Rightarrow$ increasing confidence?
- Wide priors mean *a-priori* the parameters could land anywhere.
- We should be proportionally more reassured when they land close to one another if the priors are wide

How do we deal with the prior dependency in R?

Option 1 Take the Bayesian route, accept the prior dependency, and spend time trying to justify why a given set of priors are "physical".

Option 2 Try to find a principled way of removing this prior dependency

- One of the critical observations is that one can only hide tension by widening priors. Narrowing them will only ever show tension if it is present.
- If we could define "Narrowest reasonable priors" and find that $R < 1$, then this would indicate tension.

R: a Gaussian example

- Given two Gaussians with parameter means μ_A, μ_B and parameter covariances Σ_A, Σ_B and a prior with volume V_π:

$$\log R = -\frac{1}{2}(\mu_A - \mu_B)(\Sigma_A + \Sigma_B)^{-1}(\mu_A - \mu_B) \\ + \log V_\pi - \log \sqrt{|2\pi(\Sigma_A + \Sigma_B)|}$$

- Like evidence, R composed of "Goodness of fit", and "Occam factor".
- Ideally want would remove this Occam factor (ratio of prior to posterior volume).

KL divergence $\mathcal{D}$, Information $\mathcal{I}$, suspiciousness S

- The KL divergence quantifies the compression from prior to posterior:

$$\mathcal{D} = \int P(\theta|D) \log \frac{P(\theta|D)}{P(\theta)} d\theta = \left\langle \log \frac{\text{Posterior}}{\text{Prior}} \right\rangle_{\text{Posterior}}$$

- It bears many similarities to an Occam factor, for a Gaussian:

$$\mathcal{D} = \log V_\pi - \log \sqrt{|2\pi\Sigma|} - \frac{1}{2}d$$

- Can define equivalent of R for KL divergence, the information ratio $\mathcal{I}$

$$\log R = \mathcal{Z}_{AB} - \mathcal{Z}_A - \mathcal{Z}_B$$
$$\log \mathcal{I} = \mathcal{D}_A + \mathcal{D}_B - \mathcal{D}_{AB}$$

- Subtracting the two removes prior dependency, giving suspiciousness:

$$\log S = \log R - \log \mathcal{I}$$

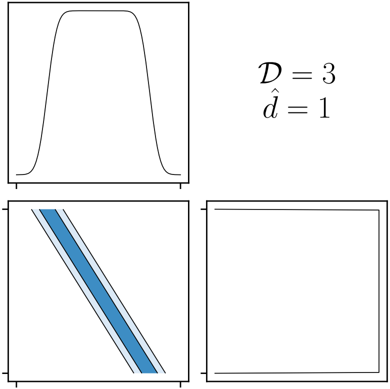
Suspiciousness S

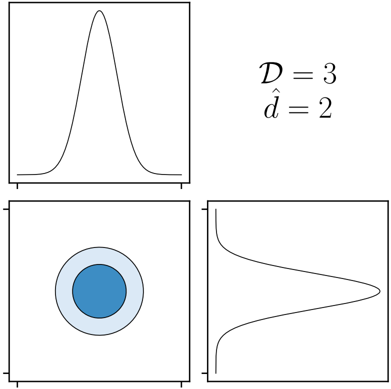
- For a Gaussian:

$$\log S = \frac{d}{2} - \frac{1}{2}(\mu_A - \mu_B)(\Sigma_A + \Sigma_B)^{-1}(\mu_A - \mu_B).$$

- We thus find that our original idea for tension $X_d^2 = d - 2\log S$.
- However S is composed of evidences $\mathcal{Z}$ and KL divergences $\mathcal{D}$, which are Gaussian-independent concepts.
- The only thing remaining to determine is d, the "number of parameters".

Dimensionality d





- Intuition should tell us that the d we need is the effective number of parameters (i.e. should not include unconstrained ones).
- Like the evidence, or the KL divergence, this "Model dimensionality" should be a sought-after inference quantity.

Dimensionality $\tilde{d}$

- KL divergence is the mean of the Shannon information I:

$$\mathcal{D} = \int P(\theta|D) \log \frac{P(\theta|D)}{P(\theta)} d\theta = \left\langle \log \frac{\text{Posterior}}{\text{Prior}} \right\rangle_{\text{Posterior}}$$

$$I = \log \frac{\text{Posterior}}{\text{Prior}}$$

- Model dimensionality proportional to variance of Shannon information:

$$\frac{\tilde{d}}{2} = \mathrm{var}\left(\frac{\text{Posterior}}{\text{Prior}} \right)_{\text{Posterior}}$$

- Examples from real data:

$$\begin{aligned} \tilde{d}_{\text{Planck}} &= 15.8 \pm 0.3 && (21) \\ \tilde{d}_{\text{DES}} &= 14.0 \pm 0.3 && (26) \\ \tilde{d}_{\text{BAO}} &= 2.95 \pm 0.07 && (6) \\ \tilde{d}_{\text{S}H_0\text{ES}} &= 0.93 \pm 0.03 && (6) \end{aligned}$$

Headline results: ΛCDM

- Can calibrate X_d^2 as on the same scale as χ_d^2 to give a p-value-like quantity, termed "Tension probability" p, or σ

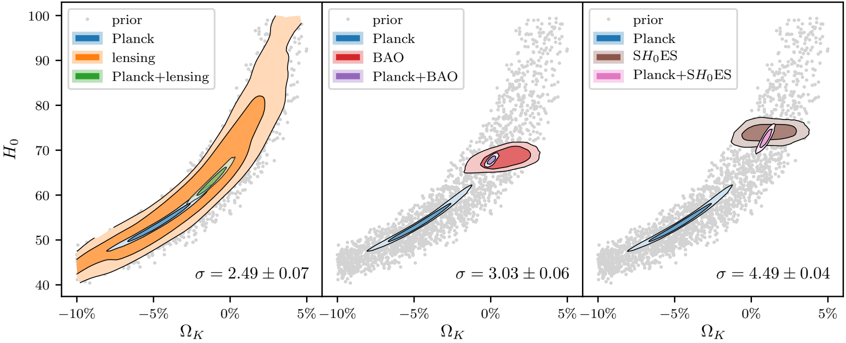
$$
\begin{aligned}
\text{Planck 2015 vs BAO}: &\quad p = 42 \pm 4\% & \sigma = 0.8 \pm 0.1 \\
\text{Planck 2015 vs DES}: &\quad p = 3.2 \pm 1.0\% & \sigma = 2.1 \pm 0.1 \\
\text{Planck 2015 vs S}H_0\text{ES}: &\quad p = 0.25 \pm 0.17\% & \sigma = 3.1 \pm 0.3 \\
\text{Planck 2018 vs S}H_0\text{ES}: &\quad p = 0.001 \pm 0.0001\% & \sigma = 4.42 \pm 0.03
\end{aligned}
$$

- Under this metric, SH_0ES is unambiguously inconsistent. BAO is consistent, and *DES* is inconsistent, but only just. This is pleasingly similar to ones intuition.
- arXiv:1902.04029

Headline results: $K\Lambda$CDM $\equiv \Lambda$CDM $+ \Omega_K$

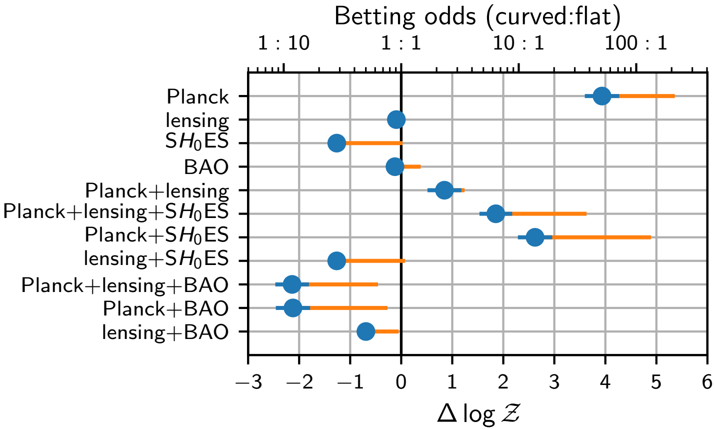
- Cosmologies with curvature have different tension (arXiv:1908.09139):

$$\begin{aligned} \text{Planck vs CMB lensing}: \quad & p = 1.3 \pm 0.3\% \qquad \sigma = 2.5 \pm 0.1 \\ \text{Planck vs BAO}: \quad & p = 0.25 \pm 0.05\% \qquad \sigma = 3.0 \pm 0.1 \end{aligned}$$



- One should be suspicious about combining CMB lensing and Planck in curved models.

Evidence: Is the universe curved?



- We only think the universe is flat because Planck+lensing and/or BAO tell us so, but these datasets are in tension.
- Planck alone tells us the universe is closed at 50:1

What about A_{lens}?

- The Planck papers do discuss this, but phrase in terms of A_{lens}
- A_{lens} is a "fudge factor" that artificially increases the effective smoothing by lensing on the primordial power spectrum if $A_{\mathrm{lens}} \neq 1$
- Planck CMB prefers $A_{\mathrm{lens}} \sim 1.2$, lensing wants $A_{\mathrm{lens}} \sim 1$ (can also be defined as a $\sim 2\sigma$ tension)
- It is the same effect that causes the preference for closed universes.

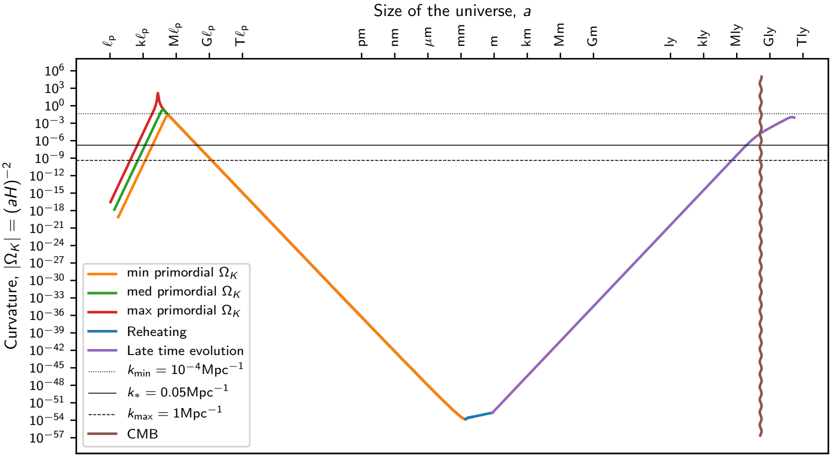
How do we fix curvature tension?

- There could be several explanations:
- Systematic in Planck CMB
  - Efstathiou & Gratton's new likelihood looks like it may relax these tensions a little (results pending release of likelihood)
- Systematic in Planck lensing
  - The lensing likelihood includes a fiducial ΛCDM model.
  - *In theory* this shouldn't matter, but re-doing the analysis with a different fiducial model is a mammoth task.
- Statistical fluctuation?
  - With tensions at 2–3σ there is always a possibility that we just got unlucky with our universe.
  - There are a lot of these 2–3σ tensions in the Planck data.
  - If there were an explanation that resolved multiple tensions simultaneously then this could be decicive evidence for a new concordance model.
- It may be that we need to wait for the next generation CMB experiment to resolve the issue (SO, LiteBird, PICO?, CORE?)

Theory of curved universes

- We generally work with flat cosmologies for a few reason
  1. Data tell us it is (or does it?)
  2. Theory of inflation tells us it is (somewhat circular reasoning)
  3. The theory is **much** easier in flat cosmologies
- I've been interested in the effect of curvature at the start of inflation for a while
- Even small amounts of curvature can have observational effects on the primordial power spectrum

History of curved universes



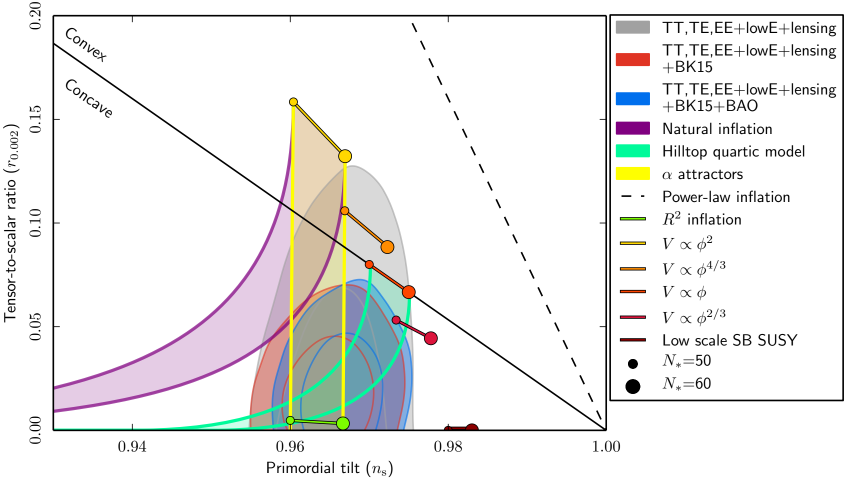
The primordial power spectrum

- Inflationary theory interacts primarily with observation via the primordial power spectrum of curvature perturbations $\mathcal{P}_\mathcal{R}(k)$ (along with the tensor to scalar ratio r)
- ΛCDM assumes a near scale-invariant power spectrum

$$\mathcal{P}_\mathcal{R}(k) = A_s \left(\frac{k}{k_s}\right)^{n_s - 1}$$

- Inflation theories generally predict values of (r, n_s) which can be compared with the phenomenological fit.
- A_s is constrained by (amongst other things) the amplitude of the cosmic microwave background power spectrum

Predictions from inflation



Primordial power spectrum computation

- Background variables:

$$H^2 = \frac{1}{3}\left(\frac{1}{2}\dot{\phi}^2 + V(\phi)\right), \qquad 0 = \ddot{\phi} + 3H\dot{\phi} + \frac{d}{d\phi}V(\phi)$$

- Mukhanov variable $v = z\mathcal{R}$ ($z = a\dot{\phi}/H$) evolves according to:

$$0 = v'' + \left(k^2 - \frac{z''}{z}\right)v \quad \Leftrightarrow \quad 0 = \mathcal{R}'' + 2\frac{z'}{z}\mathcal{R}' + k^2\mathcal{R}$$

  In terms of cosmic time t:

$$0 = \ddot{\mathcal{R}} + \left(2\frac{z'}{z} + H\right)\dot{\mathcal{R}} + k^2\mathcal{R}$$

- During inflation $\eta \to \eta_*$ friction term $\to \infty$, mode freezes.
- Primordial power spectrum is $P_\mathcal{R}(k) \propto \lim_{\eta\to\eta_*} |\mathcal{R}|^2$
- Initial conditions for evolution set by quantum mechanics (more later)

Mukhanov Sazaki equation in curved universes

- In flat space:

$$H^2 = \frac{1}{3}\left(\frac{1}{2}\dot{\phi}^2 + V(\phi)\right), \qquad 0 = \ddot{\phi} + 3H\dot{\phi} + \frac{d}{d\phi}V(\phi), \qquad z = \frac{a\dot{\phi}}{H}$$

$$0 = \ddot{\mathcal{R}} + \left(2\frac{z'}{z} + H\right)\dot{\mathcal{R}} + k^2\mathcal{R}$$

- With curvature ($K = -1, 0, +1$):

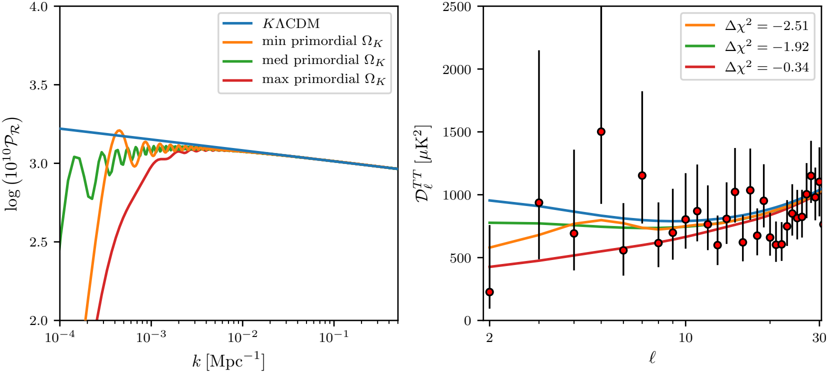
$$H^2 = \frac{1}{3}\left(\frac{1}{2}\dot{\phi}^2 + V(\phi)\right) - \frac{K}{a^2}, \qquad 0 = \ddot{\phi} + 3H\dot{\phi} + \frac{d}{d\phi}V(\phi) \qquad z = \frac{a\dot{\phi}}{H}$$

$$0 = \ddot{\mathcal{R}} + \frac{\left(H + \frac{2\dot{z}}{z}\right)\mathcal{D}^2 - \frac{3KHz^2}{2a^2}}{\mathcal{D}^2 - \frac{Kz^2}{2a^2}}\dot{\mathcal{R}} + \frac{K\left(1 + \frac{z^2}{2a^2} - \frac{2}{H}\frac{\dot{z}}{z}\right)\mathcal{D}^2 + \frac{K^2z^2}{2a^2} - \mathcal{D}^4}{a^2(\mathcal{D}^2 - \frac{Kz^2}{2a^2})}\mathcal{R}$$

$$\mathcal{D}^2 = \nabla^2 + 3K = -k^2 + 3K$$

- arXiv:1907.08524

Primordial power spectrum



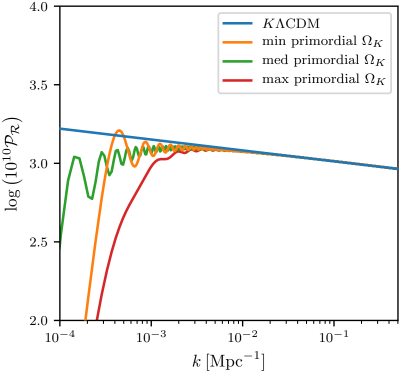
Initial conditions in curved cosmologies

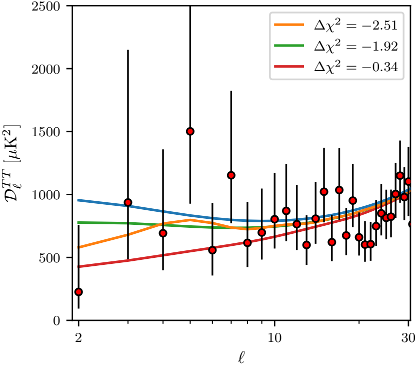
- In the flat case, usuallyinitial conditions via "Bunch Davies vacuum" to initialise Mukhanov evolution:

$$v = \frac{1}{\sqrt{2k}} e^{-ik\eta}$$

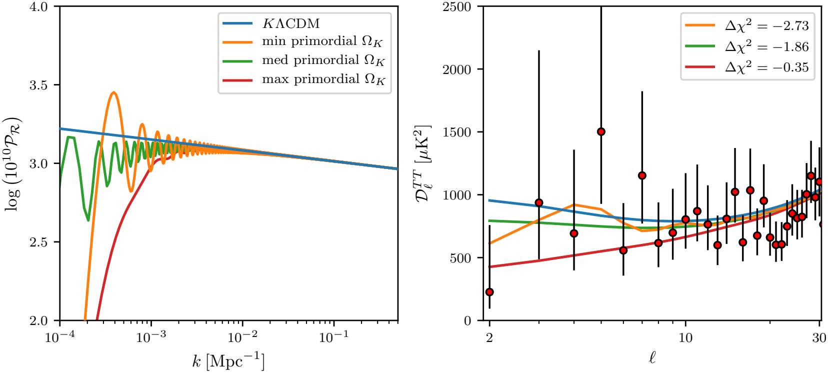
- These are only valid for modes deep in the horizon $k \gg aH$, where the spacetime is de-Sitter (effectively Minkowski)
- In the curved case, the large spatial modes have $k \sim aH$
- Need new theory of quantum initial conditions for rapidly changing spacetimes.
- Instead of defining vacuum as "particle-less" (a relative concept), could define it in terms of "lowest energy"
- Minimising the renormalised stress-energy yields observationally distinct initial conditions arXiv:1607.04148.

Bunch Davies initial conditions





Renormalised stress-energy initial conditions



Conclusions

- The suspiciousness S has become one of the standards for measuring tensions multiple simultaneous parameters.
- Significant tensions remain for cosmologies that involve curvature, comparable in scale with other tensions that are discussed.
- Curvature+inflation yields a rich set of theoretical problems with potentially observable consequences.

Extra slides

Bayesian evidence $\mathcal{Z}$: Prior dependency

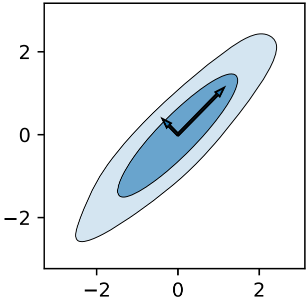
- Bayesian evidences are prior dependent:

$$\mathcal{Z} = \int P(D|\theta)P(\theta)d\theta \approx \langle \text{Likelihood} \rangle_{\text{Posterior}} \times \frac{\text{Posterior volume}}{\text{Prior volume}}$$

- They balance "goodness of fit" via likelihood with "complexity" through Occam penalty.
- Models that include too many fine-tuned parameters are disfavoured, unless they provide a much better fit.
- Corollary: Unconstrained parameters are not penalised.
- Widen prior $\Rightarrow$ reduce evidence (providing prior does not cut into posterior).
- Bayesians vs Frequentists $\leftrightarrow$ Feature vs Bug.

The problem with Principle Component Analysis

- Compute eigenvectors and eigenvalues of covariance matrix.
- These aim to describe "directions" in parameter space
- This procedure is not covariant:



"Principal Component Analysis" is a dimensionally invalid method that gives people a delusion that they are doing something useful with their data. If you change the units that one of the variables is measured in, it will change all the "principal components"! It's for that reason that I made no mention of PCA in my book. I am not a slavish conformist, regurgitating whatever other people think should be taught. I think before I teach. David J C MacKay.