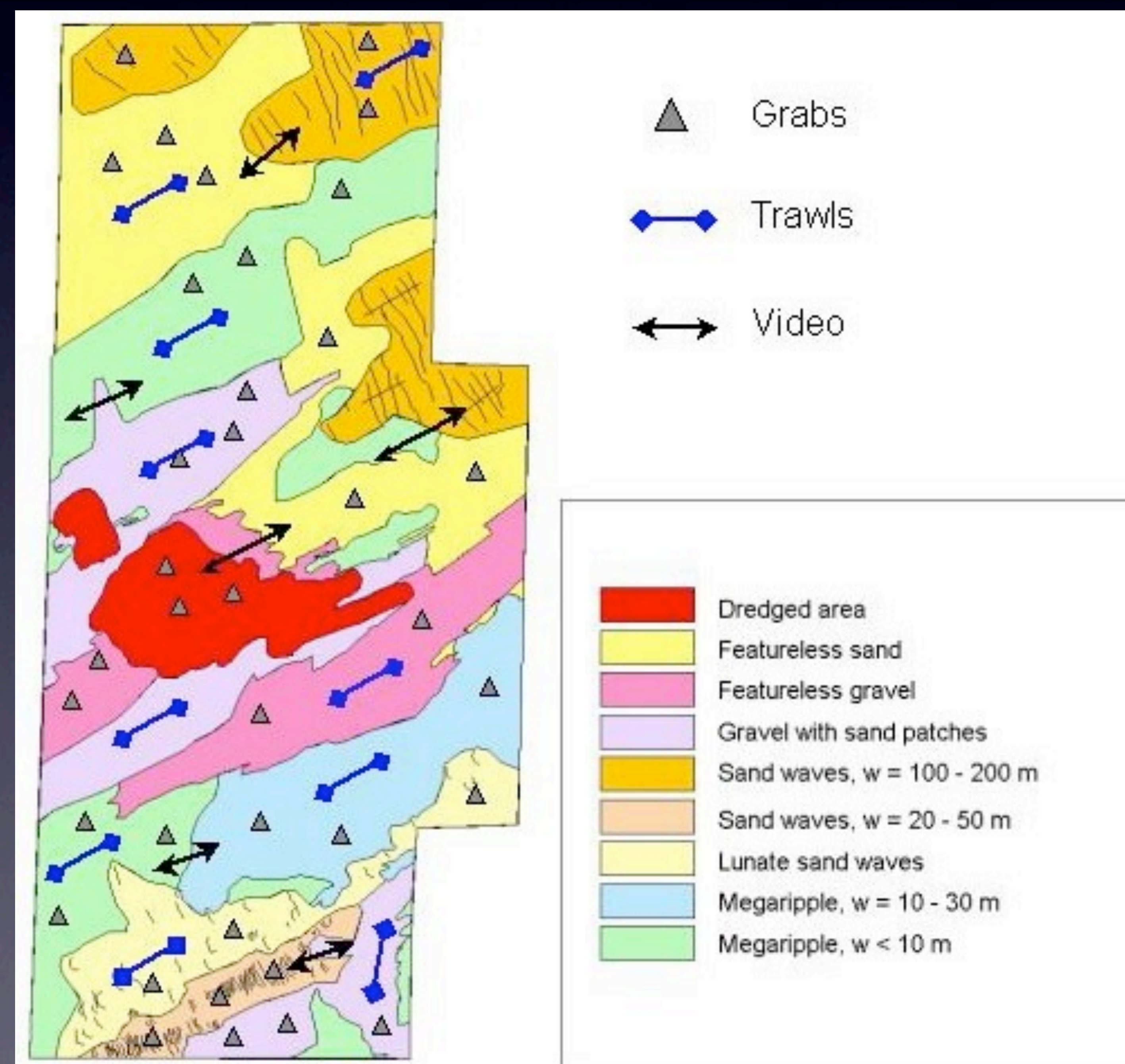
Forecasting and experimental design

University of Sussex

Possible survey design for ground-truth sampling of a draft physical map interpreted from a sidescan sonar surveys. The area is approximately 4 x 10 km.

MESH MAPPING EUROPEAN SEABED HABITATS



Microsoft-free presentation

Experimental forecasting

Forecasting is the quantification of how well future experiments can answer particular science questions. This might, for instance, be interesting to funding agencies that wish to compare competing proposals.

This is achieved by defining Figures of Merit that can be associated to each proposed experiment. There are a variety of different options that can be used, some based on parameter estimation and some on model selection.

First, we ought to decide what question we hope to answer.

Some possible questions

- I’m happy with my model, which features a parameter θ. How well will my experiment measure θ?
- I’ve got a few uncertain parameters, and I’d like to learn more about all of them. How well will I do with this experiment?

- I’m not sure whether effect A is relevant to my data, but if it is, I need to include a new parameter θ_A. Will my data be able to confirm that I need to include this parameter?
- Or, alternatively, suppose I suspect that effect A is irrelevant. Will the data be able to confirm my suspicion, if it is correct?
- I’ve got two competing models to explain my data. Will I be able to exclude one of them?

FoMs: parameter estimation

- I'm happy with my model, which features a parameter θ. How well will my experiment measure θ?
- I've got a few uncertain parameters, and I'd like to learn more about all of them. How will I do?

Here obviously we want to predict the size of the uncertainties we will obtain. I assume that we have some way of modelling the proposed experiment which yields an estimate of the expected likelihood (otherwise forecasting is impossible).

Fisher matrix approach

- Taylor expand the log likelihood around the Maximum Likelihood parameter values (θ^{ML}):

$$\ln L(\theta) \approx \ln L(\theta^{ML}) + \tfrac{1}{2}\sum_{ij}(\theta_i - \theta_i^{ML})^t H_{ij}(\theta_j - \theta_j^{ML})$$

$$H_{ij} = \left.\frac{\partial \ln L}{\partial \theta_i}\frac{\partial \ln L}{\partial \theta_j}\right|_{\theta^{ML}}$$

- Taking the expectation of *L* over many data realizations, we replace the Maximum Likelihood with the fiducial parameter value.
- The Fisher matrix is defined as the expectation of the Hessian:

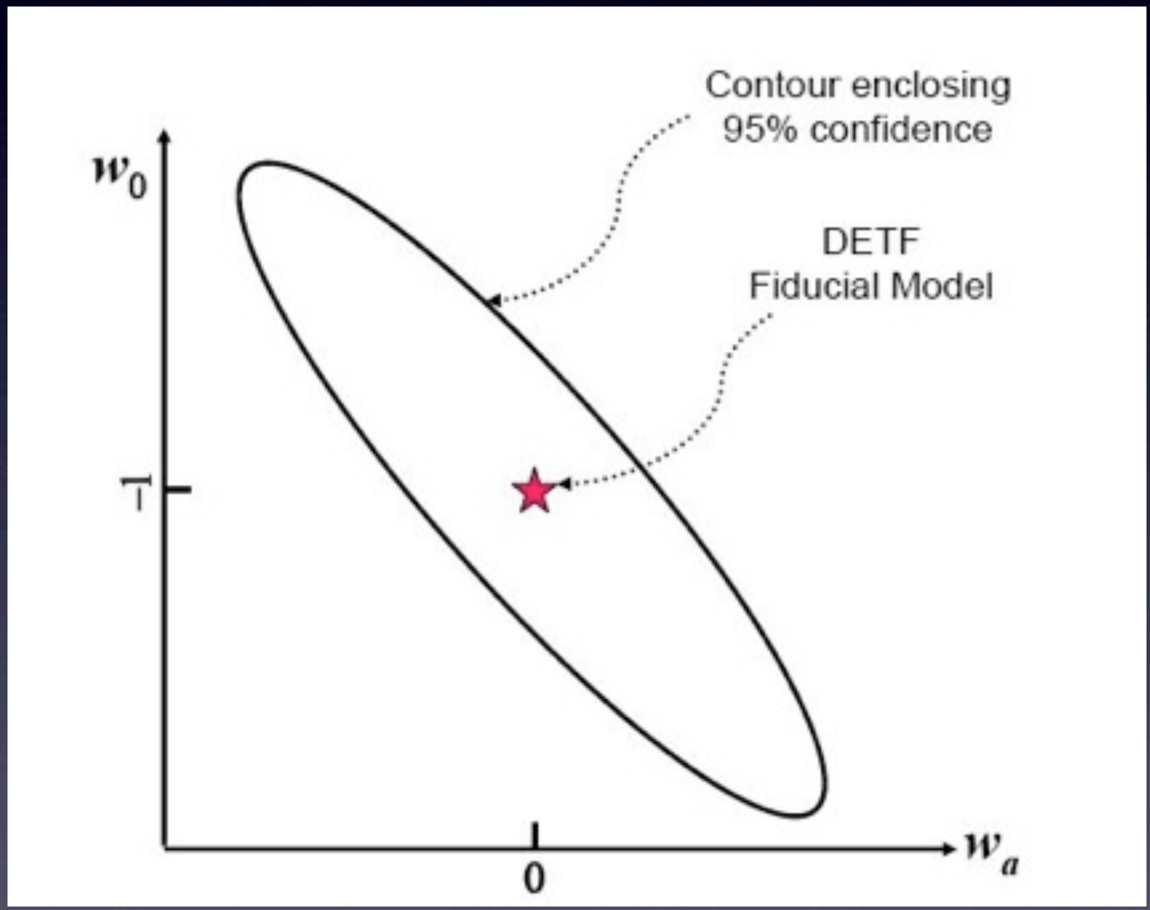
$$F_{ij} \equiv \langle H_{ij} \rangle$$

Figures of Merit (aka Utility)

- The determinant of the Fisher matrix, $|F|$ (often called D–optimality), which is inversely proportional to the square of the parameter volume enclosed by the posterior.
  - A common variation is to use the logarithm of the determinant, $\ln |F|$.
- The trace of the Fisher matrix, tr$F$, or its logarithm: this is proportional to the sum of the variances, and is often called A–optimality.
- The information gain H from performing the experiment (also often called Kullback-Leibler divergence), between the prior and posterior.

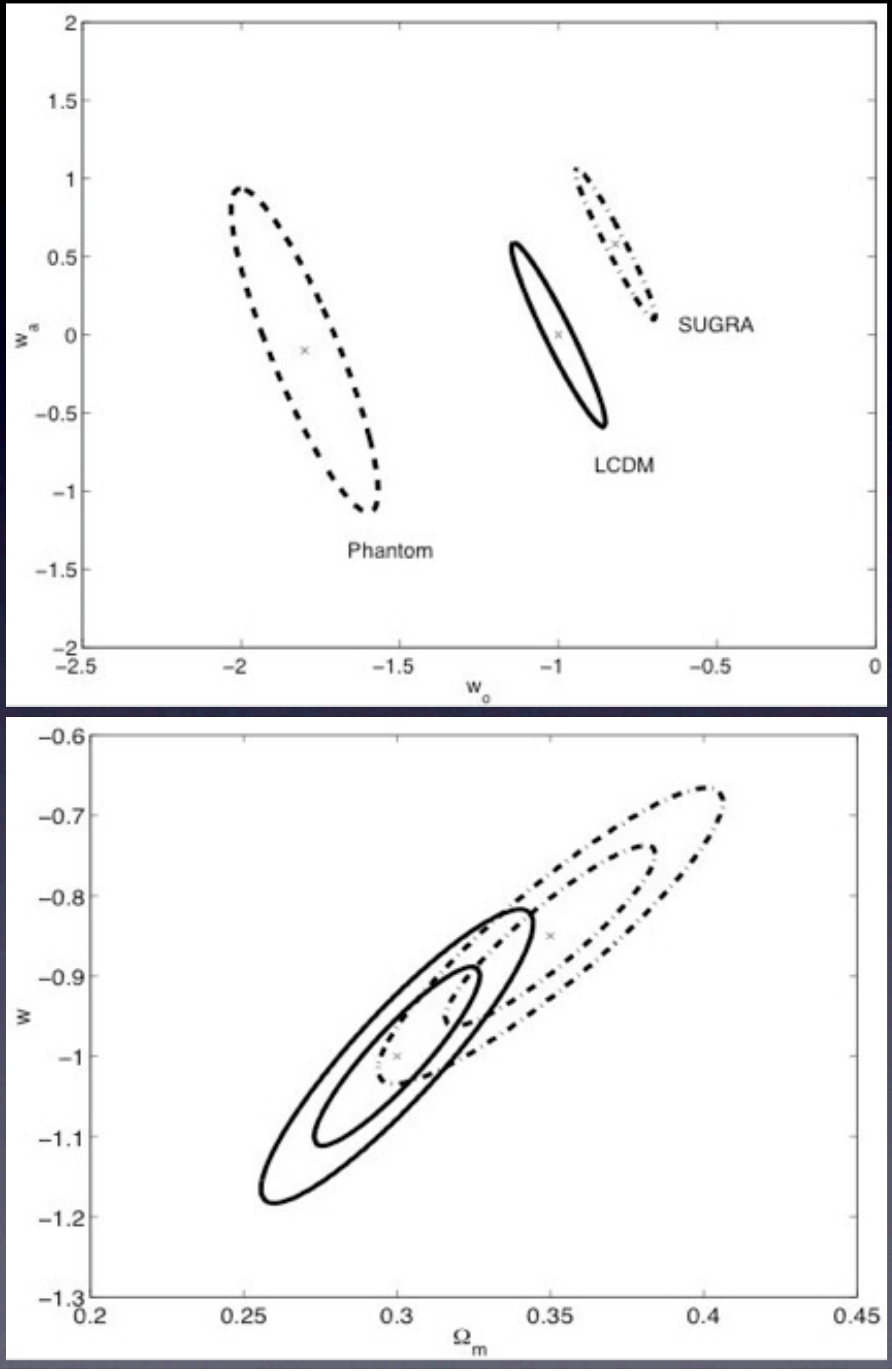
Dark Energy Task Force Figure of Merit

- Constraining the equation of state, w, and its evolution in time is seen as the primary goal.
- The DETF created a Figure of Merit to compare different surveys and approaches (Albrecht et al. 2006).
- It is the inverse of the 95% confidence contour in the w_0, w_a plane (D-optimal).
- Often quoted as $[\sigma(w_a) \times \sigma(w_p)]$-1, which is in fact sqrt(det[F_{DE}]) where [F_{DE}] is the marginalized 2×2 Fisher matrix for the dark energy parameters w_0 and w_a.



Effectiveness

- The errors on *w* (and hence the FoM) of a survey depends on the fiducial cosmology.
- And even the conclusions that you draw from the data may change with the cosmology.



FoMs: model selection

- I'm not sure whether effect A is relevant to my data, but if it is, I need to include a new parameter θ_A. Will my data be able to confirm that I need to include this parameter?
- Or, alternatively, suppose I suspect that effect A is irrelevant. Will the data be able to confirm my suspicion, if it is correct?
- I've got two competing models to explain my data. Will I be able to exclude one of them?

To explore these questions, we have to decide which model is to be taken as true, simulate observational data (possibly across its parameter space), and assess the value of the Bayesian evidence that our experiment will generate.

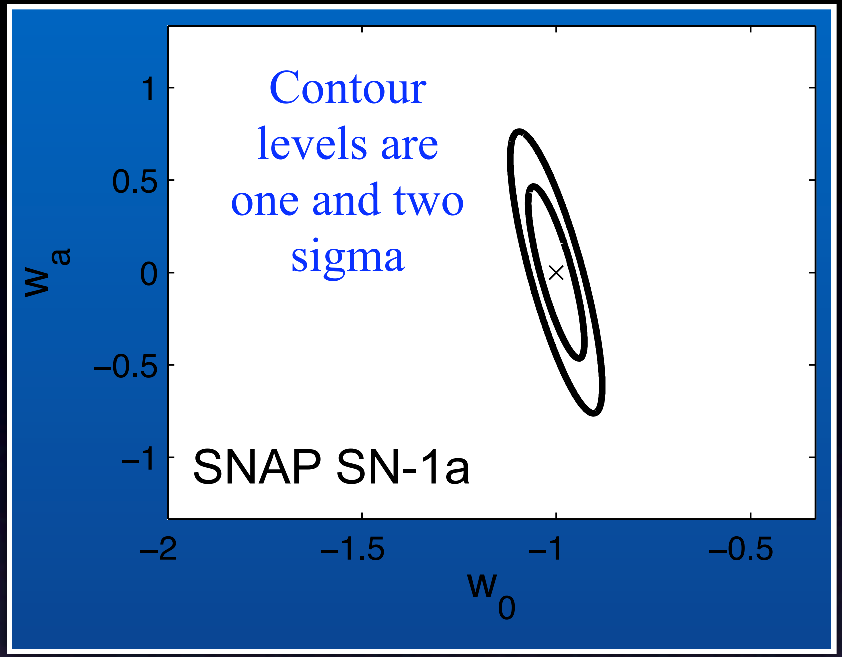
Mukherjee, Parkinson, Corasaniti, Liddle, Kunz, MNRAS, astro-ph/0512484

Forecasts for dark energy

$$w = w_0 + (1 - a) w_a$$

Parameter estimation question:

Suppose dark energy is described by a two-parameter model with $w_0 = -1$ and $w_a = 0$. How tight do I expect my constraints on those parameters to be?

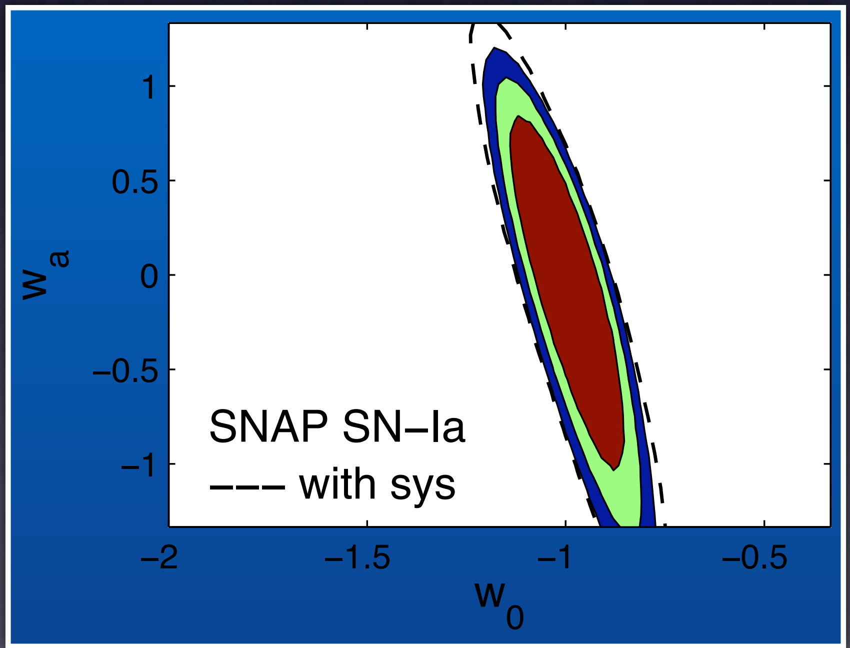


Model selection questions:

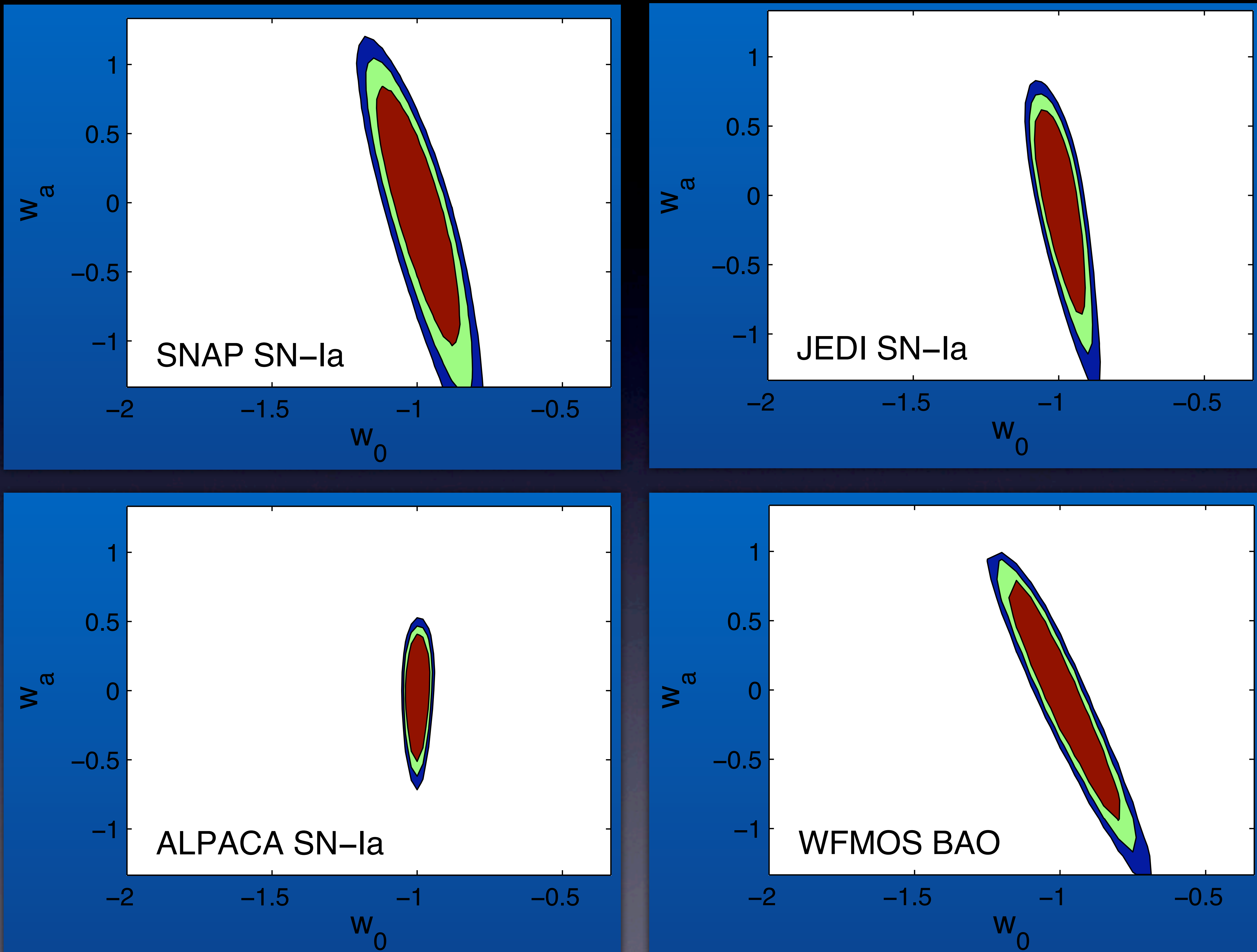
If the dark energy model is right, will my experiment support it over ΛCDM?

If it turns out that ΛCDM is right, is my experiment good enough to exclude the evolving dark energy model?

If ΛCDM is excluded, can I distinguish between quintessence and modified gravity models?



Red: Λ mildly favoured
Green/blue: indecisive
White: DE favoured


SNAP SN−Ia
JEDI SN−Ia
ALPACA SN−Ia
WFMOS BAO
w_0
w_a

(Almost) current dark energy data

Liddle, Mukherjee, Parkinson, and Wang, PRD, astro-ph/0610126

CMB shift+BAO(SDSS)+SN

LambdaCDM

Constant W

W_0-W_a

data used WMAP+SDSS+	$\Delta \ln E$	H	$\chi^2_{\rm min}$	Model parameter constraints
				Model I: Λ
Riess04	0.0	5.7	30.5	$\Omega_{\rm m} = 0.26 \pm 0.03,\ H_0 = 65.5 \pm 1.0$
Astier05	0.0	6.5	94.5	$\Omega_{\rm m} = 0.25 \pm 0.03,\ H_0 = 70.3 \pm 1.0$
				Model II: constant w, flat prior $-1 \le w \le -0.33$
Riess04	-0.1 ± 0.1	6.4	28.6	$\Omega_{\rm m} = 0.27 \pm 0.04,\ H_0 = 64.0 \pm 1.4,\ w < -0.81, -0.70^a$
Astier05	-1.3 ± 0.1	8.0	93.3	$\Omega_{\rm m} = 0.24 \pm 0.03,\ H_0 = 69.8 \pm 1.0,\ w < -0.90, -0.83^a$
				Model III: constant w, flat prior $-2 \le w \le -0.33$
Riess04	-1.0 ± 0.1	7.3	28.6	$\Omega_{\rm m} = 0.27 \pm 0.04,\ H_0 = 64.0 \pm 1.5,\ w = -0.87 \pm 0.1$
Astier05	-1.8 ± 0.1	8.2	93.3	$\Omega_{\rm m} = 0.25 \pm 0.03,\ H_0 = 70.0 \pm 1.0,\ w = -0.96 \pm 0.08$
				Model IV: w_0–w_a, flat prior $-2 \le w_0 \le -0.33,\ -1.33 \le w_a \le 1.33$
Riess04	-1.1 ± 0.1	7.2	28.5	$\Omega_{\rm m} = 0.27 \pm 0.04,\ H_0 = 64.1 \pm 1.5,\ w_0 = -0.83 \pm 0.20,\ w_a = --^b$
Astier05	-2.0 ± 0.1	8.2	93.3	$\Omega_{\rm m} = 0.25 \pm 0.03,\ H_0 = 70.0 \pm 1.0,\ w_0 = -0.97 \pm 0.18,\ w_a = --^b$
				Model V: w_0–w_a, $-1 \le w(a) \le 1$ for $0 \le z \le 2$
Riess04	-2.4 ± 0.1	9.1	28.5	$\Omega_{\rm m} = 0.28 \pm 0.04,\ H_0 = 63.6 \pm 1.3,\ w_0 < -0.78, -0.60^a,\ w_a = -0.07 \pm 0.34$
Astier05	-4.1 ± 0.1	11.1	93.3	$\Omega_{\rm m} = 0.24 \pm 0.03,\ H_0 = 69.5 \pm 1.0,\ w_0 < -0.90, -0.80^a,\ w_a = 0.12 \pm 0.22$

Conclusion: LambdaCDM currently favoured but all models still alive

Future forecasts informed by current data

Trotta, astro-ph/0504022; Liddle, Mukherjee, Parkinson, and Wang, astro-ph/0610126

Bayesian philosophy: continual updating of probabilities as new data comes in.
⇒ Use current probabilities to forecast future experiment outcomes

- If LambdaCDM is right, are upcoming experiments (eg DES, WFMOS, SNAP) good enough to favour it decisively?
- What is the probability that upcoming experiments will robustly detect dark energy evolution?
- If future experiments are still inconclusive, how tight will be the limits they can impose on dark energy properties?

Future forecasts informed by current data

Trotta, astro-ph/0504022; Liddle, Mukherjee, Parkinson, and Wang, astro-ph/0610126

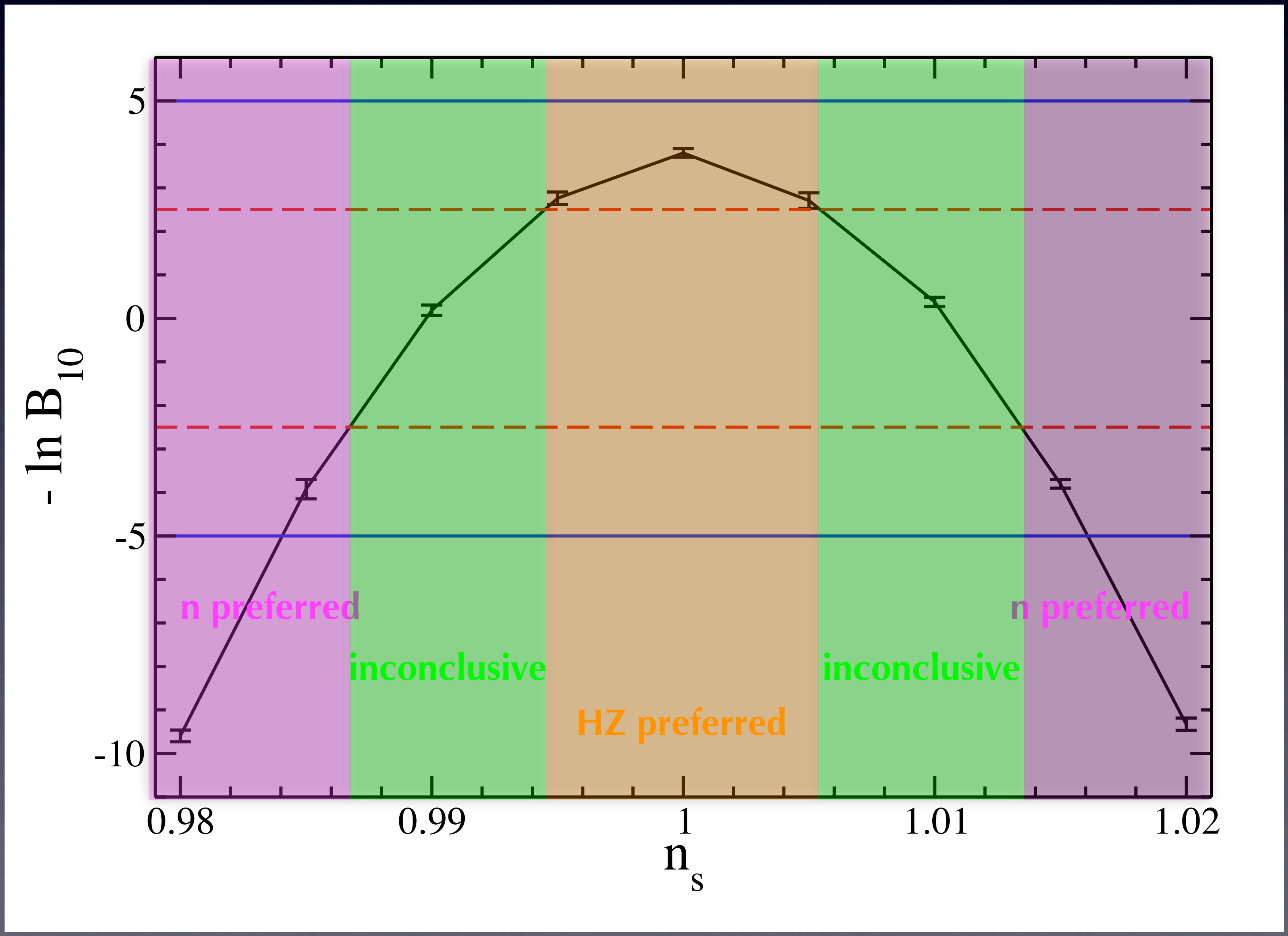
Under particular prior assumptions we made (the effect of whose variation is readily tested), the answers are ...

- If LambdaCDM is right, are upcoming experiments (eg DES, WFMOS, SNAP) good enough to favour it decisively? — YES
- What is the probability that upcoming experiments will robustly detect dark energy evolution? — About 25%
- If future experiments are still inconclusive, how tight will be the limits they can impose on dark energy properties? — Tighter than you expect!

Model selection forecasts for Planck

Pahud, Liddle, Mukherjee, and Parkinson, PRD, astro-ph/0605004

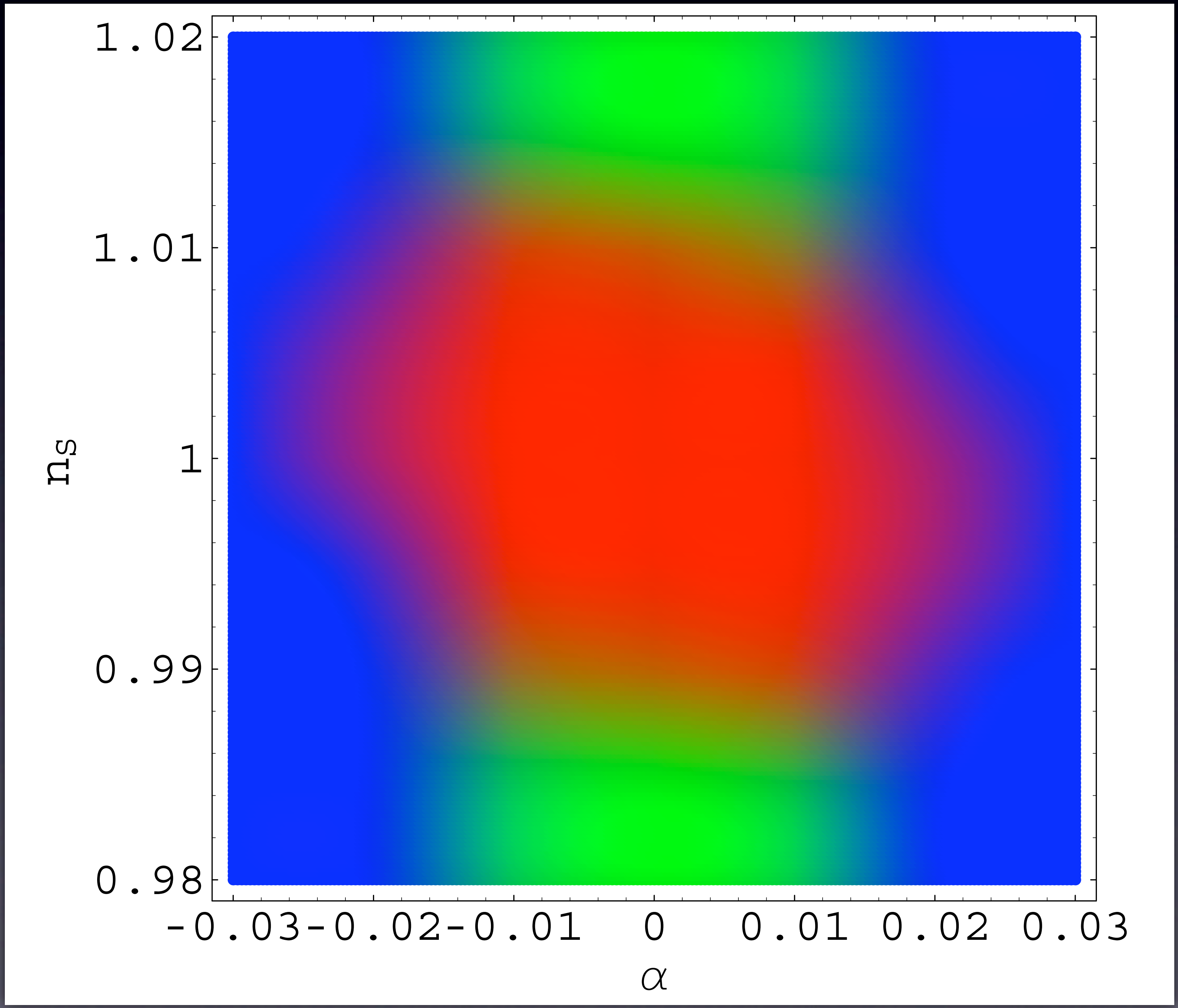
We can also do model selection forecasts for the Planck satellite. For the spectral index it looks like this:



Model selection forecasts for Planck

Pahud, Liddle, Mukherjee, and Parkinson, MNRAS, astro-ph/0701481

Or with both the spectra index n_S and running α ...



Red: HZ model preferred
Green: power-law model preferred
Blue: running model preferred

Survey optimization

Optimization refers to the tuning of surveys to maximize their ability to answer particular science questions. This might, for instance, be interesting to collaborations hoping to maximize their chance of persuading a funding agency to favour their proposal.

Rather then specify a predetermined experimental configuration, as before, we now allow the experimental setup to depend on a number of parameters, usually subject to some constraints (eg maximum permitted cost, fixed telescope observing time, etc.

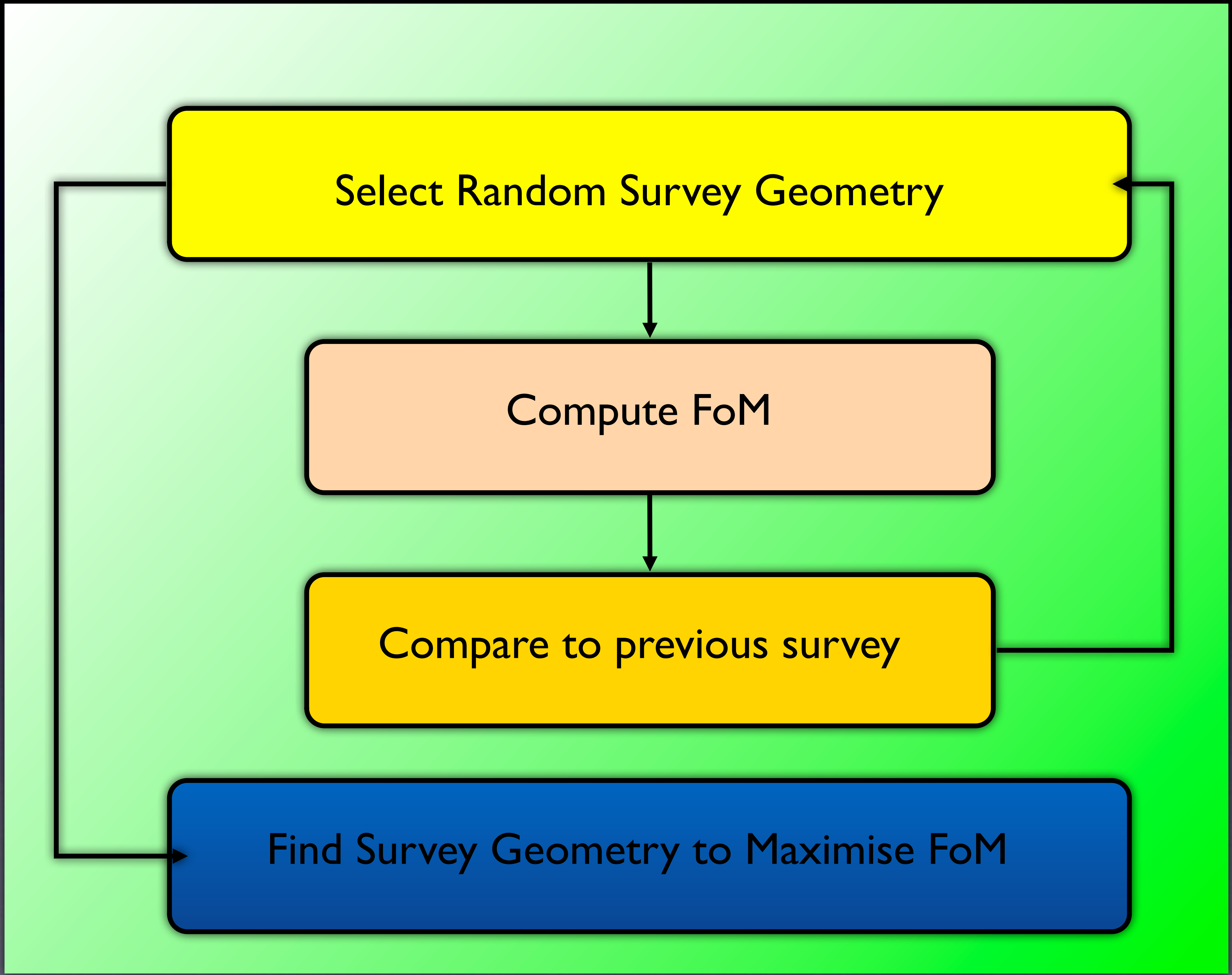
Future dark energy surveys

- Supernovae - repeated imaging with spectroscopic follow-up
  - Current: SNLS, ESSENCE, SDSS-II
  - Next gen: Pan-STARRS, DES
  - 3rd gen: LSST, JDEM
- Baryonic Acoustic Oscillations - large scale redshift survey
  - Current: WiggleZ, SDSS-II
  - Next gen: BOSS (SDSS-III), DES (photo-z), HETDEX (high-z), WFMOS, Hydrogen Sphere Survey (radio)
  - 3rd gen: LSST, JDEM, SKA (radio)
- Weak Lensing - large scale, high quality imaging survey
  - Next gen: DES, Pan-STARRS, HSC
  - 3rd gen: DUNE, JDEM, LSST

Survey design

- How do we optimize a survey to maximize its performance in constraining the dark energy?
- What survey strategy should we take; ie.
  - What type of objects should we target?
  - At which redshifts should we take measurements?
  - Should it survey a wide area at low redshift, or a small number of thin ‘pencil beam’ surveys going to a greater depth (or a mixture of the two)?
- And how do we quantify the performance of the survey?

Optimization Process



Survey optimization

We can perform the search, for instance, by Monte Carlo methods. We have a function

$$\mathbf{FoM(\alpha_i)}$$

where α_i are parameters describing the survey.

We can for instance pretend that the FoM is like a likelihood, and use something like Metropolis-Hastings. Because the FoM shape may be complex, but we are interested primarily only in the maximum, we don't have to worry about detailed balance and can use various techniques to explore around the configuration space,

e.g. a thermal annealing schedule

An example: WFMOS

Optimizing baryon acoustic oscillation surveys II: curvature, redshifts, and external datasets

David Parkinson[⋆1], Martin Kunz,[1] Andrew R. Liddle[1], Bruce A. Bassett[2,3], Robert C. Nichol[4] and Mihran Vardanyan[5]

WFMOS: Wide-field Fibre-fed Multi-Object Spectrograph

A now-defunct proposal to install a massive spectrograph, with thousands of fibres, on an 8m telescope (Gemini/Subaru).

Aim: to carry out a large galaxy redshift to use galaxy clustering (baryon acoustic oscillations) to test for dark energy evolution.

An example: WFMOS

Table 1. List of survey parameters in each redshift regime. See Parkinson et al. (2007) for detailed explanations. Note that we no longer vary the number of redshift bins, but instead divide up the redshift ranges into thin slices for the FoM calculation.

Survey Parameter	Symbol
Survey time	τ_{low}, τ_{high}
Area covered	A_{low}, A_{high}
Minimum of redshift bin	z_{low} (min), z_{high} (min)
Maximum of redshift bin	z_{low} (max), z_{high} (max)
Number of pointings	n_{p}(low), n_{p}(high)

Table 2. List of constraint parameters.

Constraint Parameter	Value
Total observing time	1500 hours
Field of view	1.5^o diameter
n_{fibres}	3000
Aperture	8m
Fibre diameter	1 arcsec
Overhead time between exposures	10 mins
Minimum exposure time	15 mins
Maximum exposure time	10 hours
Wavelength response	Priv. comm with AAO
Width of redshift slices, dz	0.05

These define the survey parameters, which can be varied, and the constraint parameters which are fixed.

An example: WFMOS

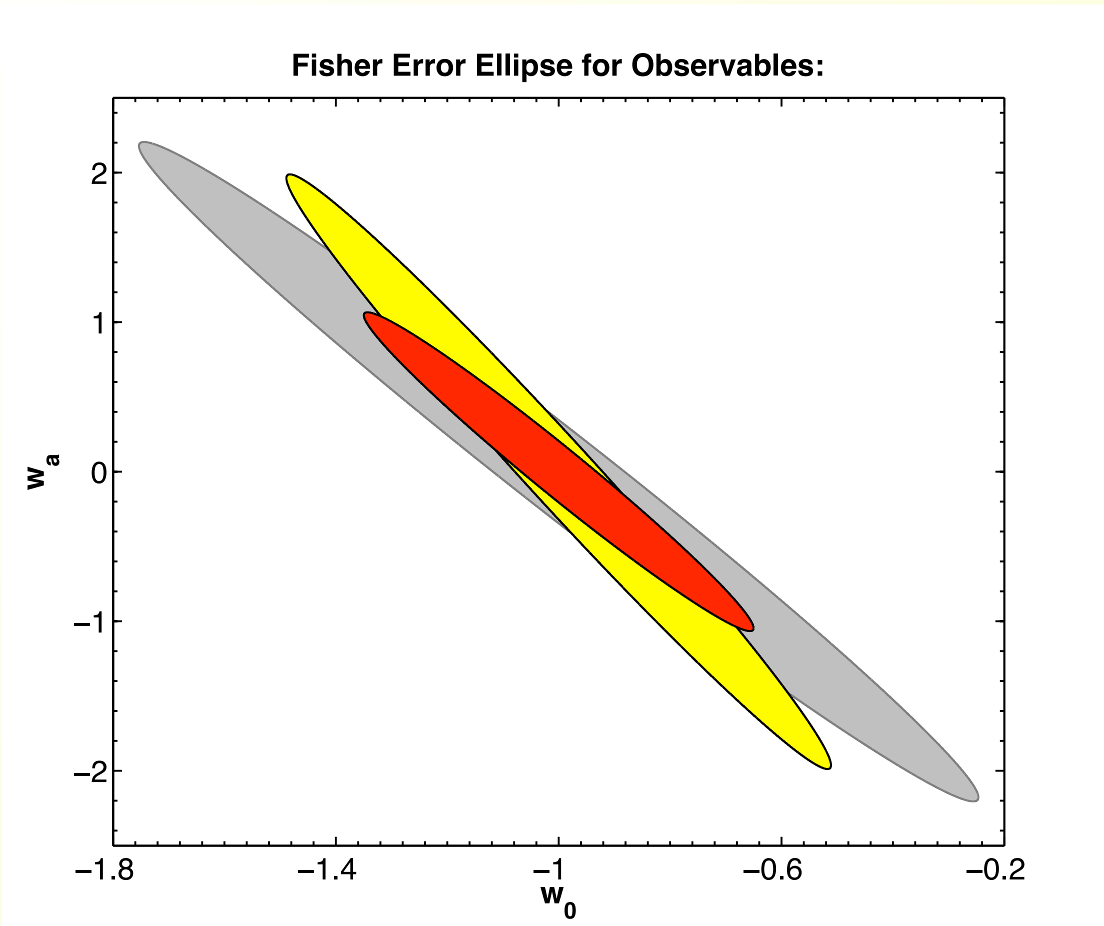


Figure 4. The 68% error ellipse on the w_0 and w_a parameters, with marginalization over curvature, for the standard WFMOS survey (grey), and the optimized one (red). Also shown (yellow) is the error ellipse were the survey optimized for a flat Universe (but the errors have been computed here marginalizing over curvature). The difference between the largest ellipse and the two smaller ones shows the improvement due to optimizing the survey for measuring the dark energy parameters, while the difference between the smaller ellipses is due to different cosmological models (flat or non-flat) used for the optimization. These constraints are calculated including prior information from Planck and SDSS.

FoMs

Unoptimized: 18

Optimized: 57

Table 6. Optimal survey Figure of Merit calculated in flat and curved cases, where the optimization has been undertaken under two different assumptions, either that Ω_k is left out or included as a nuisance parameter. The FoM in computed including prior information from Planck and SDSS.

Survey optimization	without Ω_k	with Ω_k
FoM (Ω_k set to zero)	57	48
FoM (Ω_k allowed to vary)	15	32

Conclusions

- Model selection forecasting is a powerful tool for experimental design and comparison, and is readily applied to dark energy and other experiments.
- Survey optimization offers significant potential for improved scientific and financial efficiency, certainly in an astronomical context.