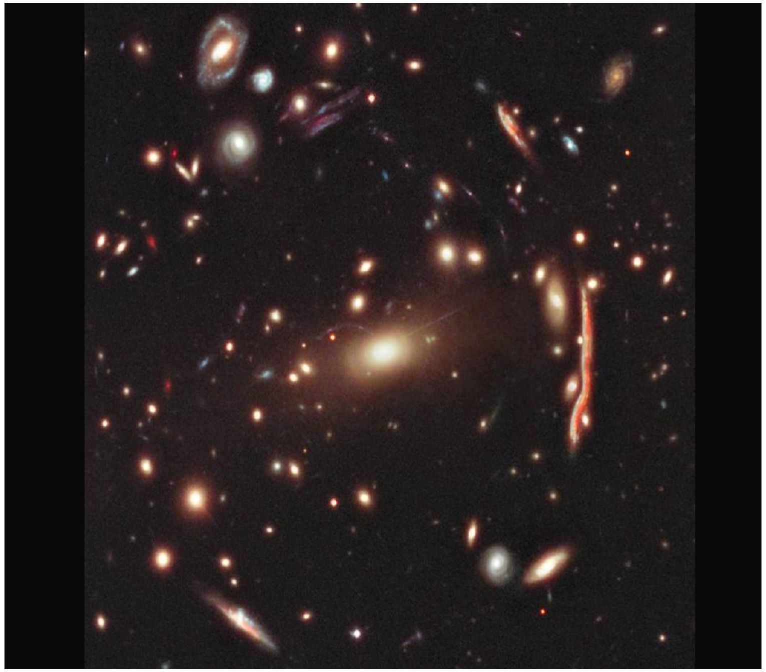
Statistical Challenges in Dark Energy Searches



CLASH MACS1206

Ofer Lahav (UCL)

Outline

- Methodology and recent trends
  in Astro-Statistics

- Examples from Dark Energy research:
  1. Cosmological parameters
  2. Photometric redshifts
  3. Galaxy classification and shapes

Modern Astro-Statistics books

- Lyons (1991)
- Press et al. (1992)
- Lupton (1993)
- Babu & Feigelson (1996)
- Sivia (1996)
- Cowan (1998)
- Starck & Murtagh (2002)
- Martinez & Saar (2002)
- Wall & Jenkins (2003)
- Saha (2003)
- Gregory (2005)
- Hobson et al. (2009)

Bayes' Theorem

- $P(A|B) = P(B|A)\,P(A) / P(B)$

- P(model | data)=
  P(data | model) P (model) / P(data)

  $\uparrow$ *Likelihood* *exp (-χ^2 /2)* $\uparrow$ *Prior* $\uparrow$ *Evidence*

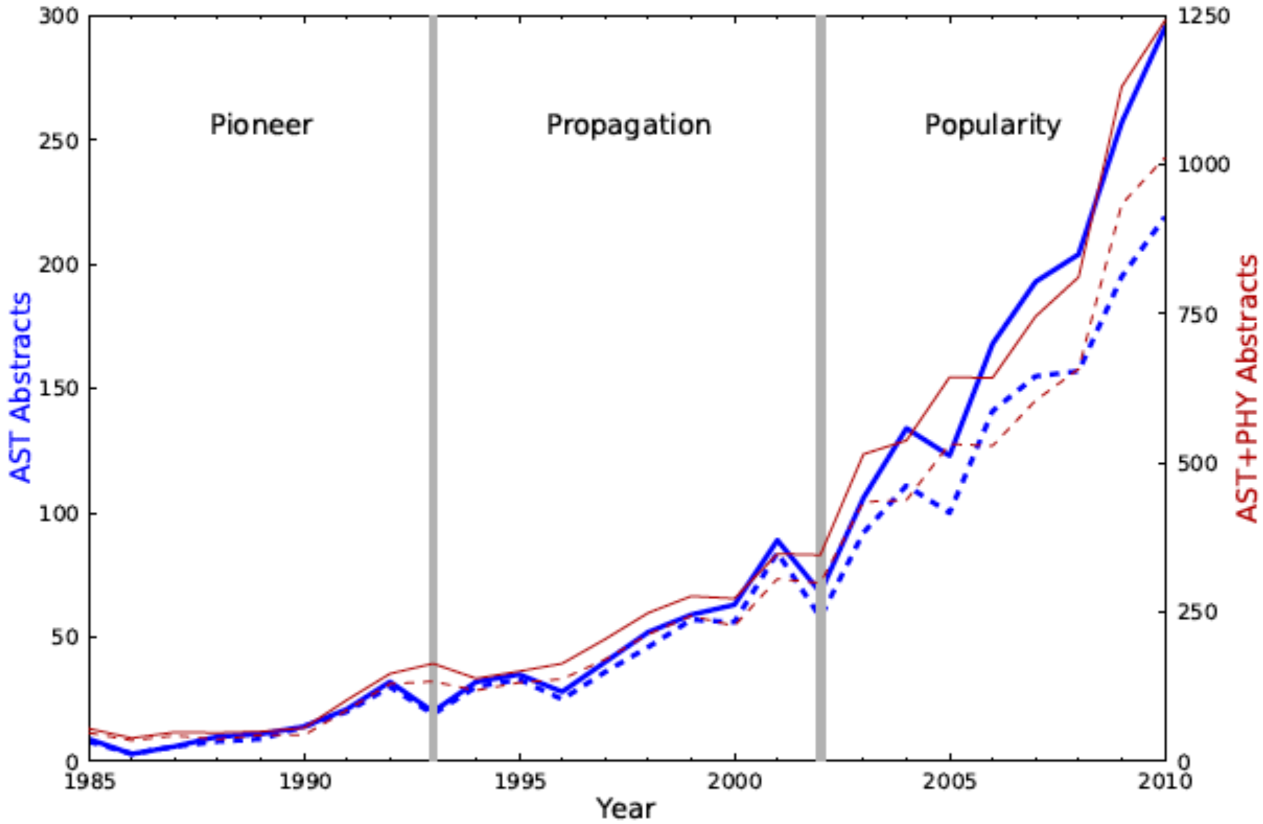
1702-1761
(paper only published in 1764)

Recent trends

- Astro-Statistics is more ‘respectable’.
- In Cosmology, Bayesian approaches are more popular (since 90s) than Frequentist methods.
- More awareness of model selection methods (e.g. Evidence, AIC, BIC, …).
- Computer intensive methods (e.g. MCMC) are more common, and free packages available.

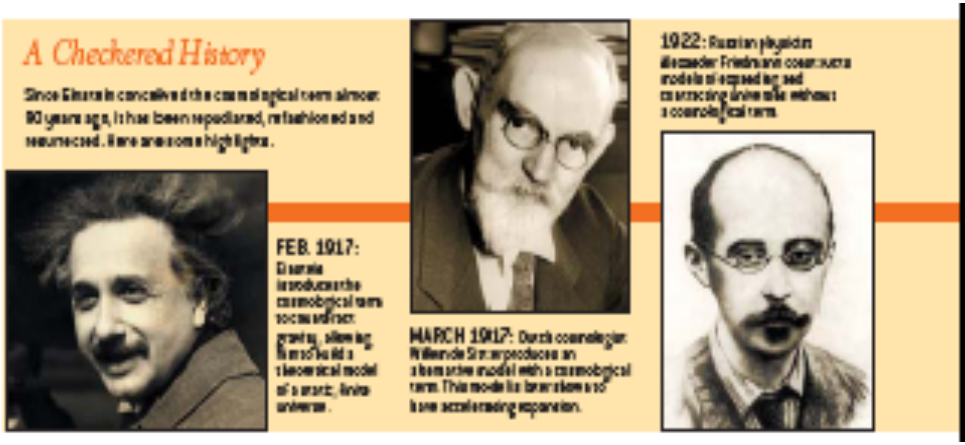
Growth of Bayesian papers

Thomas J. Loredo



Example 1:
Cosmological Parameters

The Dark Energy Paradigm



The old problem:
Theory exceeds observational limits on Λ by <u>10^{120} !</u>

New problems:
- Is Λ on the LHS or RHS?

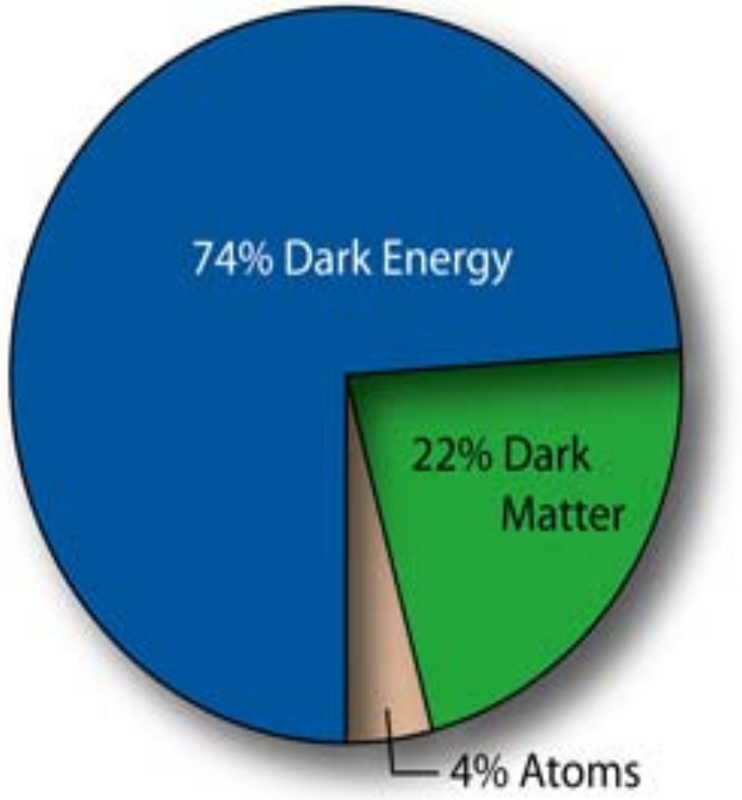
$$\boxed{R_{\mu\nu} - \frac{1}{2} R\, g_{\mu\nu} + \Lambda\, g_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}}$$

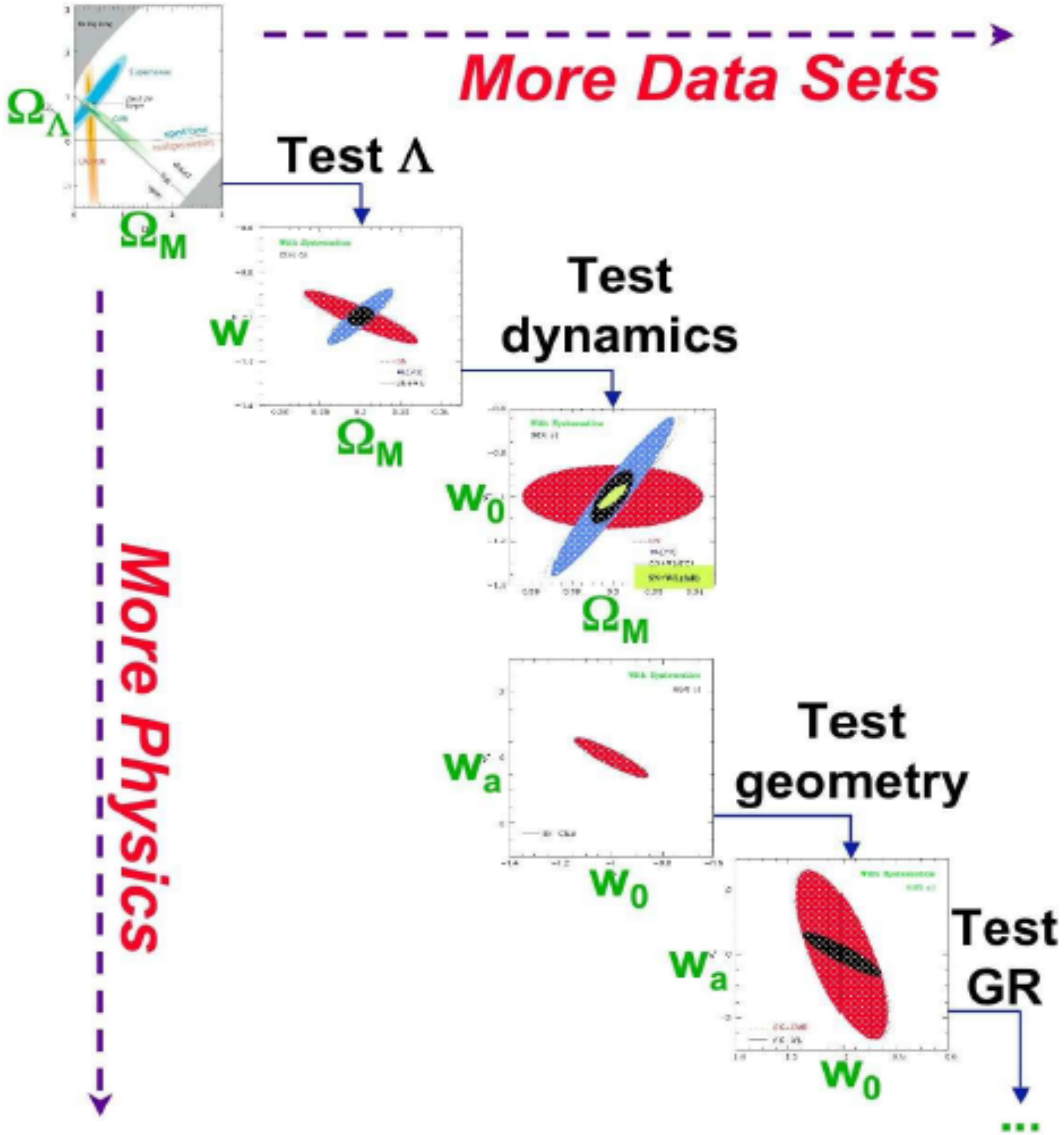
- Why are the amounts of Dark Matter and Λ so similar?

Darkness Visible

Cosmic Probes:

- Gravitational Lensing
- Peculiar Velocities
- Galaxy Clusters
- Cosmic Microwave Background
- Large Scale Structure
- Type Ia Supernovae
- Integrated Sachs-Wolf



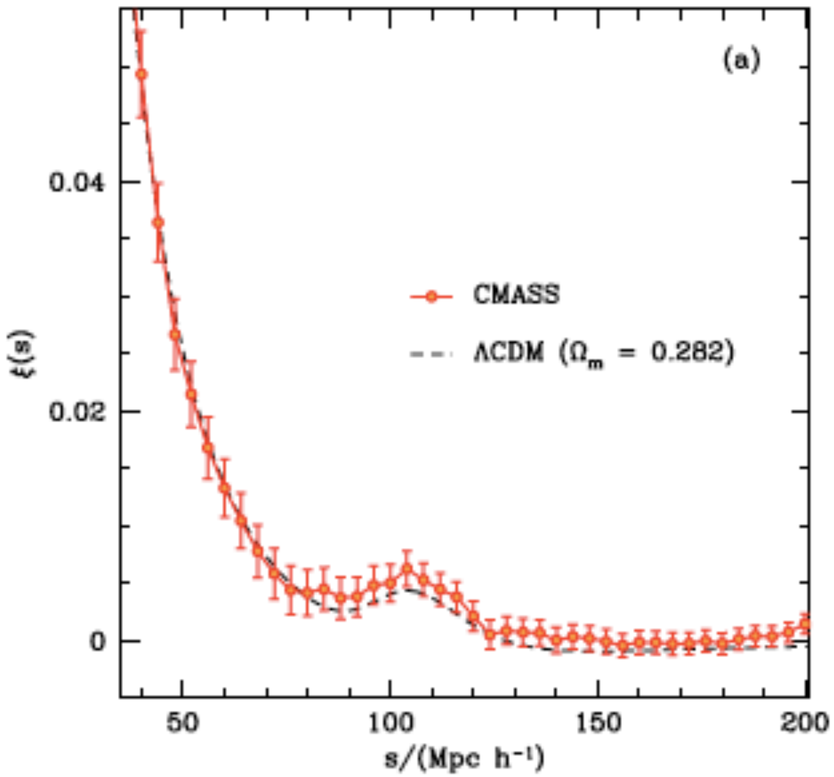


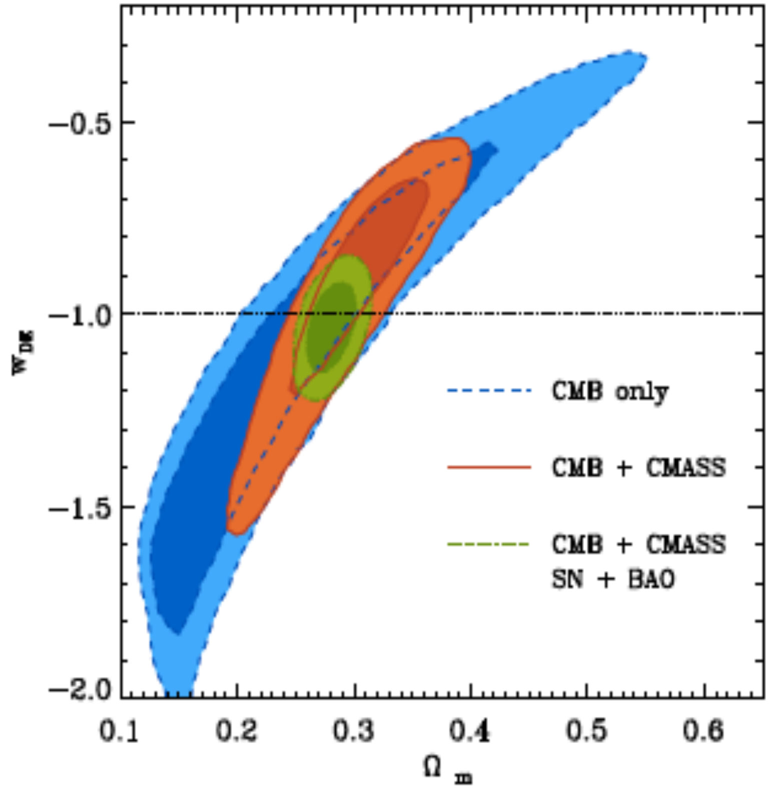
$$W = P/\rho$$

$$W = W_0 + (1-a)\, W_a$$

Linder 05

New Results - BOSS





basic parameters can be specified by $\Omega_{\rm m} = 0.285 \pm 0.009$, $100\,\Omega_{\rm b} = 4.59 \pm 0.09$, $n_{\rm s} = 0.961 \pm 0.009$, $H_0 = 69.4 \pm 0.8\,{\rm km\,s^{-1}\,Mpc^{-1}}$ and $\sigma_8 = 0.80 \pm 0.02$. The

w =-1.03 +- 0.07

Sanchez et al. 2012

The Landscape of Large Surveys

(some under construction, some proposed)

Photometric surveys: ***DES, VISTA, VST, Pan-STARRS, HSC, Skymapper, PAU, LSST, …***

Spetroscopic surveys: ***WiggleZ, BOSS, e-BOSS, BigBOSS, DESpec, HETDEX, Subaru/Sumire, VISTA/spec, SKA, …***

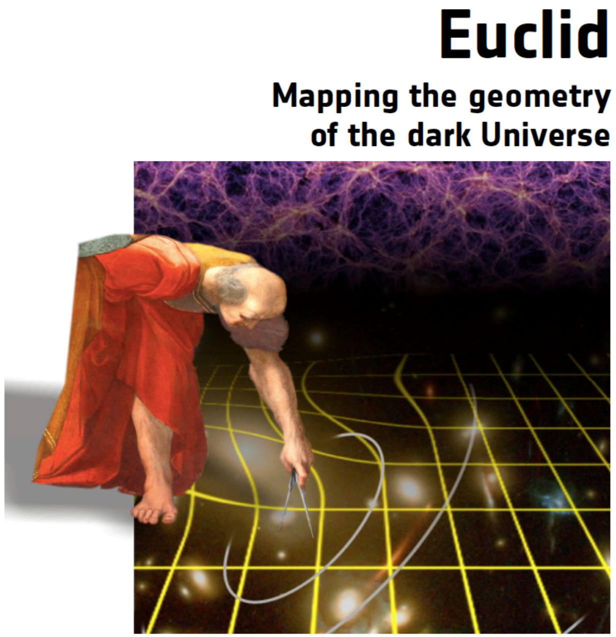
Space Missions: ***Euclid, WFIRST***

EUCLID

ESA Cosmic Vision
planned launch 2019

The key original ideas:
weak lensing from space
and photo-z from the ground
(DUNE) + spectroscopy (SPACE)

The new Euclid: 15000 sq deg
1B galaxy images + 50M spectra
(+ground based projects,
e.g. PS, DES, LSST,…)



Euclid Forecast

	Modified Gravity	Dark Matter	Initial Conditions	Dark Energy		
Parameter	γ	m_ν/eV	f_{NL}	w_p	w_a	*FoM*
Euclid Primary	0.010	0.027	5.5	0.015	0.150	430
Euclid All	0.009	0.020	2.0	0.013	0.048	1540
Euclid+Planck	0.007	0.019	2.0	0.007	0.035	4020
Current	0.200	0.580	100	0.100	1.500	~10
Improvement Factor	**30**	**30**	**50**	**>10**	**>50**	**>300**

The Dark Energy Survey

First Light in September 2012

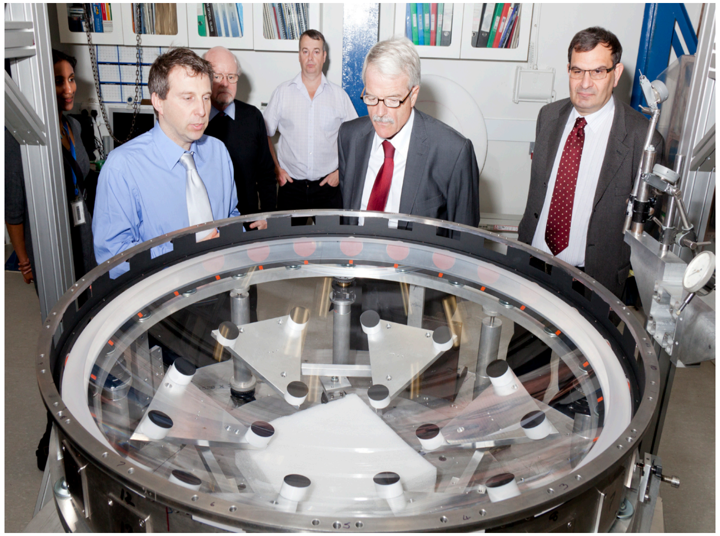
- Multi-probe approach

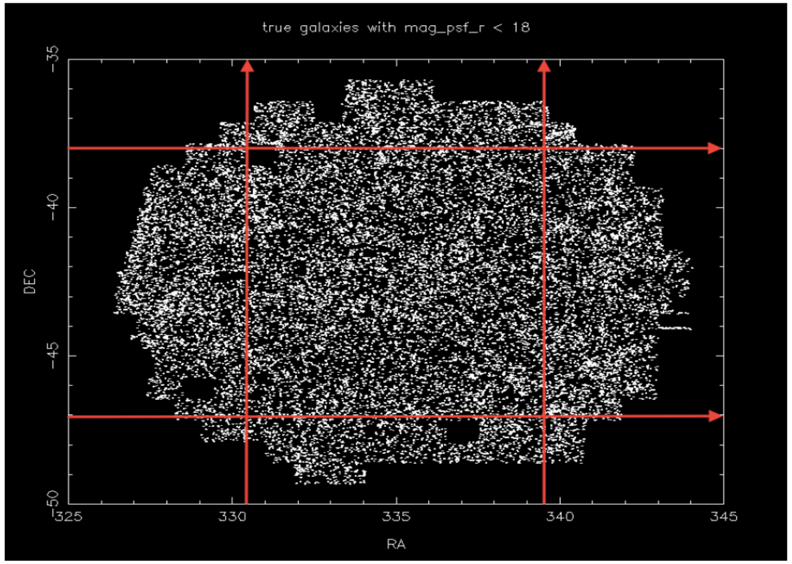
  Cluster Counts
  Weak Lensing
  Large Scale Structure
  Supernovae Ia

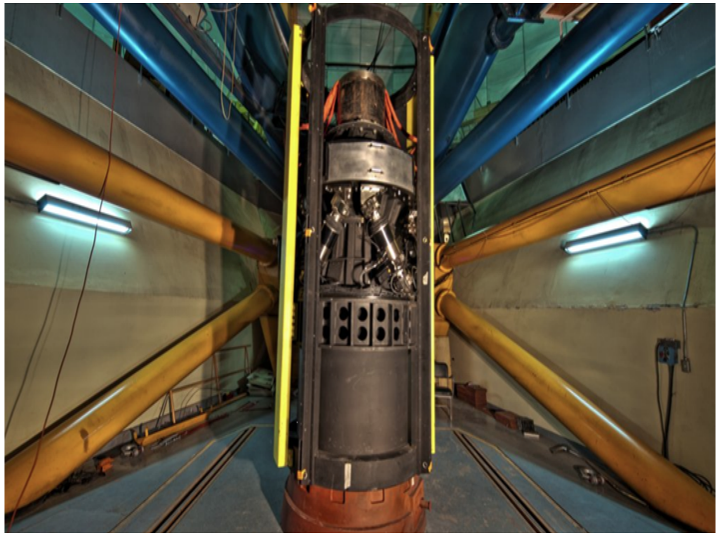
- 8-band survey

  5000 deg^2 *grizY*
  300 million photometric redshifts
  + *JHK* from VHS (1200 sq deg covered at half exposure time)
  +SPT SZ (550 clusters observed over 2500 sq deg)

The Dark Energy Survey

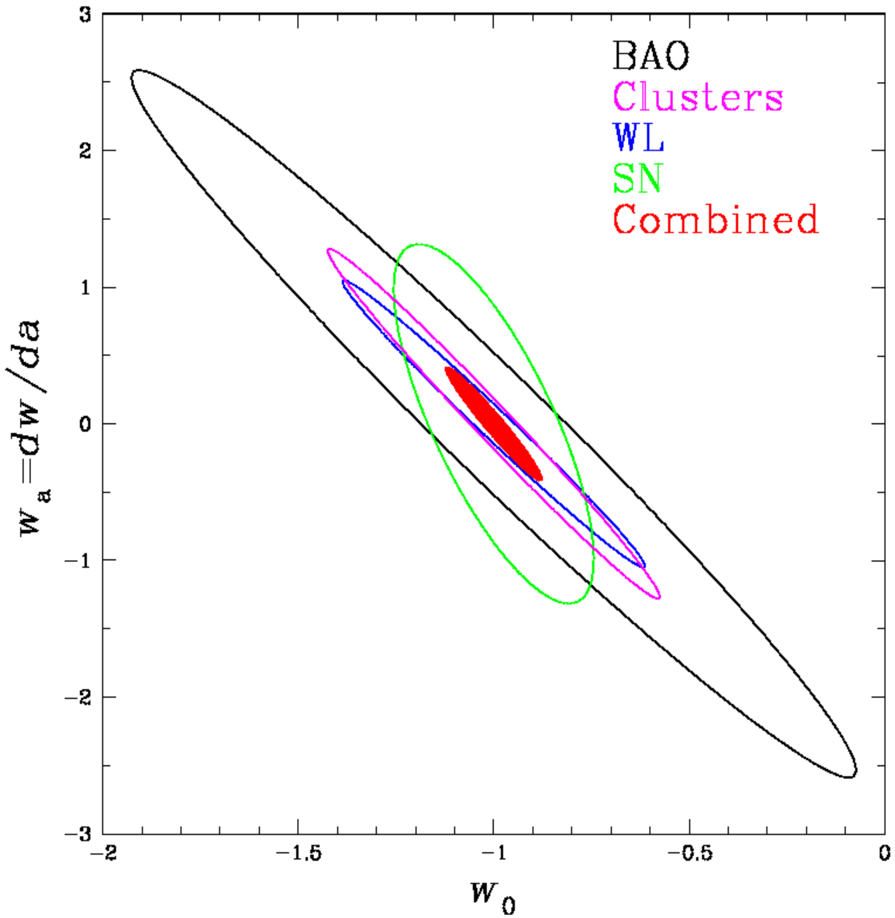






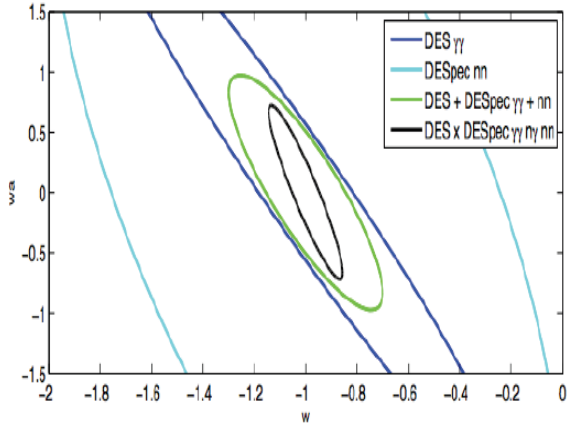
Fisher Matrix FoM as a tool for decision making by funding agencies

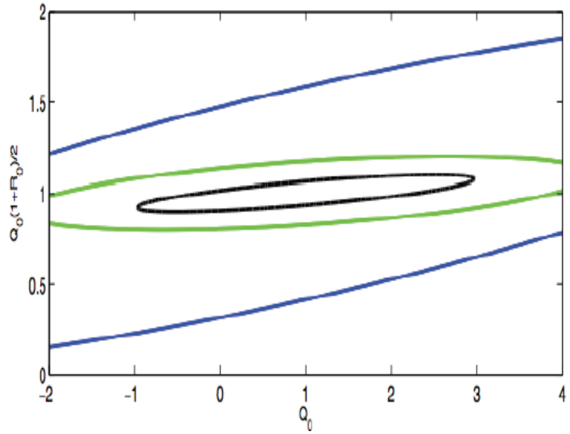
$w(z) = w_0 + w_a(1 - a)$ 68% CL



Example: FoM for Dark Energy Survey is a factor 4.6 tigther compared to near term projects

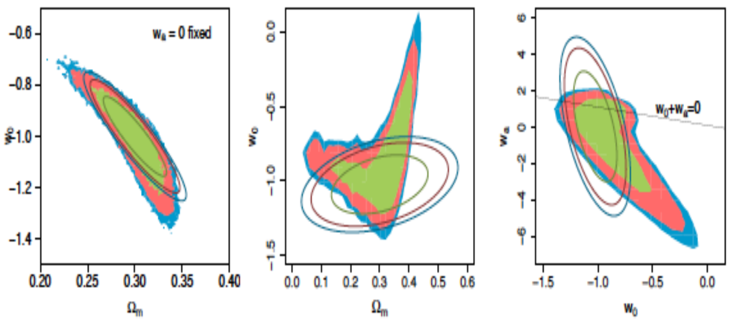
DES (WL) + DESpec (LSS)



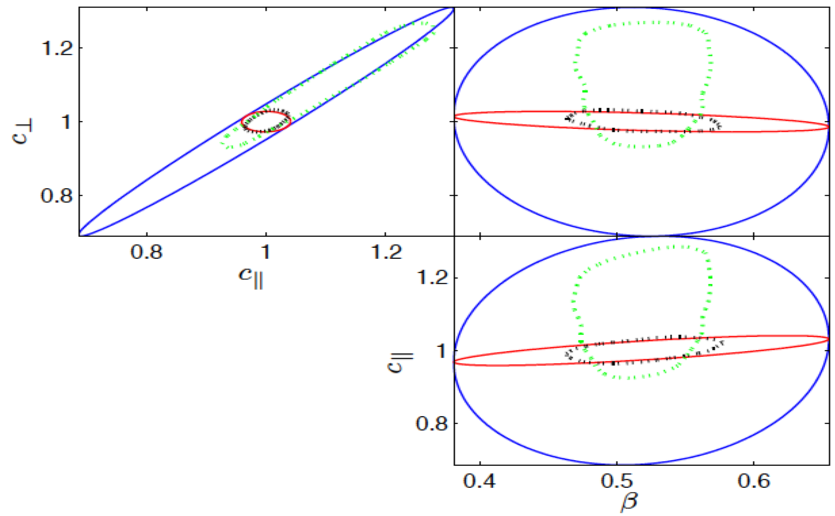


Kirk, Lahav, Bridle et al. (in preparation)
Cf. Gaztanaga et al 2012, Bernstein & Cai 2012

Fisher Matrix vs MCMC forecasting



SN Ia survey (Wolz et al. 2012)



Redshift distortion (1Gpc^3 survey) + CMB sound horizon (Hawken et al. 2012)

Use and Abuse of

- Priors
- Marginalization
- Evidence

How to choose a prior?
(e.g. on the curvature)

* Theoretical prejudice

  (e.g. "according to Inflation the universe must be flat")

* Previous observations

  (e.g. "we know from WMAP the universe is flat to within 2%")

* Parameterized ignorance (e.g. "uniform prior, Jeffrey's prior, or Entropy prior?")

Bayesian Evidence: To BE or not to BE?

$$p(d|\mathcal{M}) \equiv \int_{\Omega_{\mathcal{M}}} p(d|\theta, \mathcal{M}) p(\theta|\mathcal{M}) \mathrm{d}\theta \quad \text{(Bayesian evidence).}$$

$$B_{01} \equiv \frac{p(d|\mathcal{M}_0)}{p(d|\mathcal{M}_1)} \quad \text{(Bayes factor).}$$

e.g. B=1 weak; B=5 strong

But how sensitive to assumed priors?

Variations: AIC, BIC, DIC,…

e.g. Liddle (2007), Trotta (2008), Efstathiou (2009)

Sources of Systematics in Cosmology

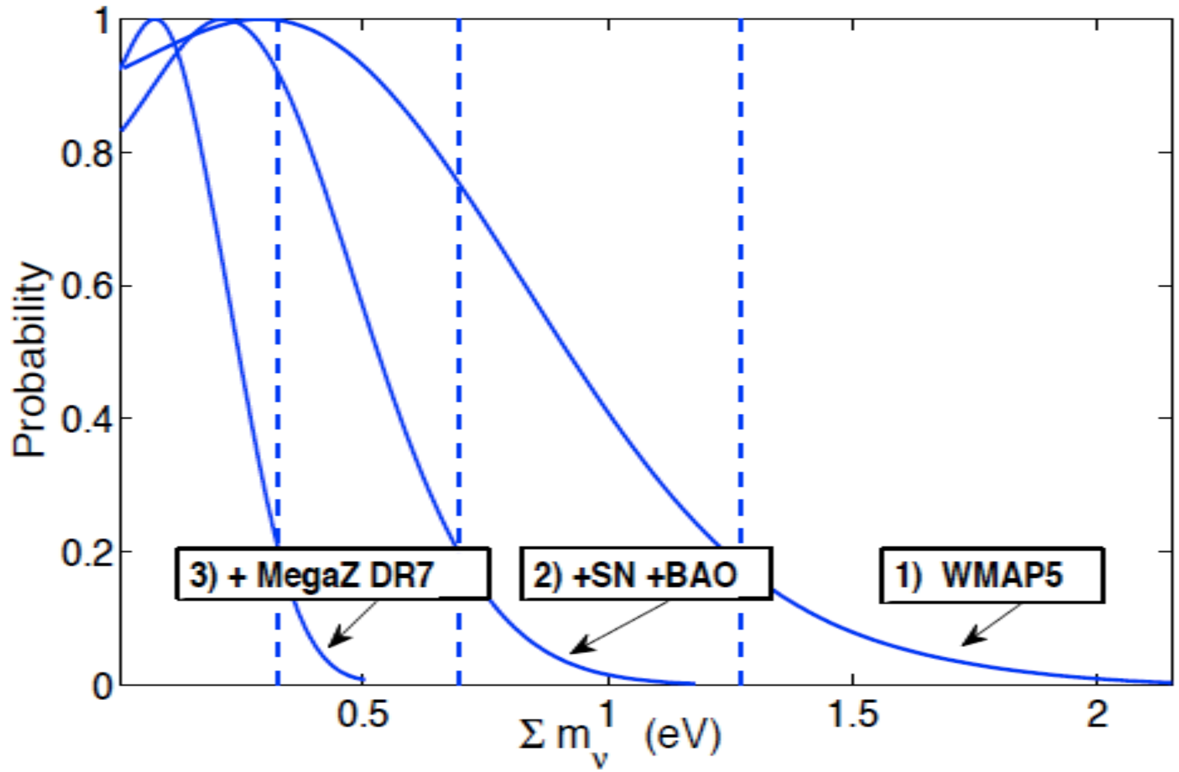
- Theoretical (the cosmological model & parameters, e.g. w/out neutrino mass)
- Astrophysical (e.g. galaxy biasing in LSS, dust in SN, intrinsic alignments in WL)
- Instrumental (e.g. image quality, photo-z)

The case for "Vanilla systematics"

- We model the whole universe with 6-12 parameters.
- How many parameters should we allow as "nuisance parameters" for unknown astrophysics – 10, 100, 1000?
- There is some knowledge from theory and simulations on galaxy biasing (and e.g. intrinsic alignments).
- A small number of physically motivated free parameters are easier for comparison with other analyses.
- These can be useful to test the1000-parameter setup (or their PCA-compressed version).

Upper limits on total Neutrino mass

Total mass < 0.28 eV (95% CL)



Thomas, Abdalla & Lahav, PRL (2010) 0911.5291

Astro-Statistics for Exo-Planets

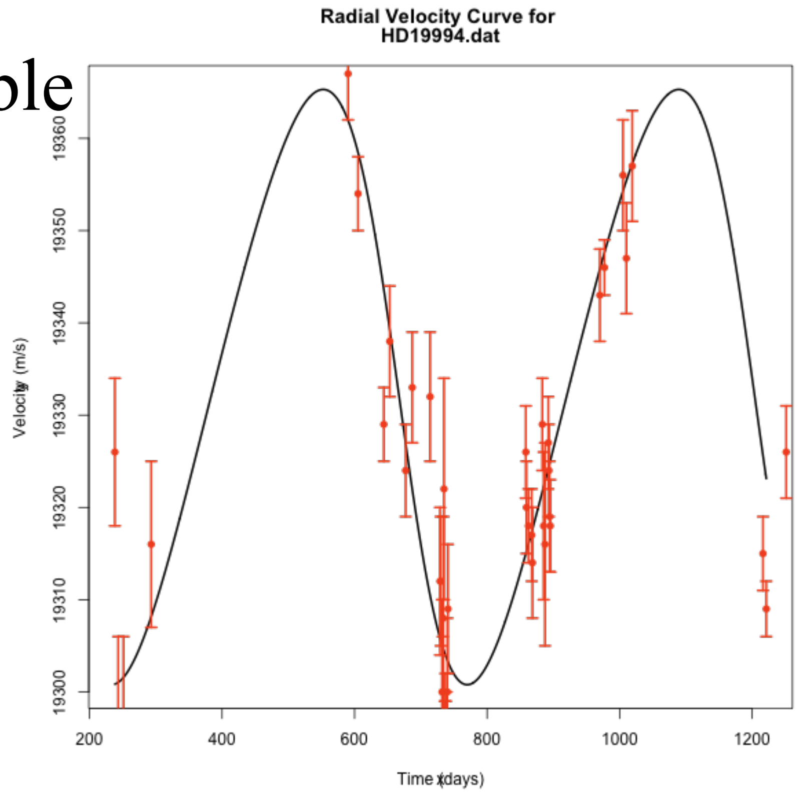
Orbital parameters from radial velocities

- ExoFit - publically available
- Bayesian approach,

  with emphasis on priors,

  implemented by

  MCMC



Balan & Lahav (MNRAS, 2009)

Cf. Gregory (2005), Ford (2006), Feroz et al. (2011)

Example 2:
Photometric Redshifts

Photo-z –Spectra cross talk

- Approximately, for a photo-z slice:

$(\delta w/ w) = 5 (\delta z/ z) = 5 (\sigma_z/z) N_s^{-1/2}$

=> the target accuracy in w

and photo-z scatter σ_z dictate the number of required spectroscopic redshifts

$N_s = 10^5 - 10^6$

Photometric redshift

- Probe strong spectral features (4000 break)
- Difference in flux through filters as the galaxy is redshifted.

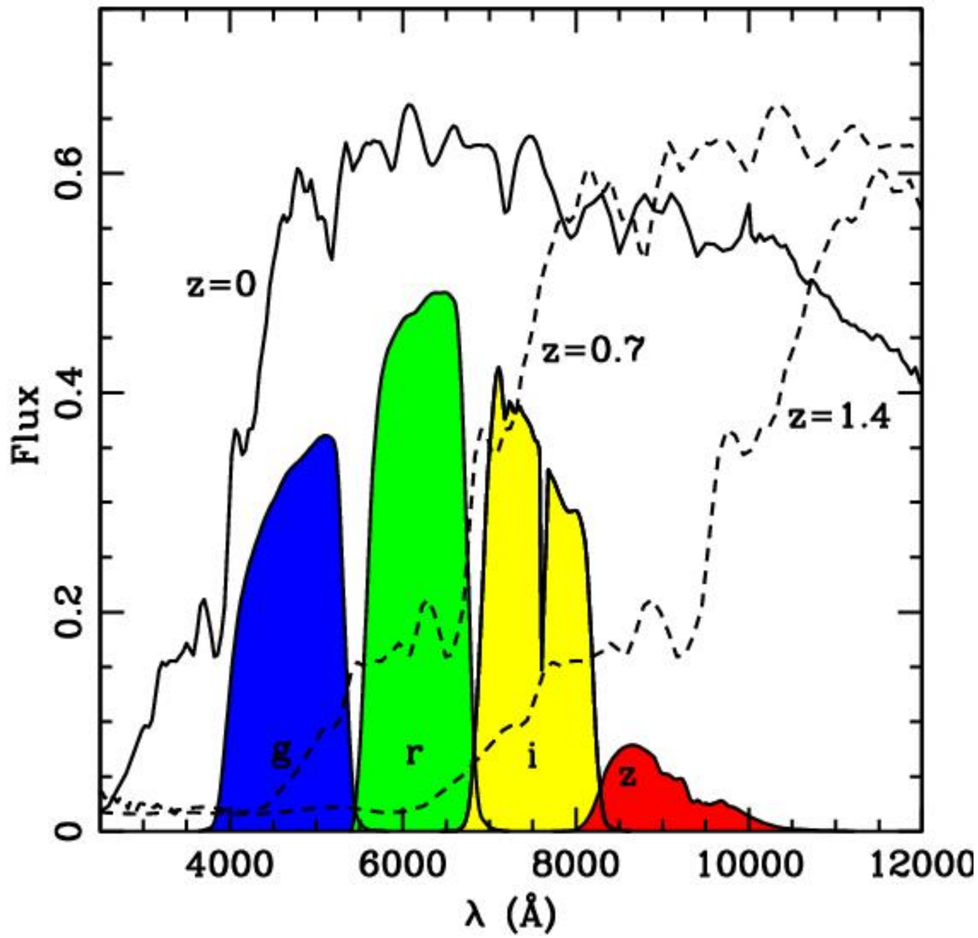
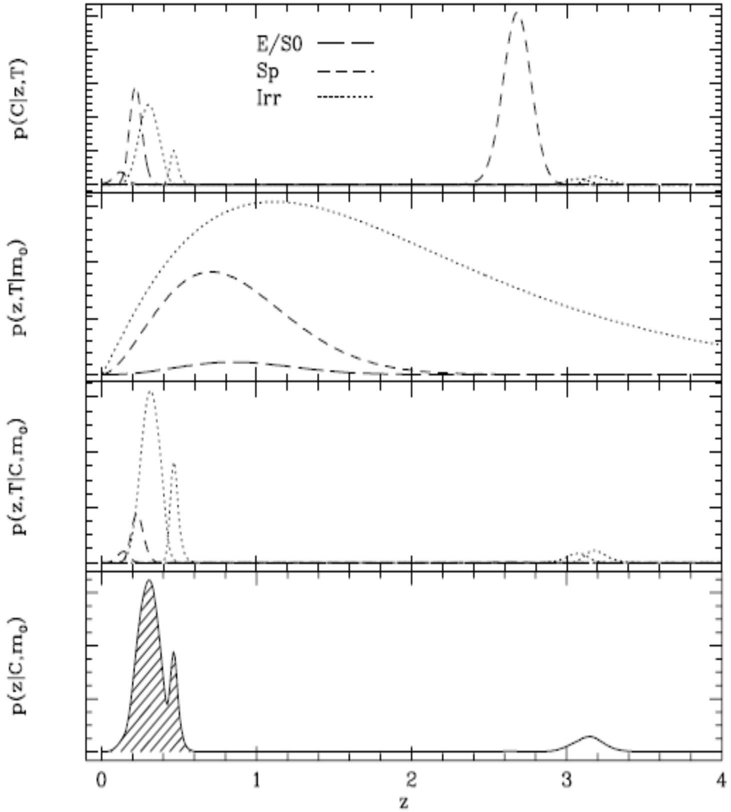


PHOTO-Z CODES

CODE	METHOD	REFERENCE
HyperZ	Template	Bolzonella et al. (2000)
BPZ	Bayesian	Benitez (2000)
ANNz	Training	Collister & Lahav (2004)
ImpZLite	Template	Babbedge et al. (2004)
SDSS Template	Hybrid	Padmanabhan et al. (2005)
ZEBRA	Hybrid, Bayesian	Feldmann et al. (2006)
LePhare	Template	Ilbert et al. (2006)

Bayesian Photo-z



Redshift z

$p(C|z,T)$ *likelihood*

$p(z,T|m_o)$ *prior*

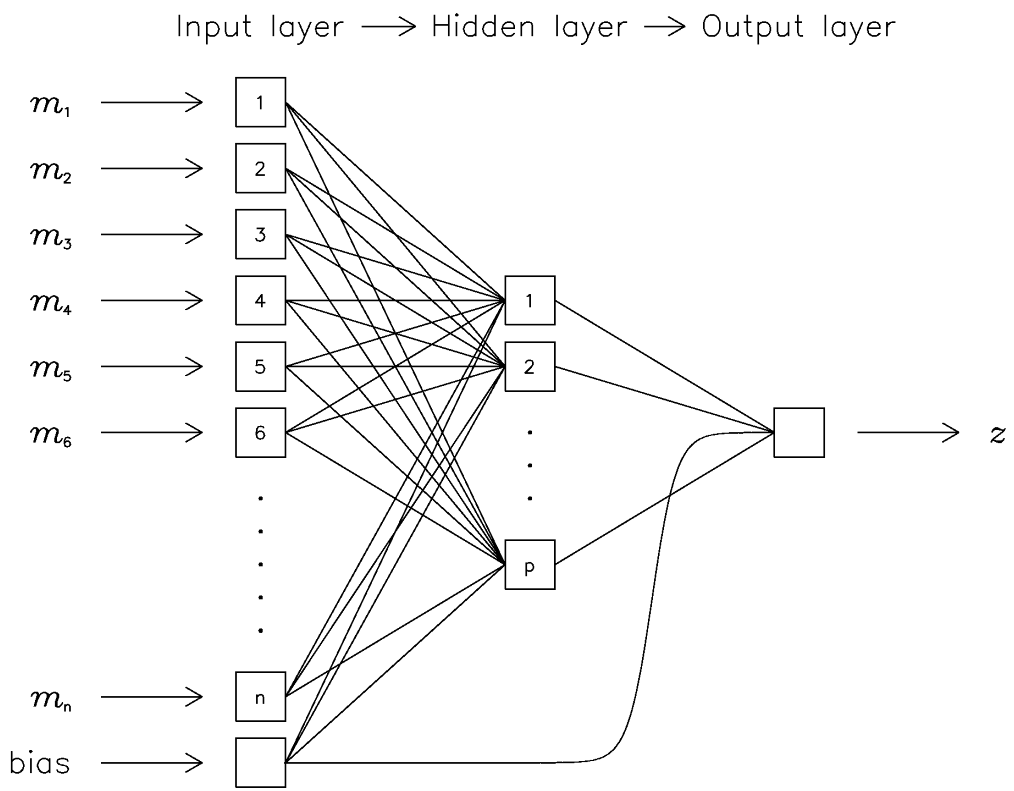
$$p(z, T \mid C, m_0) \propto p(z, T \mid m_0) p(C \mid z, T)$$

$$p(z \mid C, m_0) = \sum_T p(z, T \mid C, m_0)$$

Benitez 2000 (BPZ)

ANNz - Artificial Neural Network

Input: magnitudes



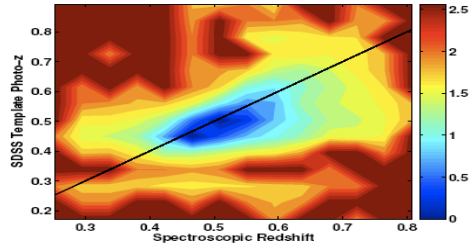
Output: redshift

Collister & Lahav 2004

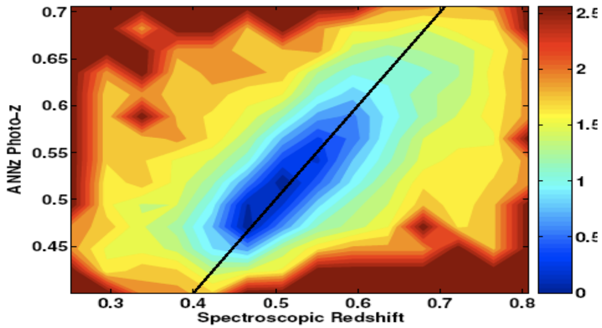
http://www.star.ucl.ac.uk/~lahav/annz.html

1.5M LRGs ("MegaZ") photo-z code comparison

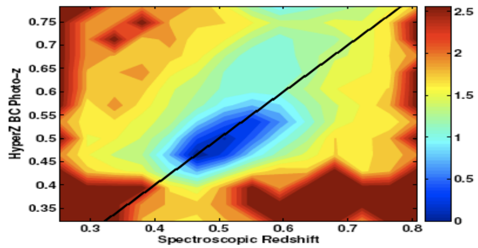
SDSS



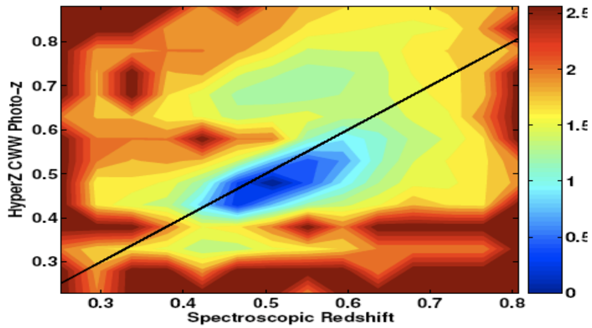
ANNz



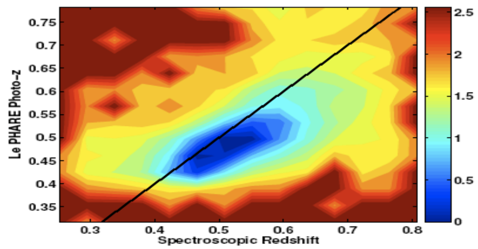
HpZ+BC



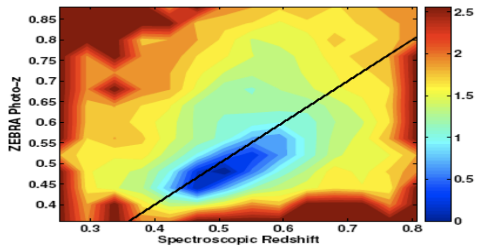
HpZ+WWC



Le PHARE



Zebra



Abdalla, Banerji, Lahav & Rashkov

Cf. PHAT (Hildebrandt et al. 2010)

Example 3:
Galaxy Classification and Shapes

GALAXY ZOO 2

- One Million galaxies classified by 100,000 people!

Is the galaxy simply smooth and rounded, with no sign of a disk?

Smooth

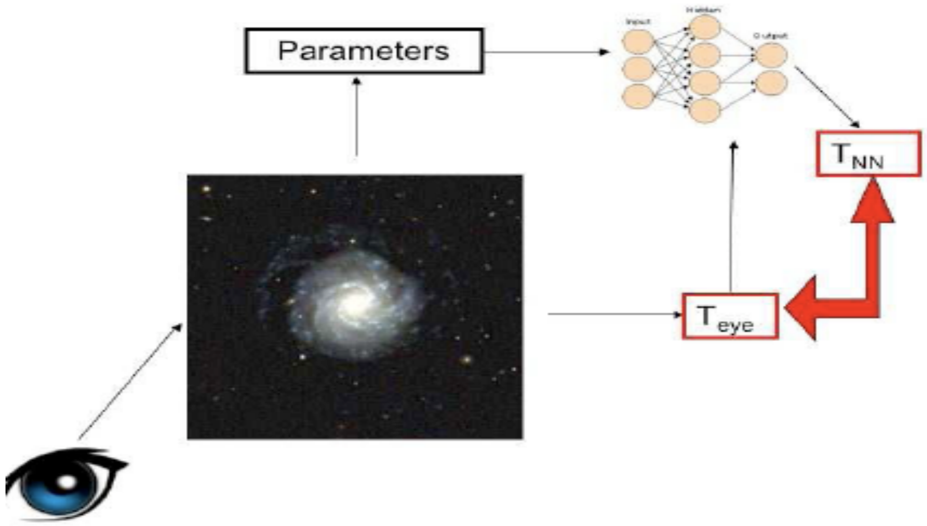
Features or disk

Star or artifact

Need help?

Lintott et al.

Galaxy zoo and machine learning



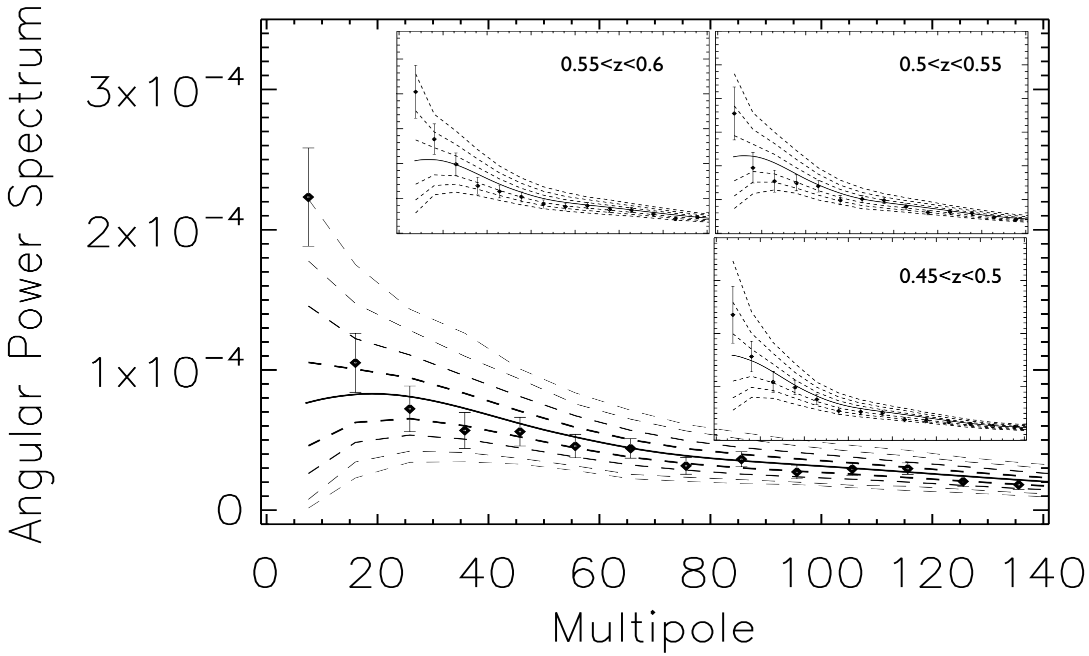
		GALAXY ZOO		
		Elliptical	Spiral	Star/Other
A	ELLIPTICAL	91%	0.08%	0.5%
N	SPIRAL	0.1%	93%	0.2%
N	STAR/OTHER	0.3%	0.3%	96%

Banerji, Lahav et al. (0908.2033)

Systematics in LSS

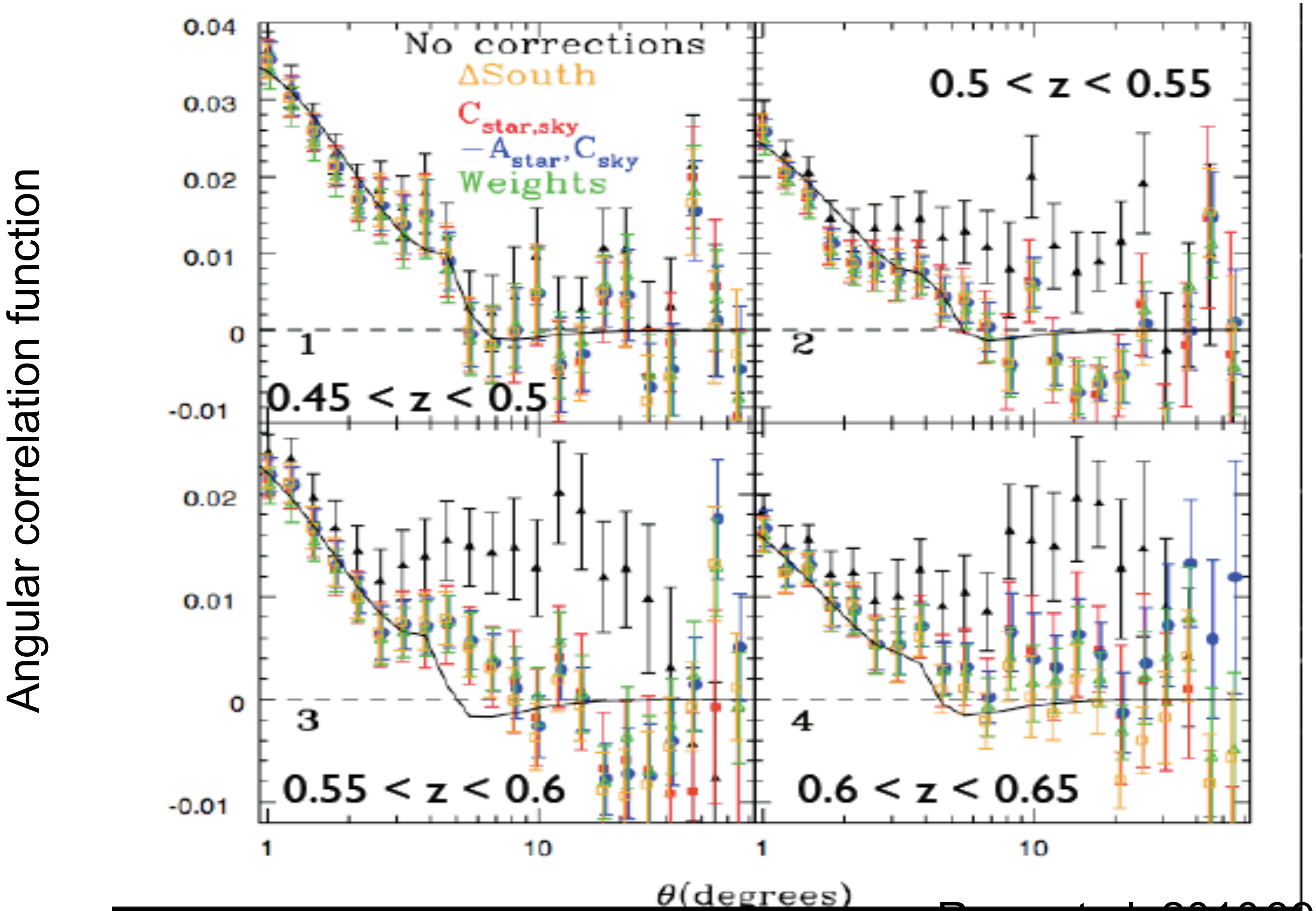
- Star-galaxy separation
- Galactic extinction
- Seeing
- Sky brightness
- Airmass
- Calibration offsets
- Others…

Excess power on Gpc scale: systematics or new physics?



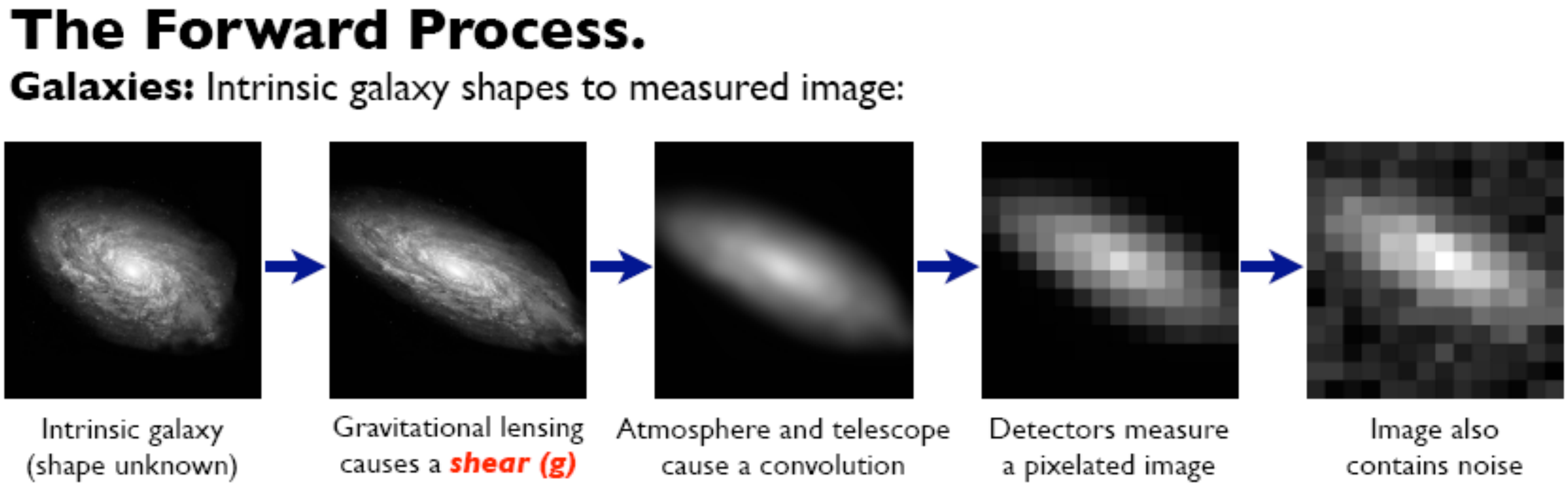
Thomas, Abdalla & Lahav (2011)
Using MegaZ-LRG (ANNz Photo-z)

Corrections to angular correlation function



Ross et al. 2010

Cosmic shear measurement



1% effect, to be measured to 1 % to get w to 1%

GREAT08 (Bridle et al.); GREAT10 (Kitching et al.);
GREAT3 (Mandelbaum & Rowe et al)

Further input much needed from statistics

- Model selection methodology
- MCMC machinery and extensions
- Detection of non-Gaussianity and shape finders
- Blind de-convolution (eg. PSF)
- Object classification
- Comparing simulations with data
- Visualisation
- VO technology

END