Bayesian model selection in cosmology

Martin Kilbinger
Institut d'Astrophysique de Paris (IAP)

Shanghai Normal University
3 August 2009

Model selection

- Traditional parameter estimation:
  Q: For a specific model with n parameters which is the most likely (best-fit) parameter and confidence interval given the data?

- Model selection:
  Q: Which of two or more models with parameters n_1, n_2, ... is the most likely to fit the data?

- Examples in cosmology:

  ★ Cosmological constant Λ vs. dark energy vs. modified gravity

  ★ Flat vs. curved

  ★ Primordial fluctuations: scale-free (n_s=1) vs. n_s=const vs. running $n_s(k)$

- Other applications: Cluster profile reconstruction, exo-planets, nuisance parameters, ...

Bayesian evidence

- Bayes' theorem

Likelihood Prior

$$p(\theta|d,m) = \frac{\mathcal{L}(d|\theta,m)\pi(\theta|m)}{E(d|m)}$$

Posterior Evidence

$$\begin{aligned} m &: \quad \text{model} \\ d &: \quad \text{data} \\ \theta &: \quad \text{model parameter} \end{aligned}$$

- Posterior is normalised

$$E(d|m) = \int \mathrm{d}^n\theta\, \mathcal{L}(d|\theta,m)\, \pi(\theta|m)$$

- Bayes factor

$$B_{12} = \frac{E(d|m_1)}{E(d|m_2)}$$

Posterior odds

How can we interpret the evidence?

- Bayes' theorem again:

$$p(m|d) \propto E(d|m)\pi(m)$$

- Ratio for two models m_1 and m_2:

$$\frac{p(m_1|d)}{p(m_2|d)} = B_{12}\frac{\pi(m_1)}{\pi(m_2)}$$

Posterior odds

Jeffreys' scale

$\lvert\log_{10} B_{12}\rvert$	Odds	Strength
0 ... 0.5	1 ... 3	weak
0.5 ... 1	3 ... 10	substantial
1 ... 2	10 ... 100	strong
>2	>100	decisive

Approximations to the Evidence

- BIC (Bayesian Information Criterion) [Schwarz 1987]

$$\mathrm{BIC} = -2\ln\mathcal{L}_{\mathrm{max}} + k\ln N_{\mathrm{data}}$$

likelihood @ maximum

penalty for large parameter space (Occam's razor)

(Similar: AIC, DIC)
Problem: Penalty independent whether parameters constrained by data or not

- Laplace approximation: likelihood Gaussian, priors large and uniform

[Lazarides, Ruiz de Austri & Trotta 2004, Heavens, Kitching & Verde 2007]

$$E \approx (2\pi)^{n/2}|F|^{-1/2}(\Delta\theta_1\ldots\Delta\theta_n)^{-1}\mathcal{L}_{\mathrm{max}}$$

$$\underbrace{\text{volume allowed by data} \quad \text{initial volume (prior)}}_{\text{Occam's razor}}$$

Laplace: prios might be small (physical parameter boundaries)

Fisher matrix not good approx.

CMB+SNIa+BAO: Dark energy, curvature

- Base model: flat ΛCDM, n_{par}=3 $(\Omega_\mathrm{m}, \Omega_\mathrm{b}, h)$
  Data: pure geometrical probes (WMAP5 distance priors, SNIa, BAO)

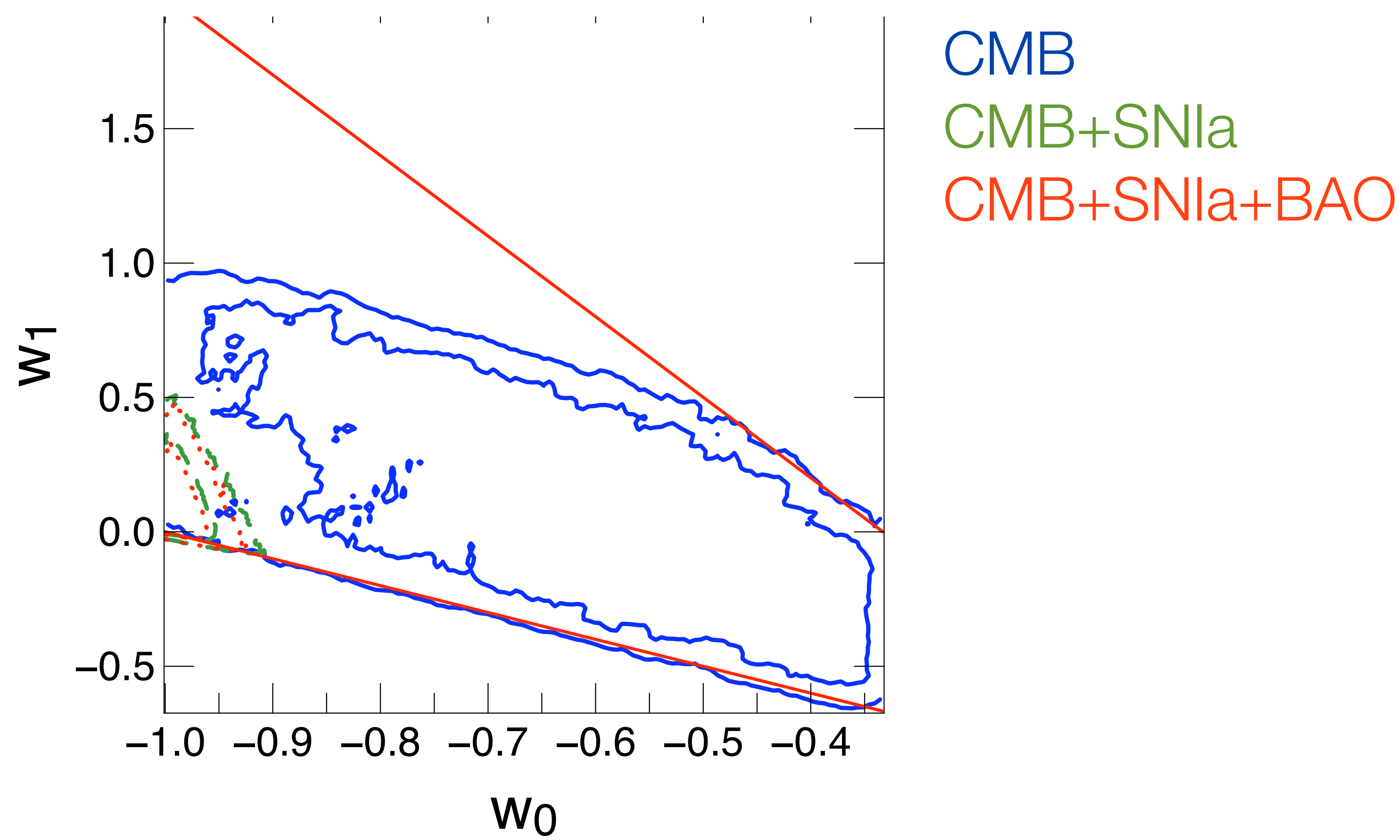
- Three models of dark energy:

$w = -1$	ΛCDM
$w = w_0$	wCDM
$w = w_0 + w_1(1-a)$	$w(z)$CDM

- Flat (Ω_K=0) vs. curved ($\Omega_K \neq 0$)

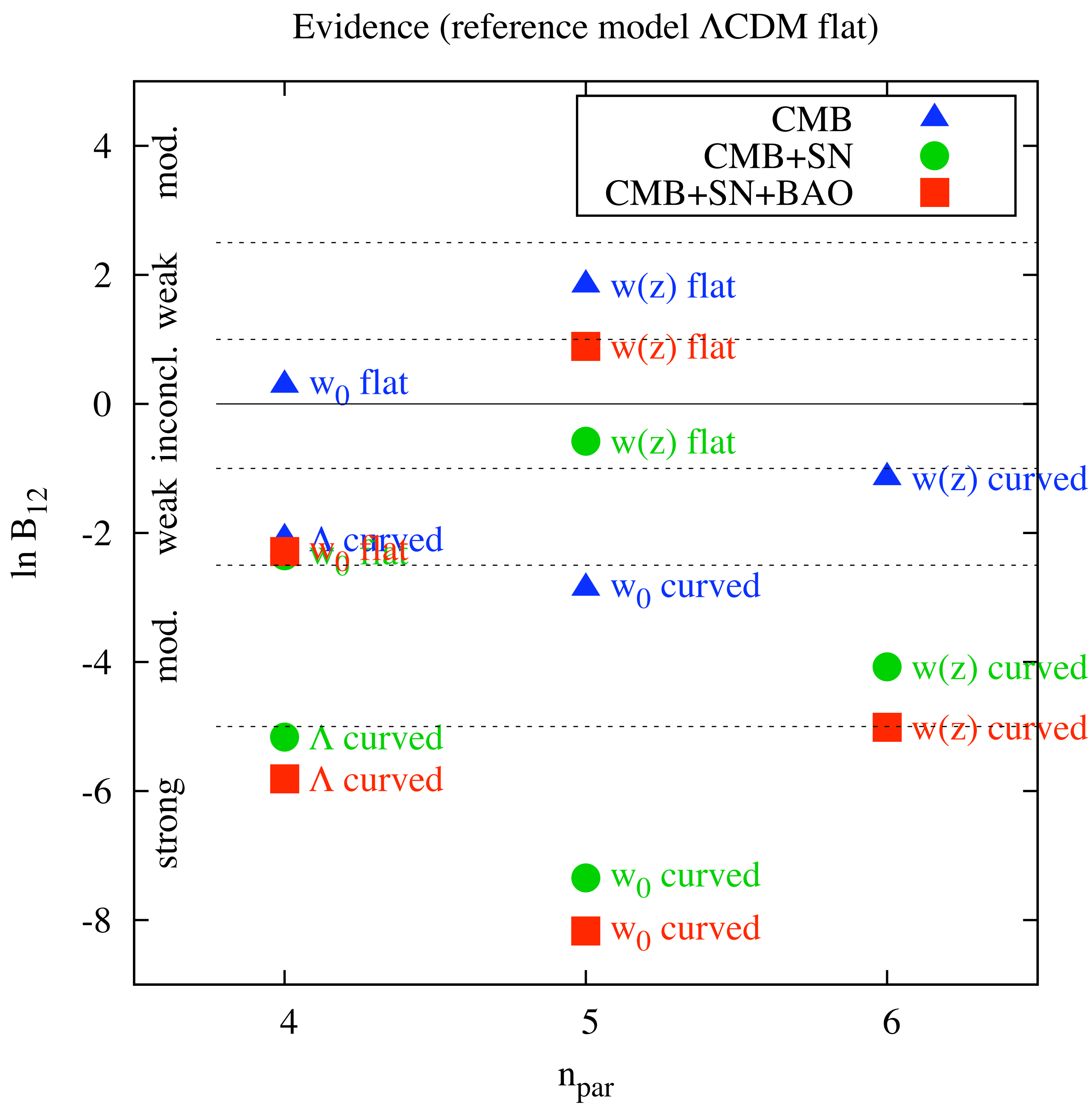
Priors

- Exclude phantom energy ($w \geq -1$), require accelerated expansion for $z < z_{acc} = 0.5$ ($w < -1/3$)



- Upper boundary for curvature: empty Universe, $\Omega_K < 1$, lower boundary $\Omega_K > -1$

Bayes’ factor: dark energy, curvature



[MK, Wraith et al. in prep]

Priomordial Universe

- Density perturbation power spectrum

$$P_\delta \propto k^{n_\mathrm{s}+1/2\alpha_\mathrm{s}\ln(k/k_0)}$$

n_s: spectral index (n_s =1: scale-free, Harrison-Zel'dovich spectrum)
α_s: 'running' of spectral index

- Tensor perturbation (gravitational waves) power spectrum

$$P_\mathrm{t} \propto k^{n_\mathrm{t}}$$

- Standard model has *n*s = const, α_s = n_t = 0, r=tensor/scalar=0

Priors

- Slow-roll approximation of inflation

$$\epsilon = \frac{m_{\rm Pl}^2}{4\pi}\left[\frac{H'}{H}\right]^2;$$

$$^{\ell}\lambda_H = \left(\frac{m_{\rm Pl}^2}{4\pi}\right)^{\ell}\frac{(H')^{\ell-1}}{H^{\ell}}\frac{\mathrm{d}^{\ell+1}H}{\mathrm{d}\phi^{\ell+1}}; \ell \leq 1,$$

  depend on potential *V* of scalar field Φ causing inflation.

- Slow-roll condition: $\epsilon \ll 1, \quad |^{\ell}\lambda_H| \ll 1$

- Choose: $\varepsilon \leq 0.1, \quad -0.1 < {}^{1}\lambda_H = \eta < 0.1, {}^{\ell}\lambda_H = 0 \text{ for } \ell > 1$

Priors

- Relations between slow-roll and power-spectrum parameters

$$n_{\rm s} = 1 + 2\eta - 4\epsilon - 2(1+\mathcal{C})\epsilon^2 - \frac{1}{2}(3-5\mathcal{C})\epsilon\eta;$$

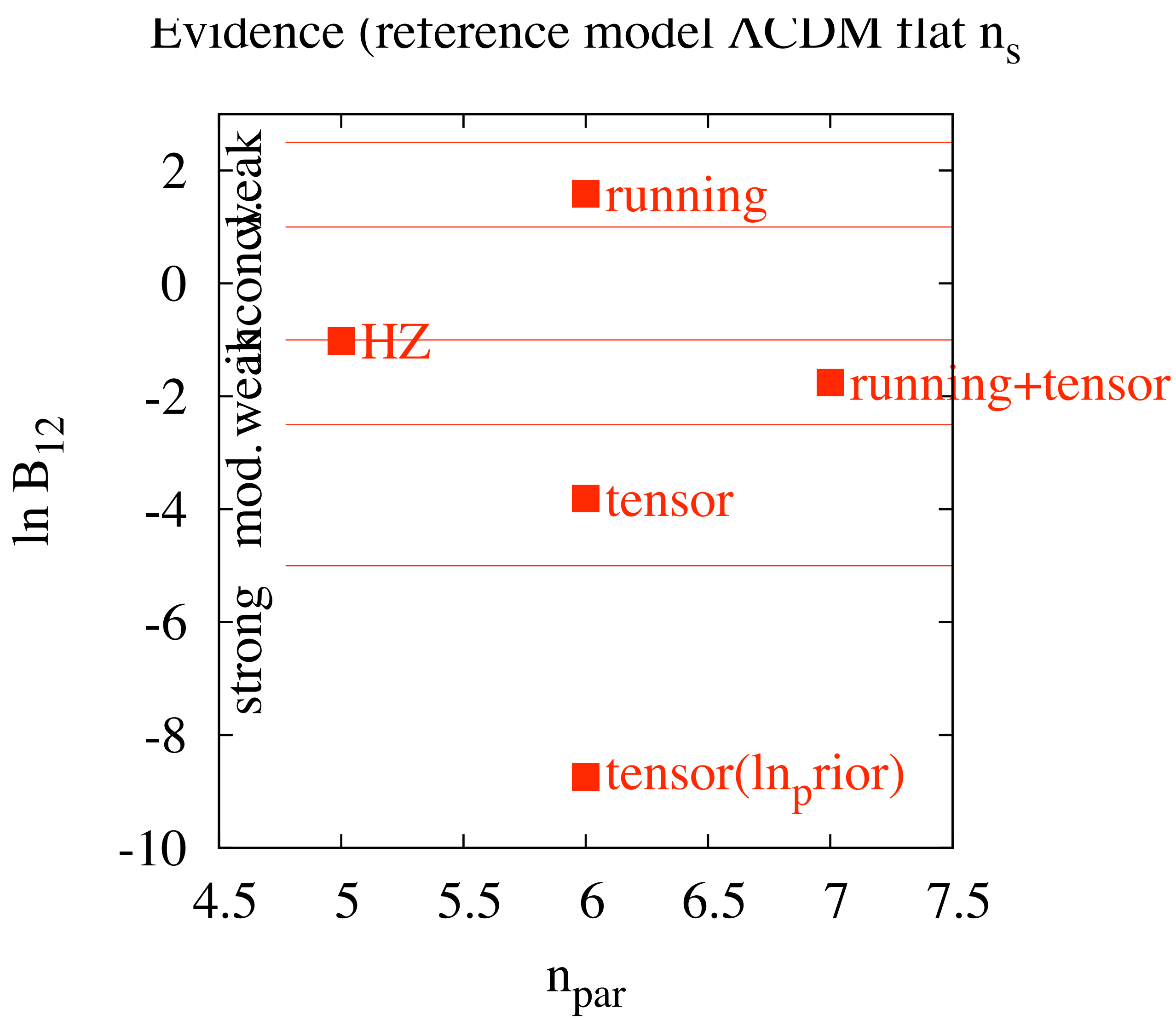
$$r = 16\epsilon\left[1 + 2\mathcal{C}(\epsilon - \eta)\right];$$

$$\alpha_{\rm s} = \frac{\epsilon}{1-\epsilon}\left(8\epsilon + 10\eta\right);$$

$$n_{\rm t} = -2\epsilon - (3+\mathcal{C})\epsilon^2 + (1+\mathcal{C})\epsilon\eta.$$

$$\mathcal{C} = 4(\ln 2 + \gamma) - 5 \approx -0.7296$$

Bayes' factor (CMB+SNIa+BAO)



	Parameter	Minimum	Maximum
linear	r	0	1.83
logarithmic	$(\ln r$	-80	0.604)

Constraints (CMB+SNIa+BAO)