Data Analysis in Cosmology

Arman Shafieloo

APCTP

Asia Pacific Center for Theoretical Physics

26 September 2014,

Topical Program

APCTP-Seoul Branch

In doing cosmology:

- Instrumentalists: People who build the instruments (telescopes, detectors, mirrors, rockets, space crafts etc)

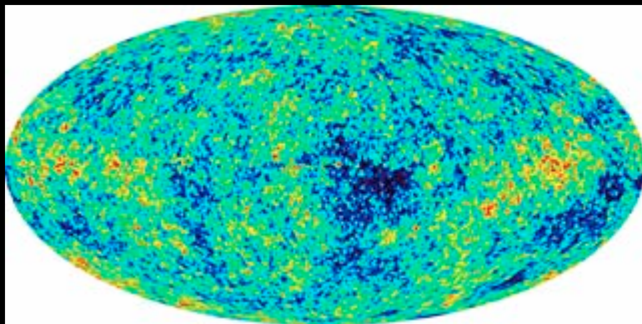
- Observers: People who run the experiments and observe the Universe and get the data

- Theorists:

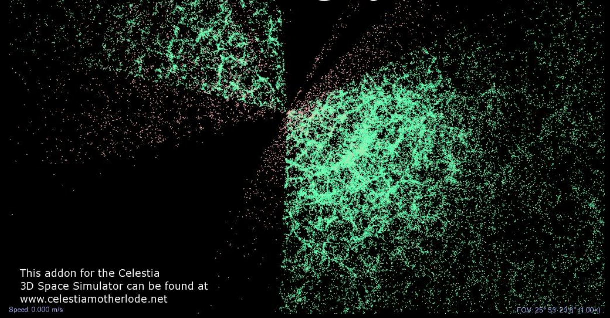
  1) People who deal with the data in order to get meaningful information, those who do simulations

  2) People who build models to explain the data

Era of Precision Cosmology



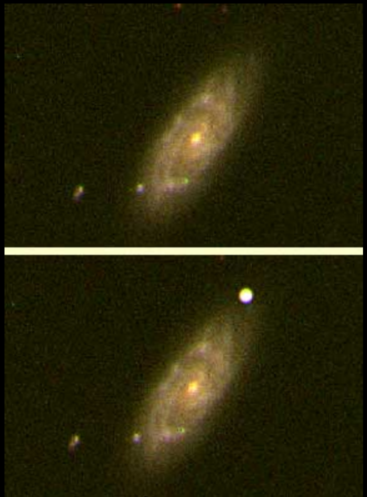
Cosmic Microwave Background (CMB)



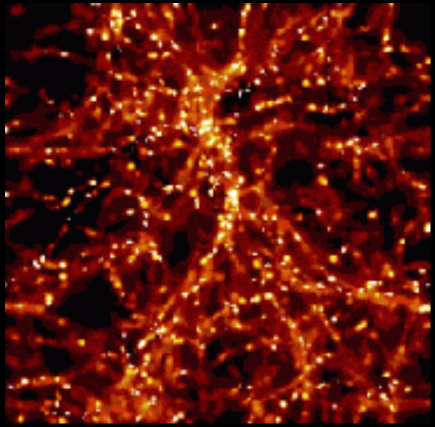
Large-scale structure

Cosmological Observations

Gravitational Lensing



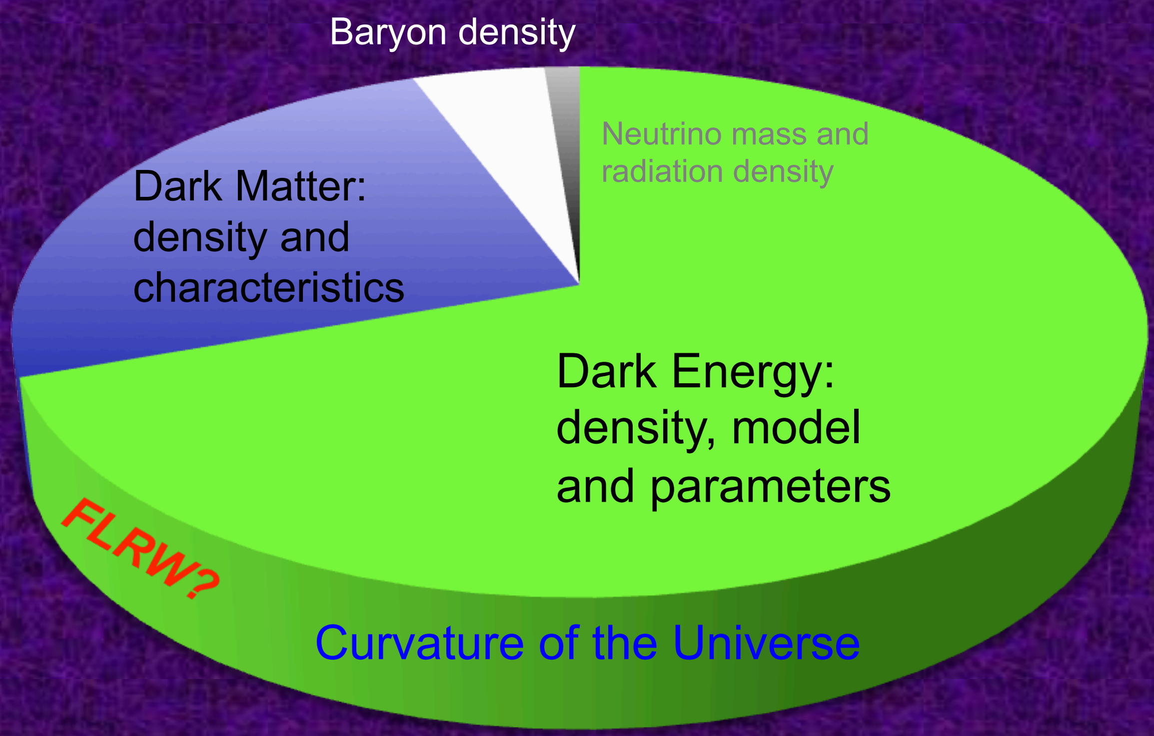
Type Ia Supernova



Lyman Alpha Forest

Era of Precision Cosmology

Combining new measurements and using statistical techniques to place sharp constraints on cosmological models and their parameters.

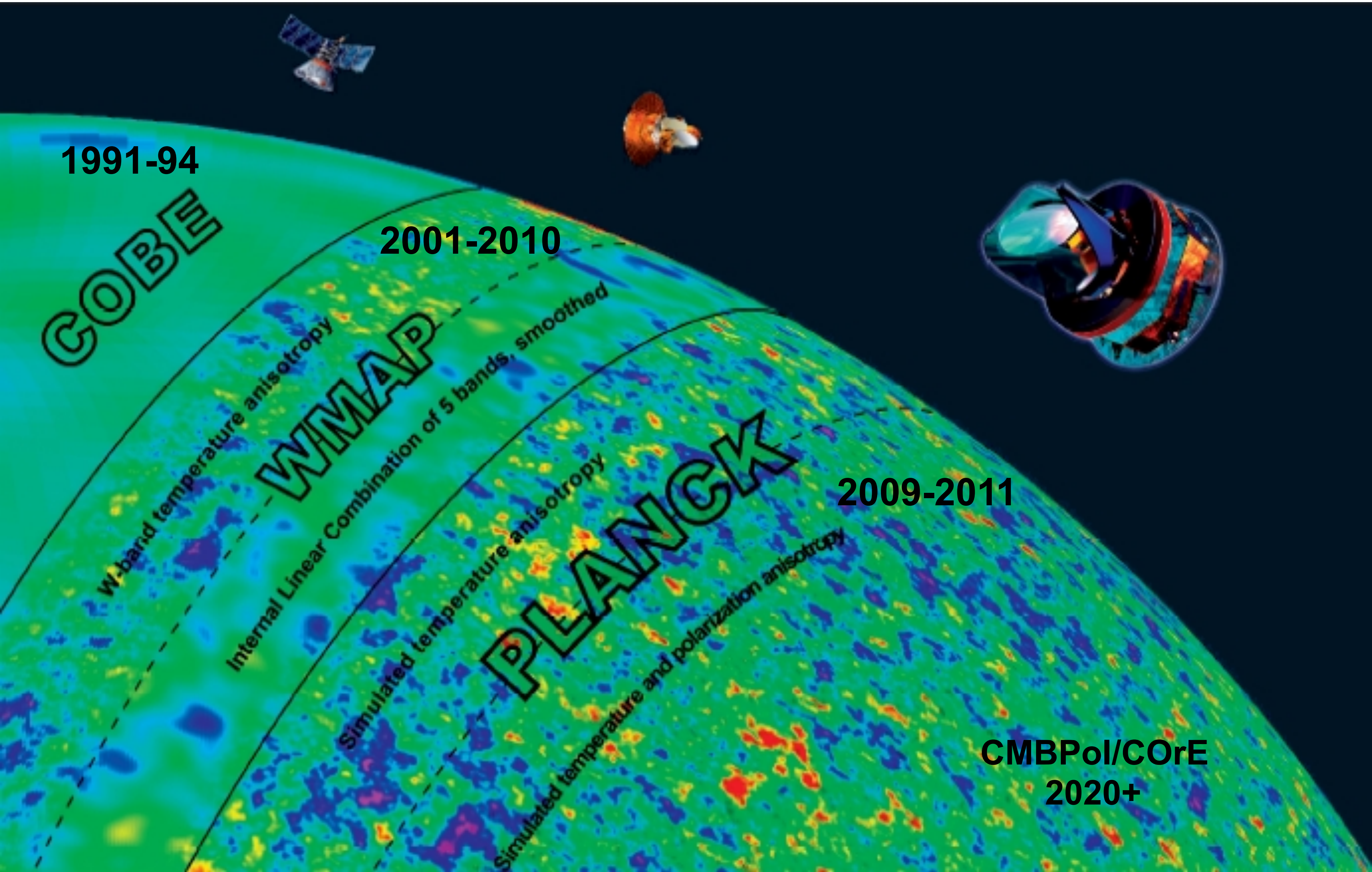


Initial Conditions:
Form of the Primordial Spectrum and Model of Inflation and its Parameters

Epoch of reionization

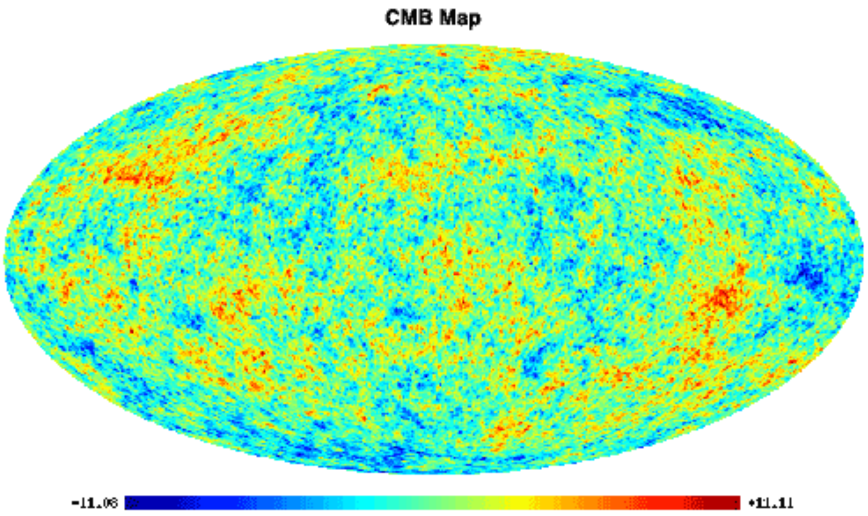
Hubble Parameter and the Current Rate of Expansion

CMB space missions



Statistics of CMB

CMB Anisotropy Sky map => Spherical Harmonic decomposition



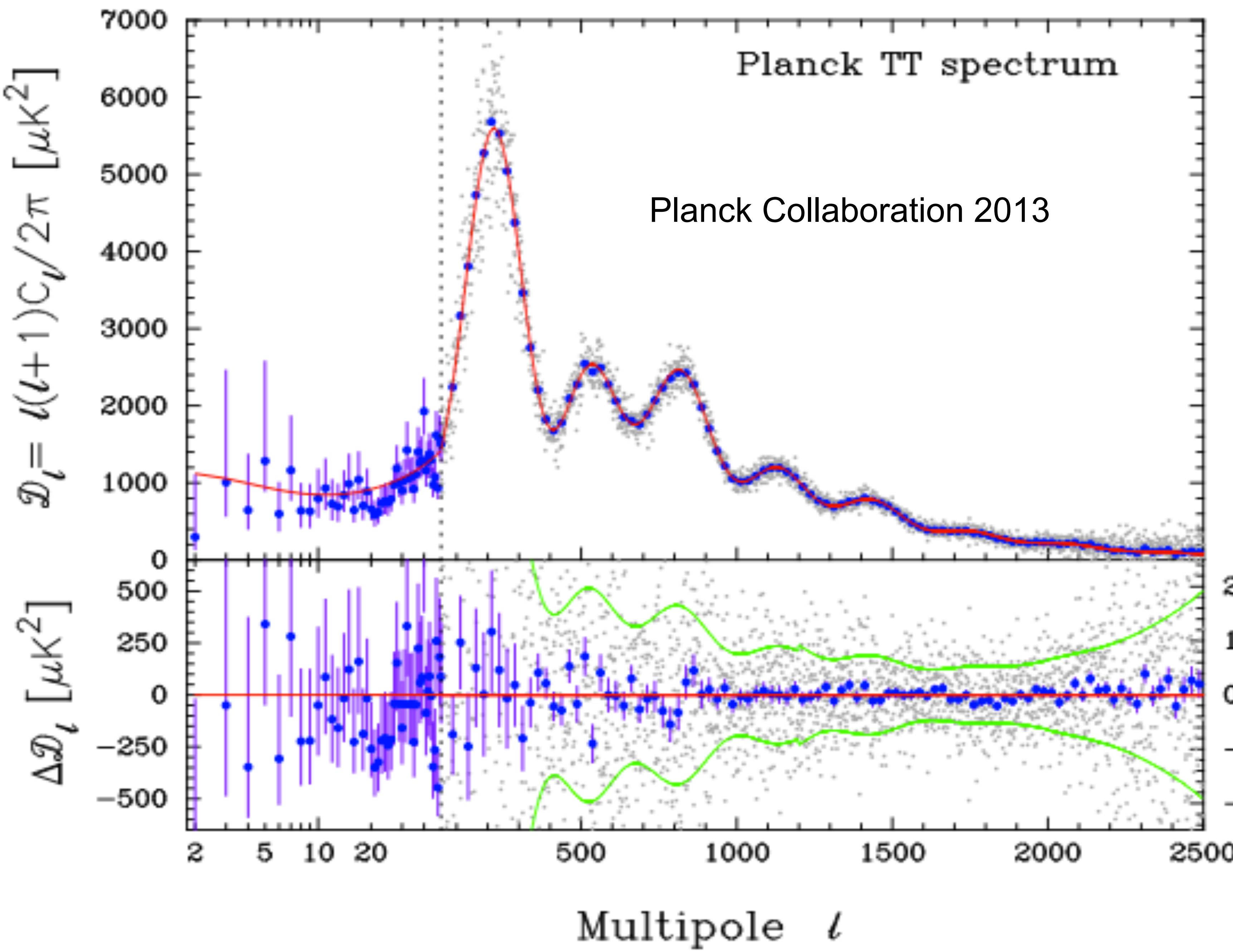
$$\Delta T(\theta,\phi) = \sum_{l=2}^{\infty} \sum_{m=-l}^{l} a_{lm} Y_{lm}(\theta,\phi)$$

$$\langle a_{lm}\, a^{*}_{l'm'} \rangle = C_l \; \delta_{ll'} \delta_{mm'}$$

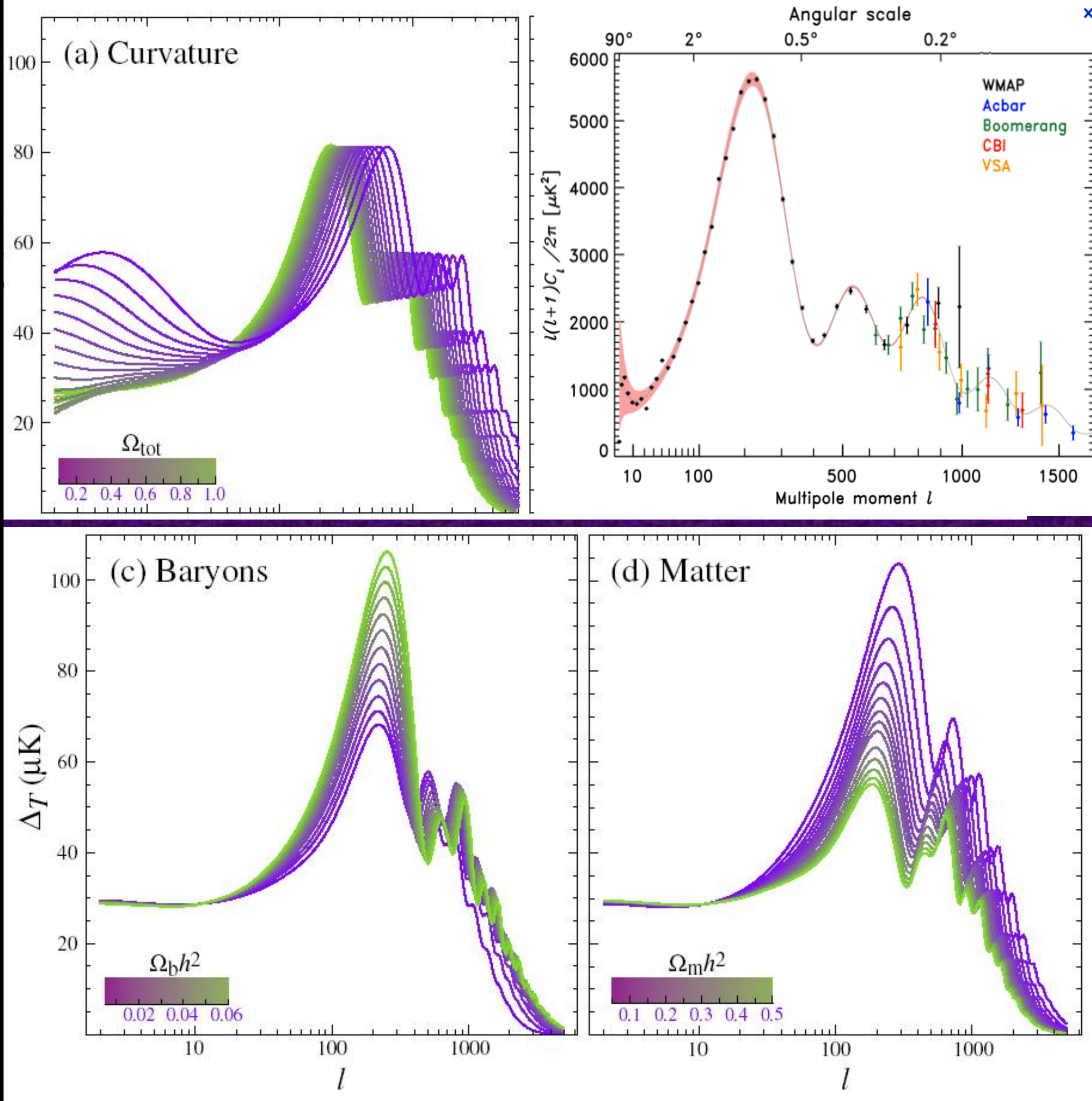
Gaussian Random field => Completely specified by

angular power spectrum $l(l+1)C_l$:

Power in fluctuations on angular scales of $\sim \pi/l$



Planck TT spectrum
Planck Collaboration 2013
$\mathcal{D}_\ell = \ell(\ell+1)C_\ell/2\pi\ [\mu K^2]$
$\Delta\mathcal{D}_\ell\ [\mu K^2]$
Multipole ℓ



Sensitivity of the CMB acoustic temperature spectrum to four fundamental cosmological parameters.

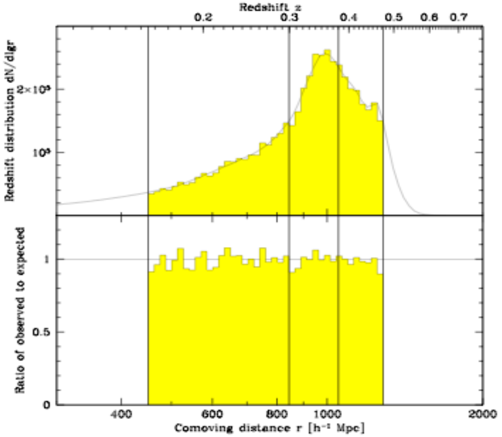
Total density

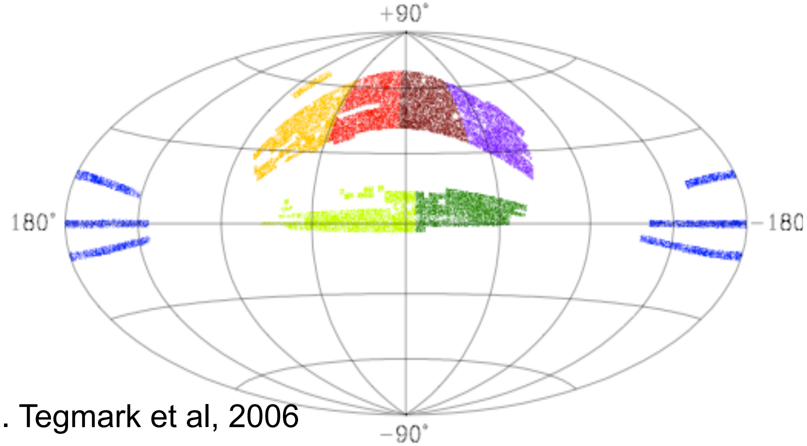
Dark Energy

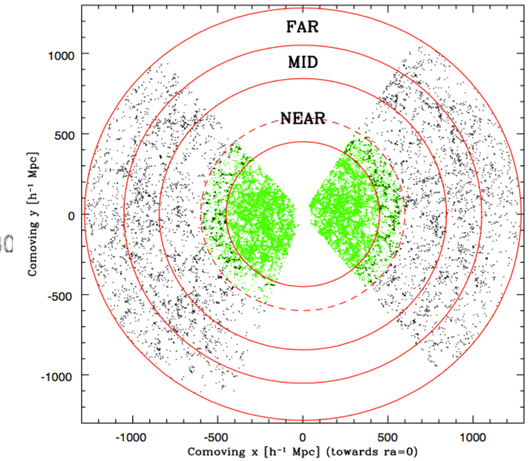
Baryon density and

Matter density.

From Hu & Dodelson, 2002



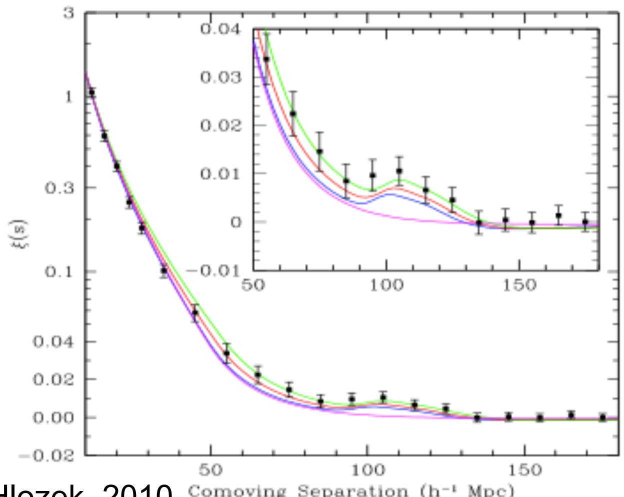




M. Tegmark et al, 2006

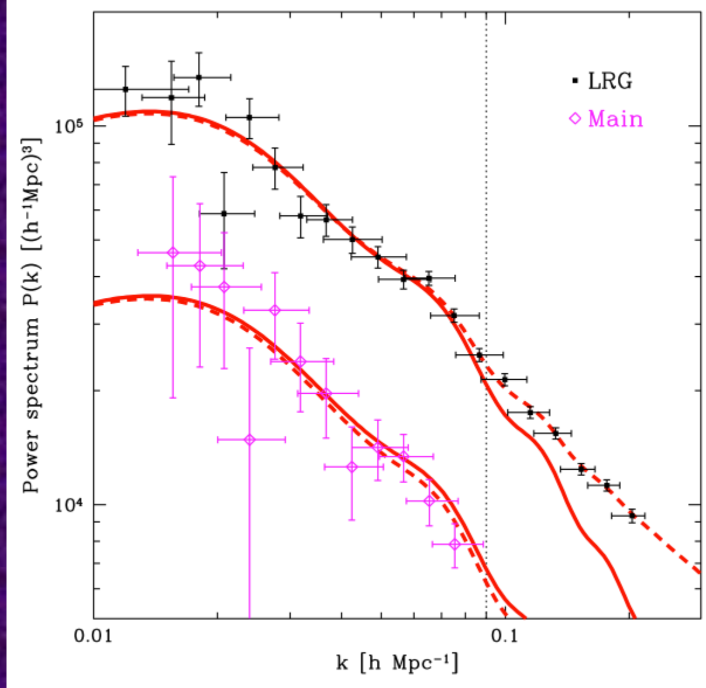
Large Scale Structure Data and Distribution of Galaxies

$$P(k) = \int_{-\infty}^{\infty} \xi(r) \exp(-ikr) r^2 dr \, .$$



Bassett & Hlozek, 2010

Fig. 1.1. The Baryon Acoustic Peak (BAP) in the correlation function – the BAP is visible in the clustering of the SDSS LRG galaxy sample, and is sensitive to the matter density (shown are models with $\Omega_m h^2 = 0.12$ **(top)**, 0.13 **(second)** and 0.14 **(third)**, all with $\Omega_b h^2 = 0.024$). The bottom line without a BAP is the correlation function in the pure CDM model, with $\Omega_b = 0$. From Eisenstein *et al.*, 2005 (52).



Large Scale Structure Data and Distribution of Galaxies

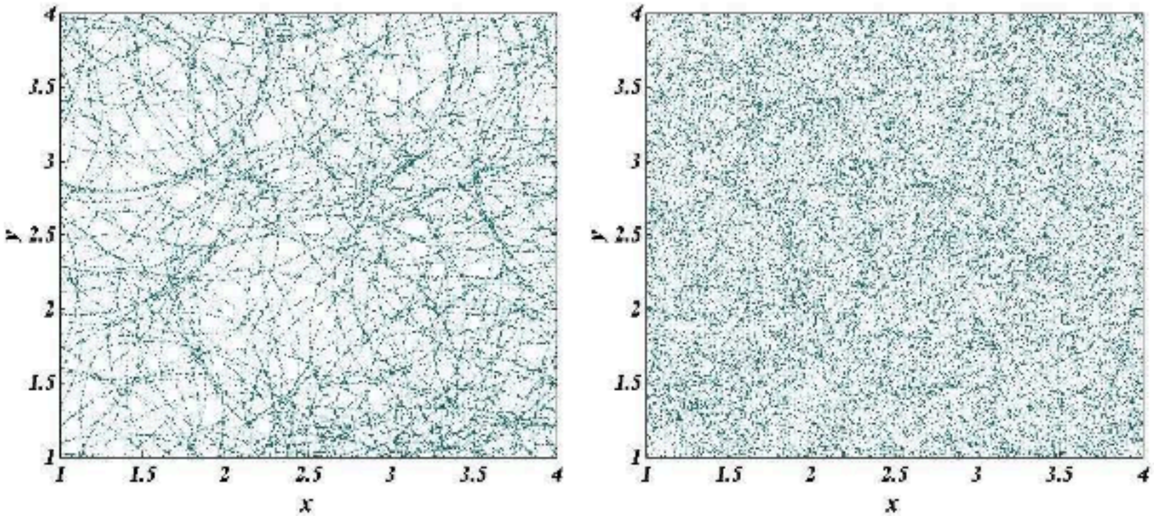
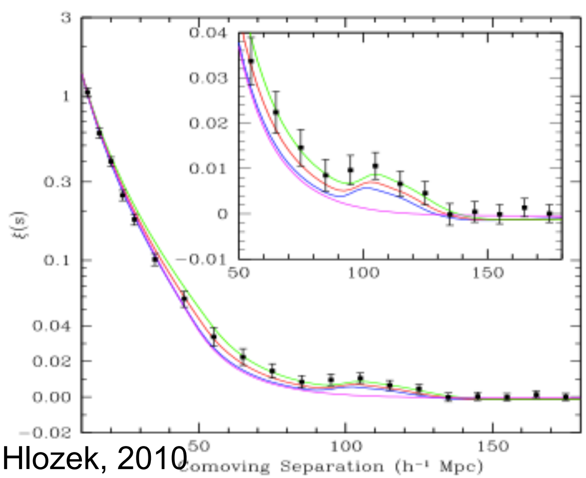


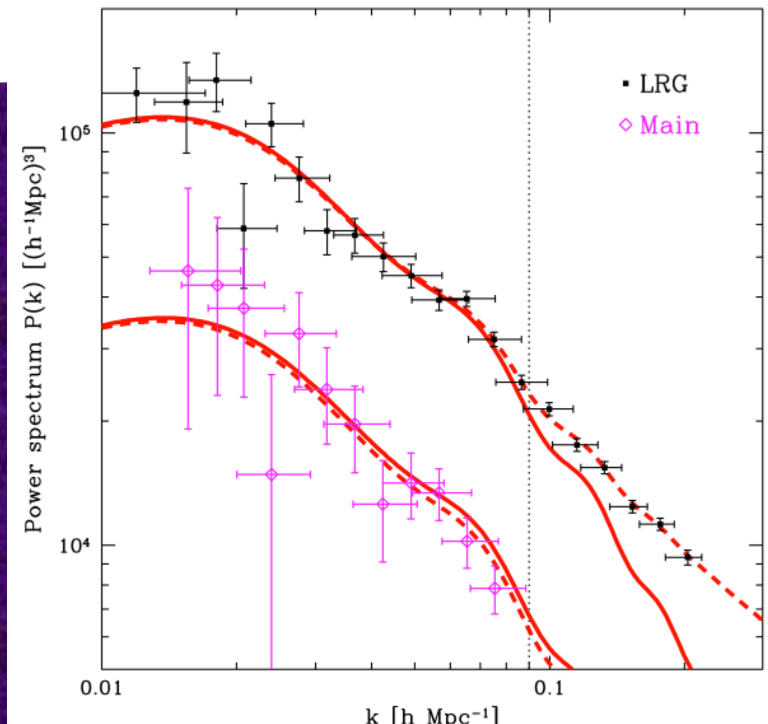
Fig. 1.5. Rings of power superposed. Schematic galaxy distribution formed by placing the galaxies on rings of the same characteristic radius L. The preferred radial scale is clearly visible in the left hand panel with many galaxies per ring. The right hand panel shows a more realistic scenario - with many rings and relatively few galaxies per ring, implying that the preferred scale can only be recovered statistically.

$$P(k) = \int_{-\infty}^{\infty} \xi(r) \exp(-ikr) r^2 dr \, .$$



Bassett & Hlozek, 2010

Fig. 1.1. The Baryon Acoustic Peak (BAP) in the correlation function – the BAP is visible in the clustering of the SDSS LRG galaxy sample, and is sensitive to the matter density (shown are models with $\Omega_m h^2 = 0.12$ **(top)**, 0.13 **(second)** and 0.14 **(third)**, all with $\Omega_b h^2 = 0.024$). The bottom line without a BAP is the correlation function in the pure CDM model, with $\Omega_b = 0$. From Eisenstein *et al.*, 2005 (52).



Measuring Distances in Astronomy

Supernovae Ia Observations

We can see supernovae up to very very large distances

- By observing their brightness and red shift we can calculate the distances and understand how far they are.

Asiago 1.82m + AFOSC

courtesy ESC

Measuring Distances in Astronomy

Standard Candles

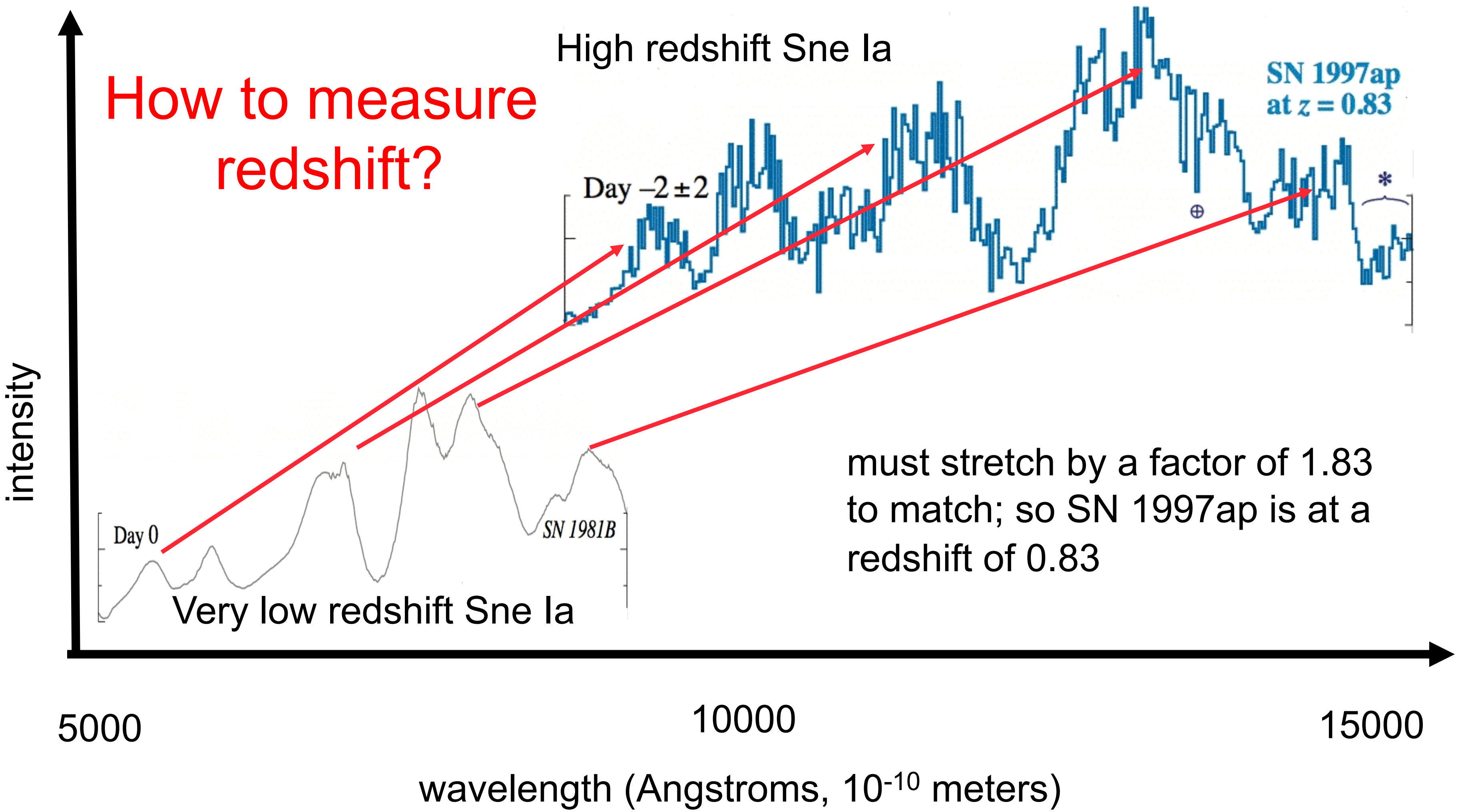
Supernovae type Ia are standard candles because we know how bright they are.

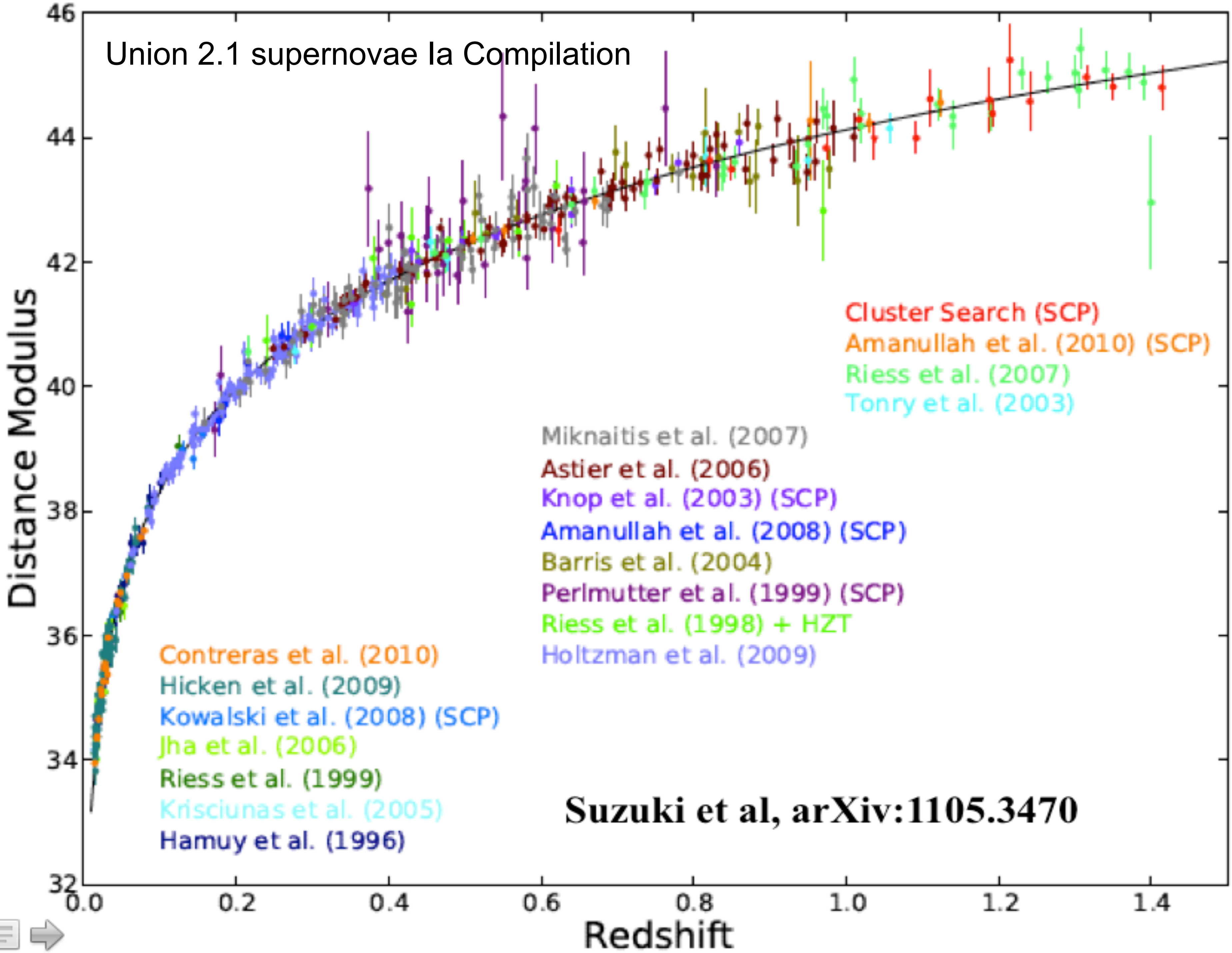
Carbon-Oxygen white dwarfs that accretes mass from a binary companion and core reaches to ignition temperature for Carbon fusion.

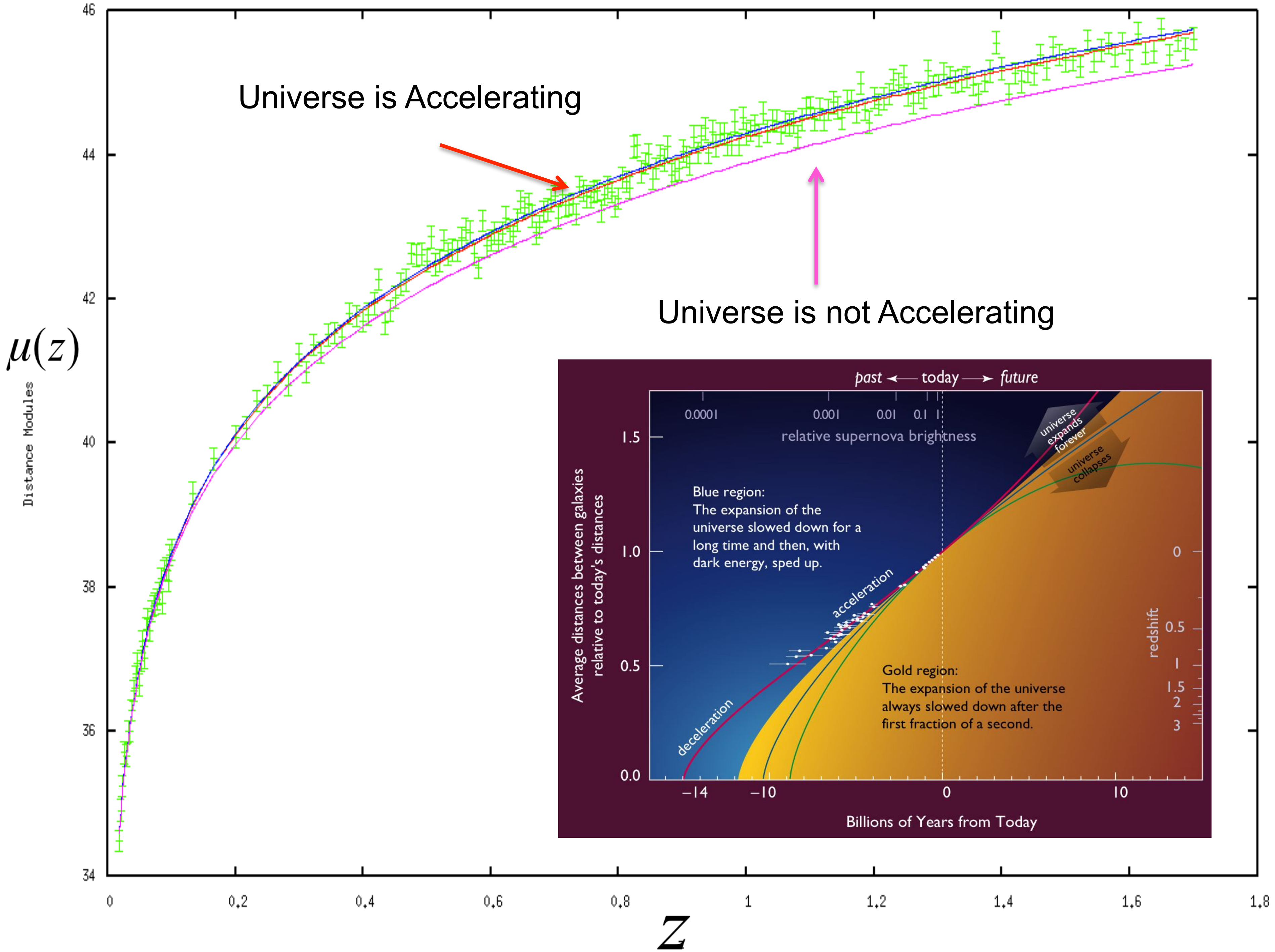
Brightness tells us distance away (lookback time)

Redshift measured in spectrum tells us expansion factor (average distance between galaxies)

Measuring Distances in Astronomy




Union 2.1 supernovae Ia Compilation
Distance Modulus
Redshift
46
44
42
40
38
36
34
32
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
Cluster Search (SCP)
Amanullah et al. (2010) (SCP)
Riess et al. (2007)
Tonry et al. (2003)
Miknaitis et al. (2007)
Astier et al. (2006)
Knop et al. (2003) (SCP)
Amanullah et al. (2008) (SCP)
Barris et al. (2004)
Perlmutter et al. (1999) (SCP)
Riess et al. (1998) + HZT
Holtzman et al. (2009)
Contreras et al. (2010)
Hicken et al. (2009)
Kowalski et al. (2008) (SCP)
Jha et al. (2006)
Riess et al. (1999)
Krisciunas et al. (2005)
Hamuy et al. (1996)
Suzuki et al, arXiv:1105.3470



Universe is Accelerating
Universe is not Accelerating
$\mu(z)$
Distance Modules
46
44
42
40
38
36
34
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
z
past ← today → future
0.0001
0.001
0.01
0.1
1
relative supernova brightness
Blue region:
The expansion of the universe slowed down for a long time and then, with dark energy, sped up.
universe expands forever
universe collapses
acceleration
deceleration
Gold region:
The expansion of the universe always slowed down after the first fraction of a second.
Average distances between galaxies relative to today's distances
1.5
1.0
0.5
0.0
redshift
0
0.5
1
1.5
2
3
−14
−10
0
10
Billions of Years from Today

Eureka! or
“What’s wrong with this?”
Adam Riess’s notebook
Fall 1997

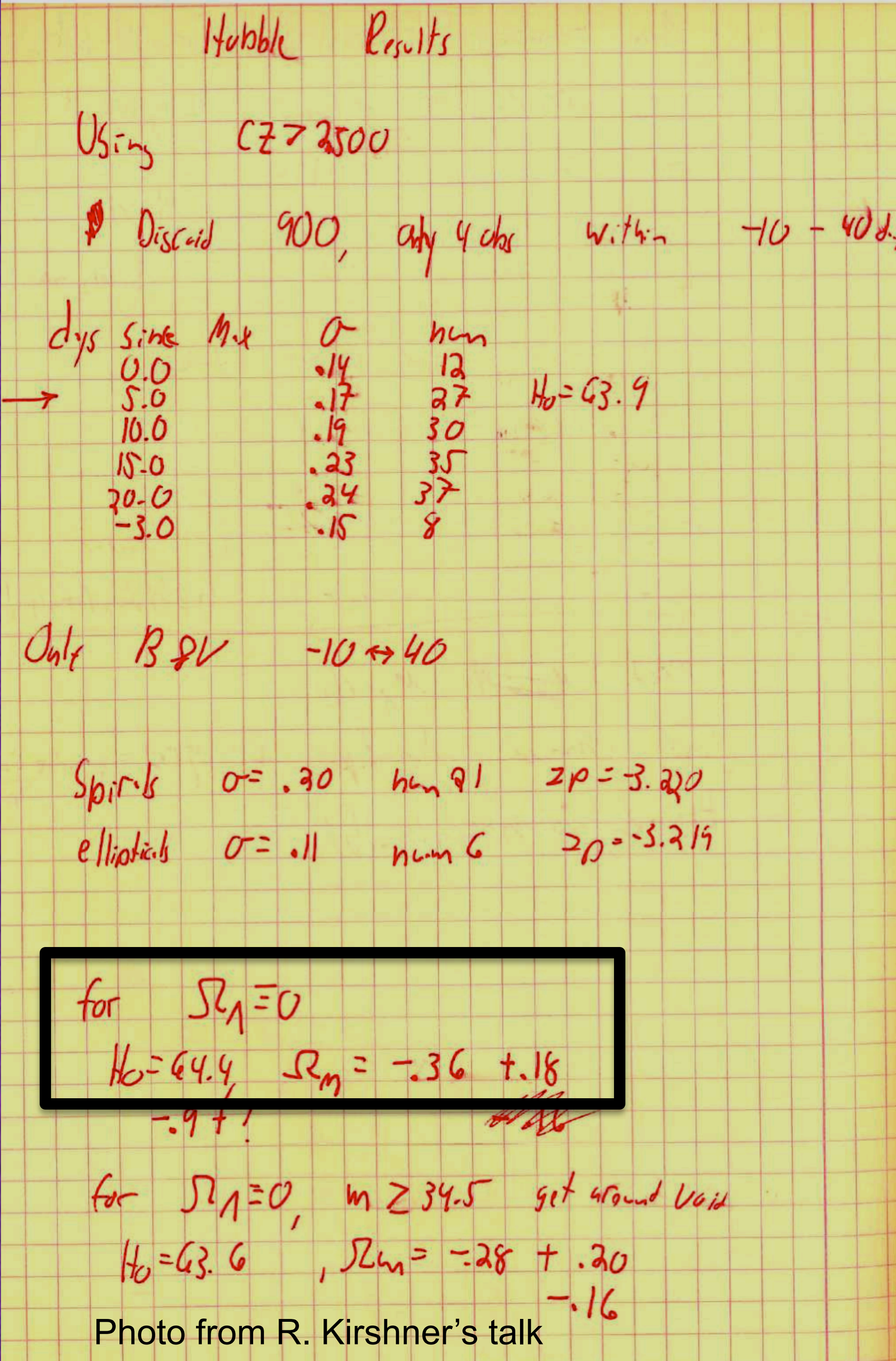
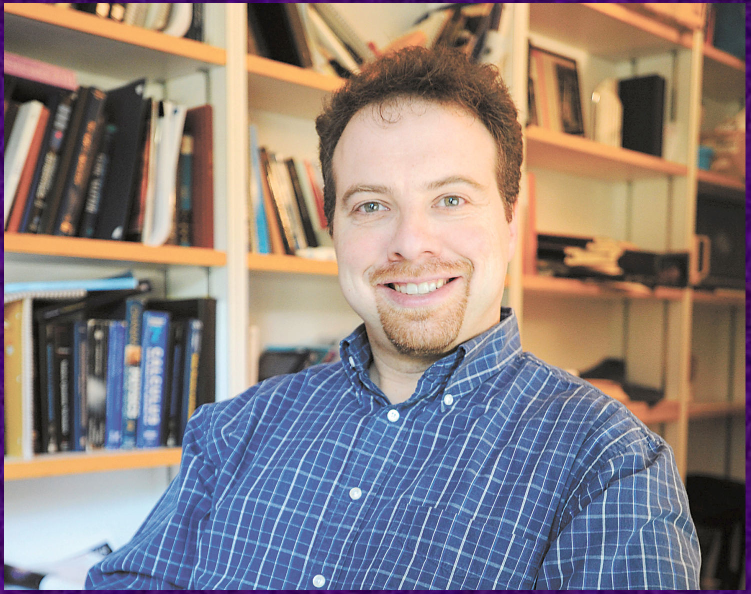
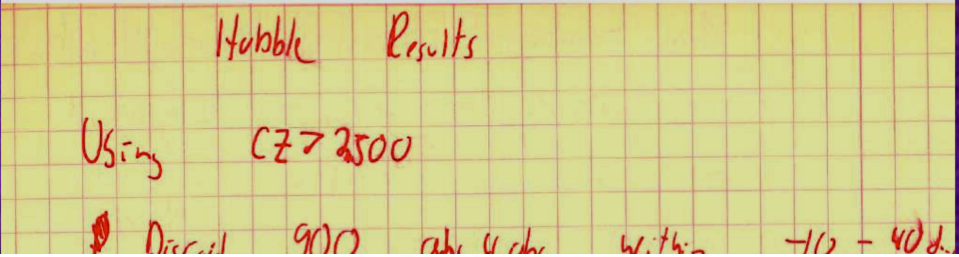
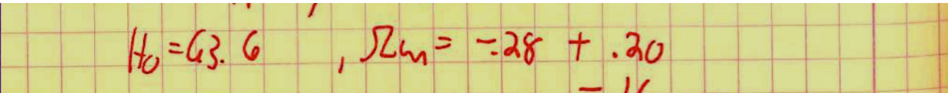


Photo from R. Kirshner’s talk

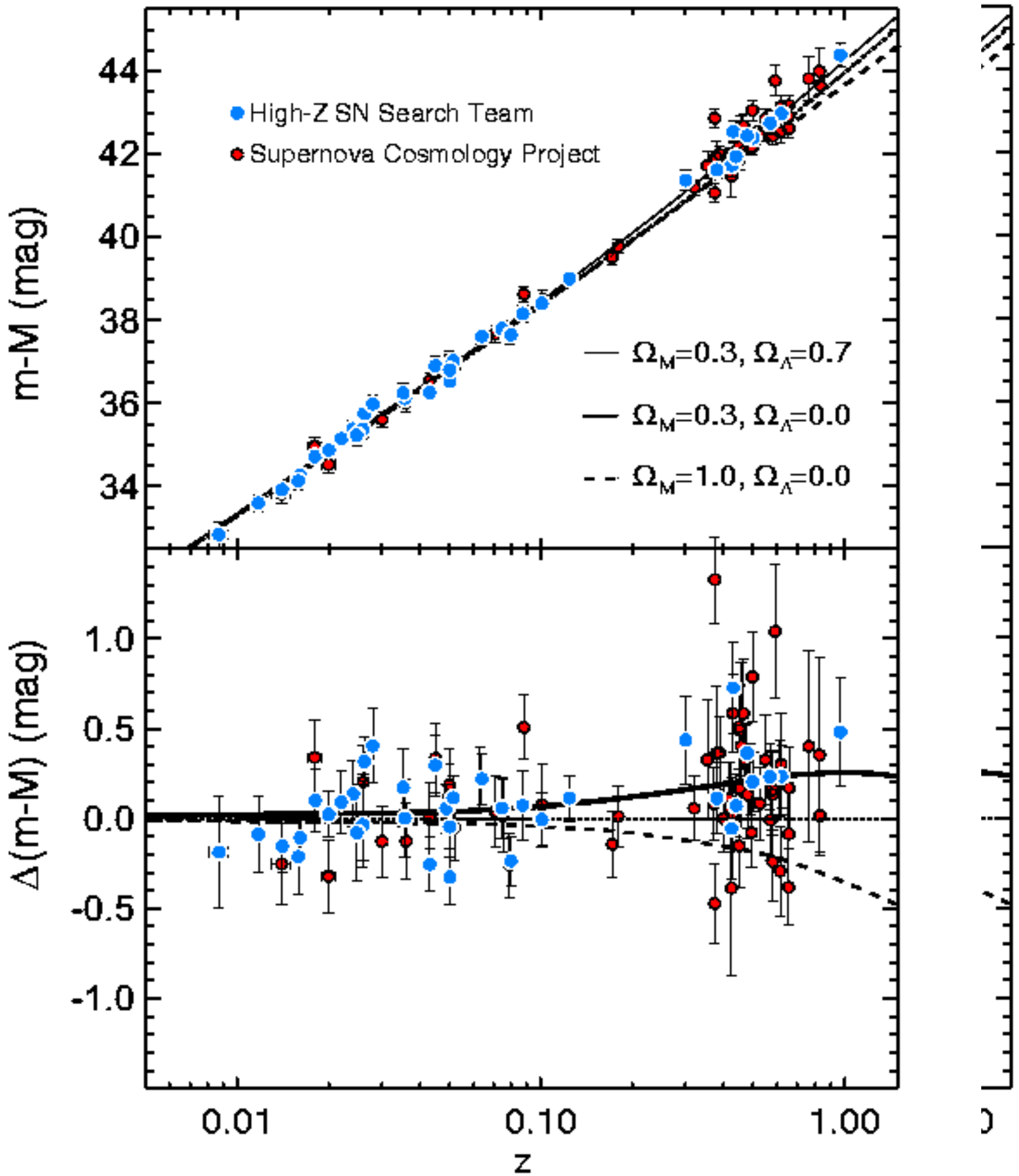


Eureka! or
“What’s wrong with this?”



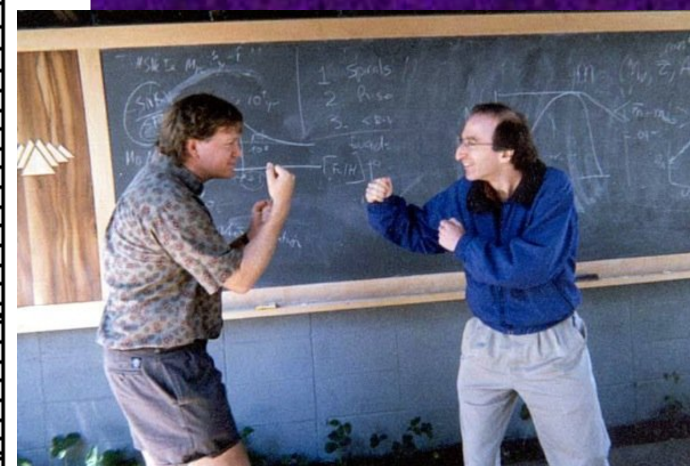


Photos from R. Kirshner’s talk



Riess et al.
Astronomical Journal
1998

Perlmutter et al.
Astrophysical Journal
1999

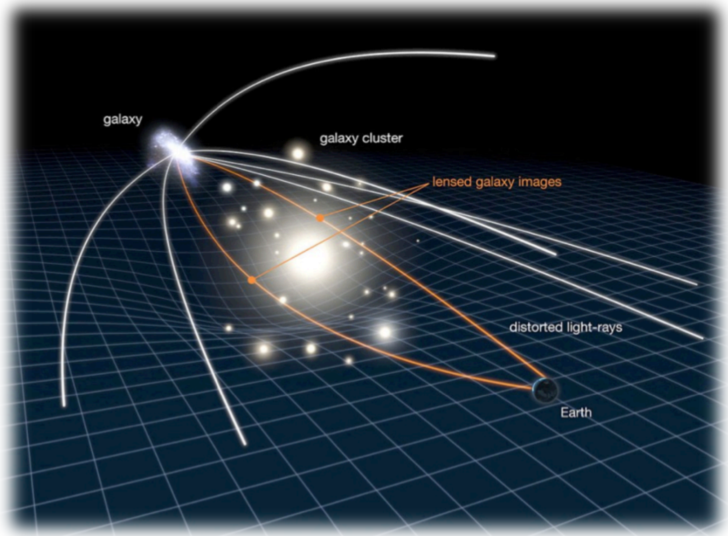


Photos from R. Kirshner’s talk

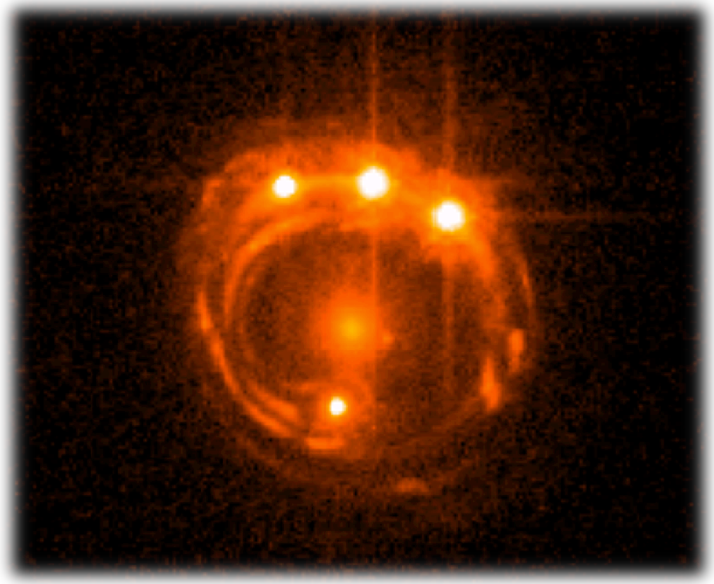
Strong lensing of objects

- Multiple images
- Magnification/Demagnification
- Time delays
  - Different geometrical path
  - Different gravitational potentials

Application for Cosmology ?



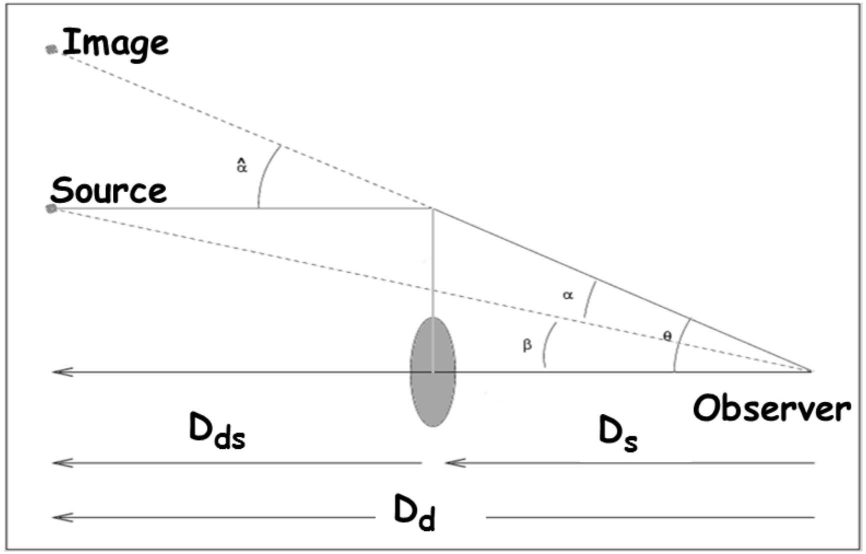
Credit: NASA, ESA, L. Calasada



quasar RXJ1131-123 by HST

Slide from A. Hojjati

Cosmology with strongly lensed variable sources



Need to measure:

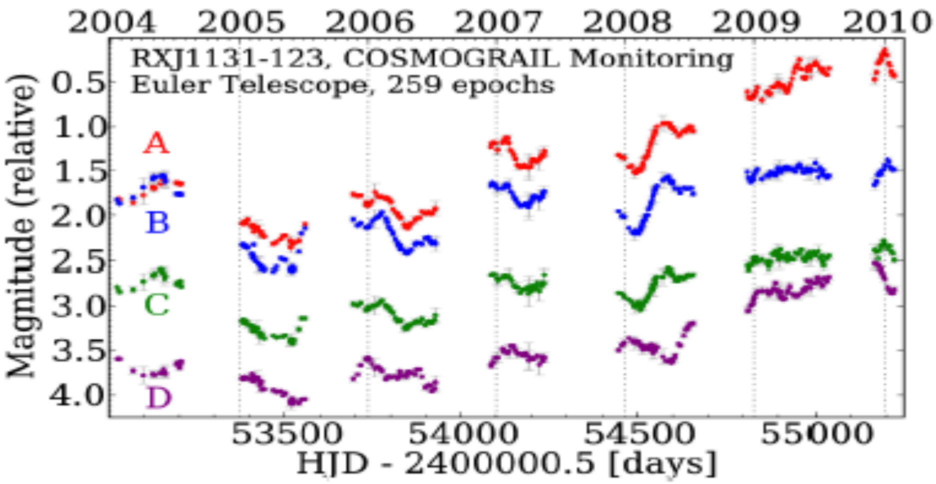
- Lens mass model
- Angular positions
- Mass along LOS
- Time delays

Fermat potential

$$t(\vec{\theta},\vec{\beta}) = \frac{1}{c}\frac{D_{\rm d}D_{\rm s}}{D_{\rm ds}}(1+z_{\rm d})\,\phi(\vec{\theta},\vec{\beta})$$

Geometric delay *Lensing potential*

$$\phi(\vec{\theta},\vec{\beta}) \equiv \left[\frac{(\vec{\theta}-\vec{\beta})^2}{2} - \psi(\vec{\theta})\right]$$



Variable sources :

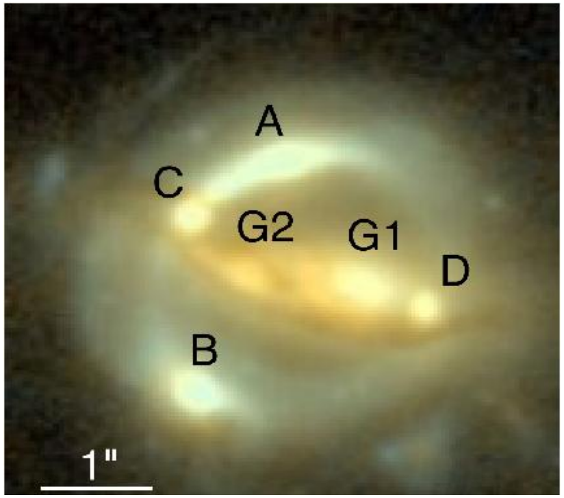
Time delays can be estimated !

Slide from A. Hojjati

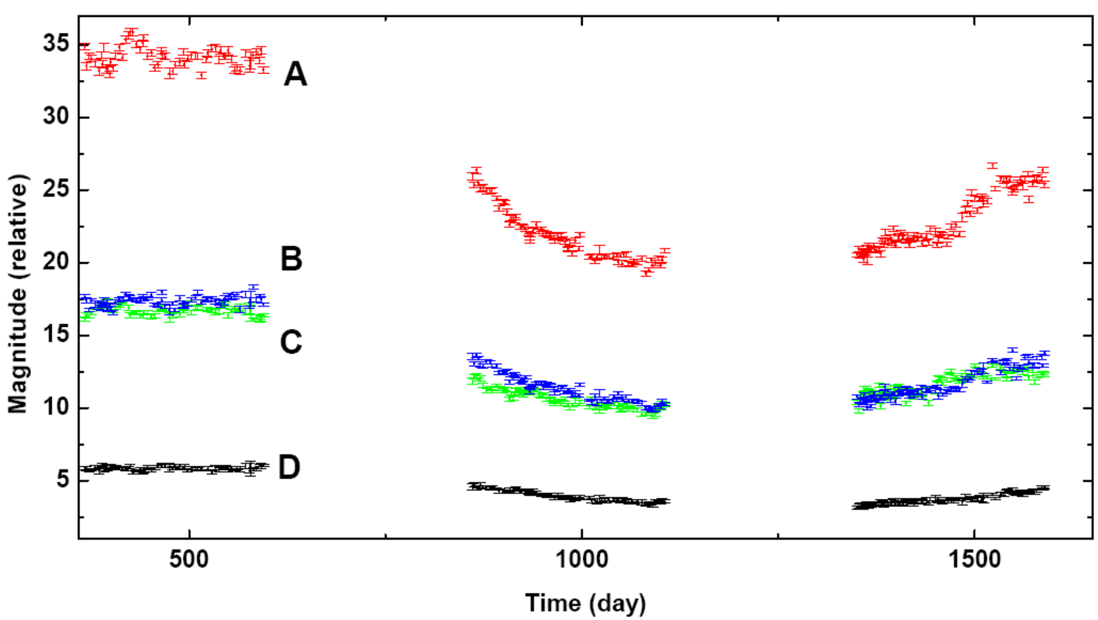
Time delay estimation

- Goal : Reconstructing the shift between multiple streams of data

- Challenges:
  - Measurement noise
  - Seasonal gap
  - Microlensing
  - Flux systematics
  - ...



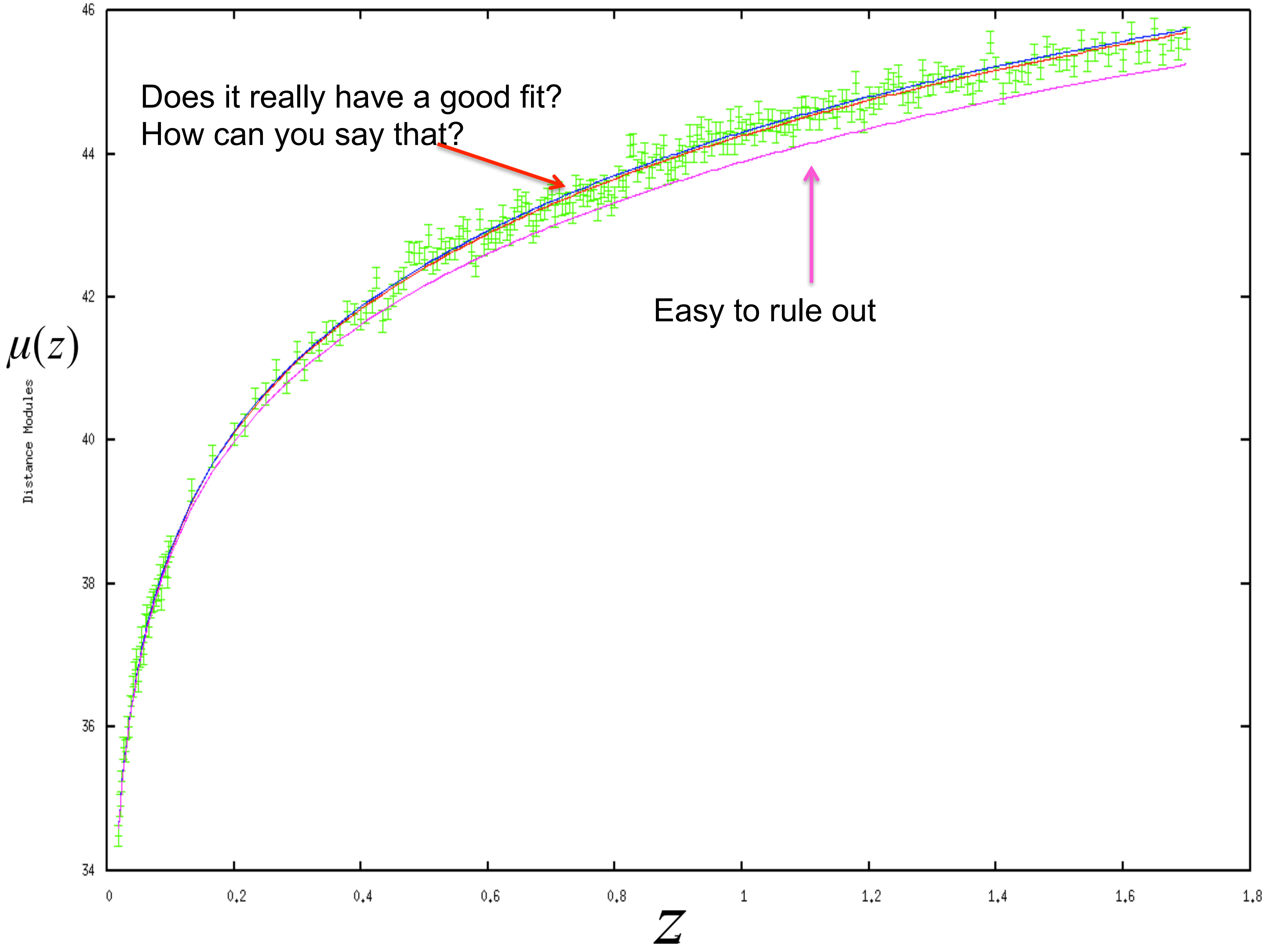
B1608+656

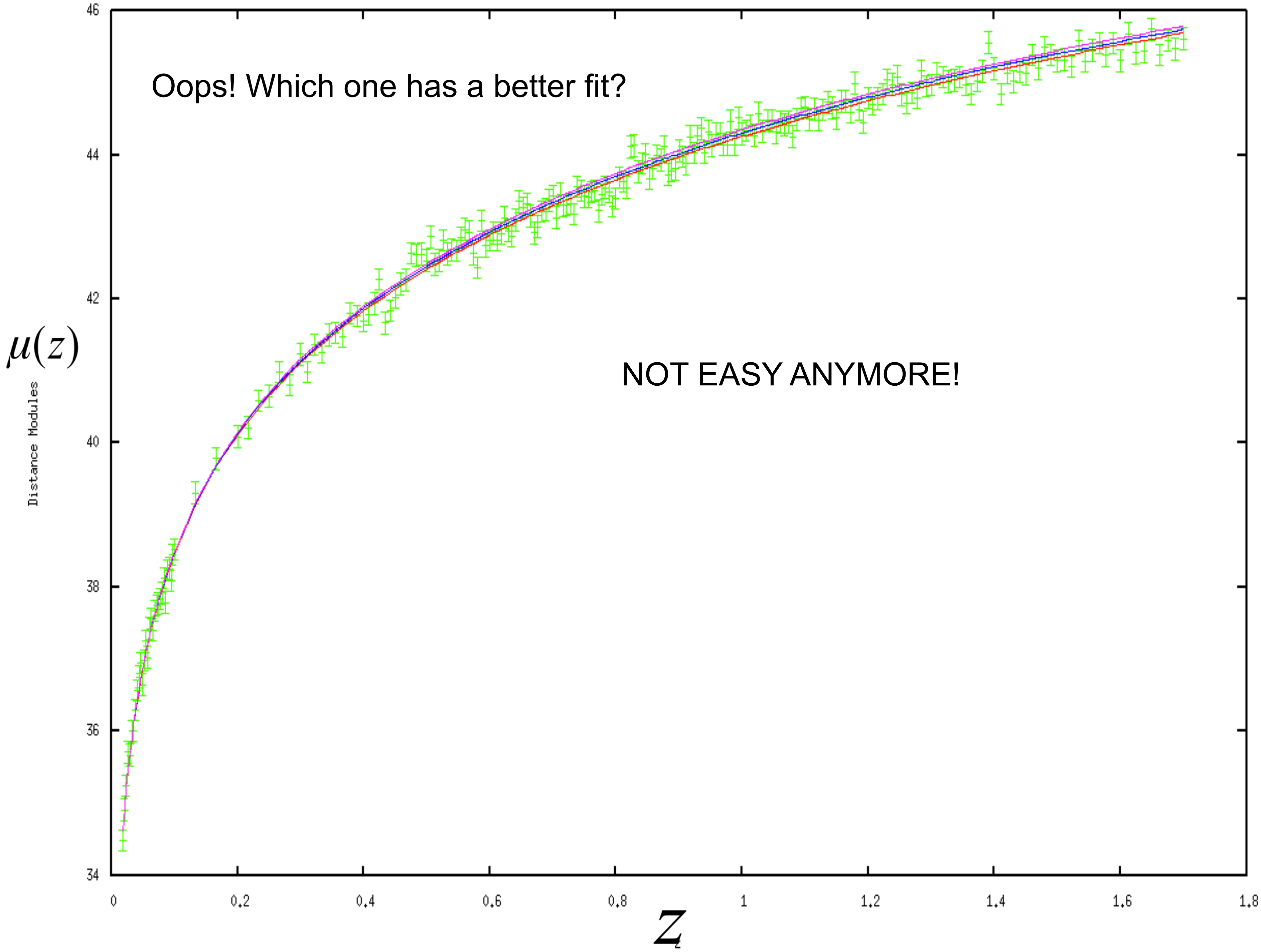


Slide from A. Hojjati

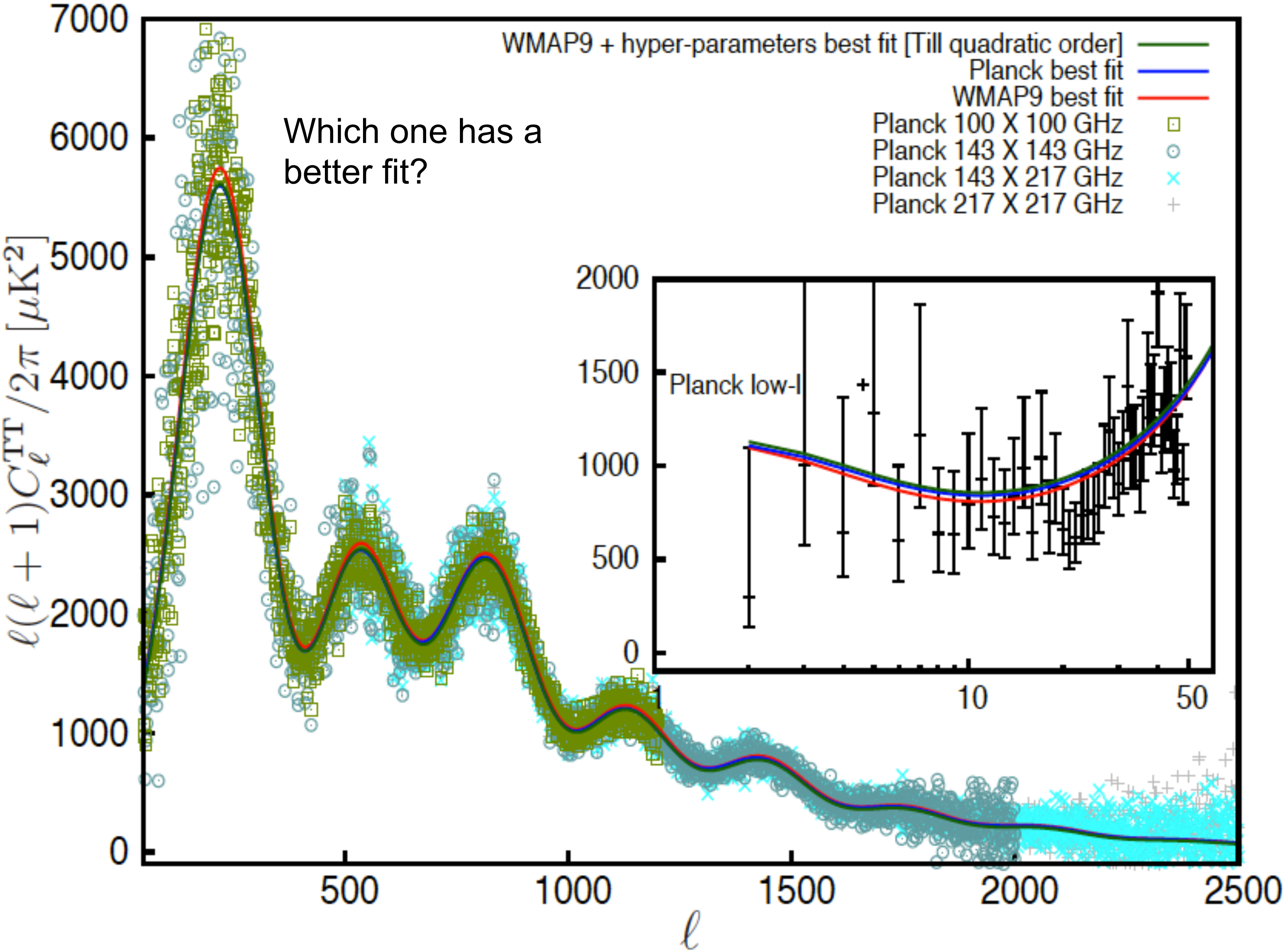
Data Analysis in Cosmology

- Reconstruction and numerical modeling
- Simulation
- Model Selection & Falsification
- Parameter Estimation

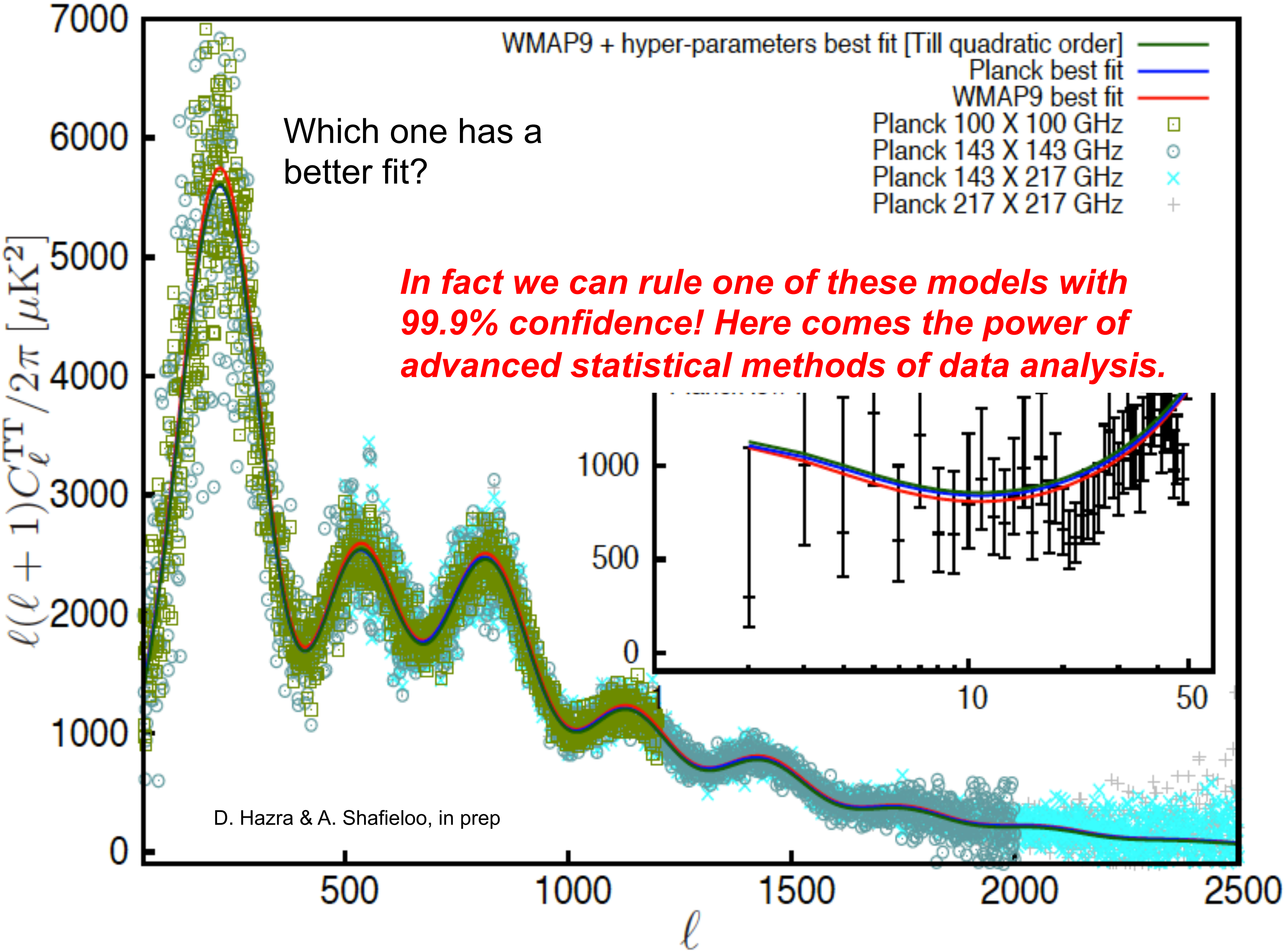

Does it really have a good fit?
How can you say that?
Easy to rule out
$\mu(z)$
Distance Modules
z
46
44
42
40
38
36
34
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8



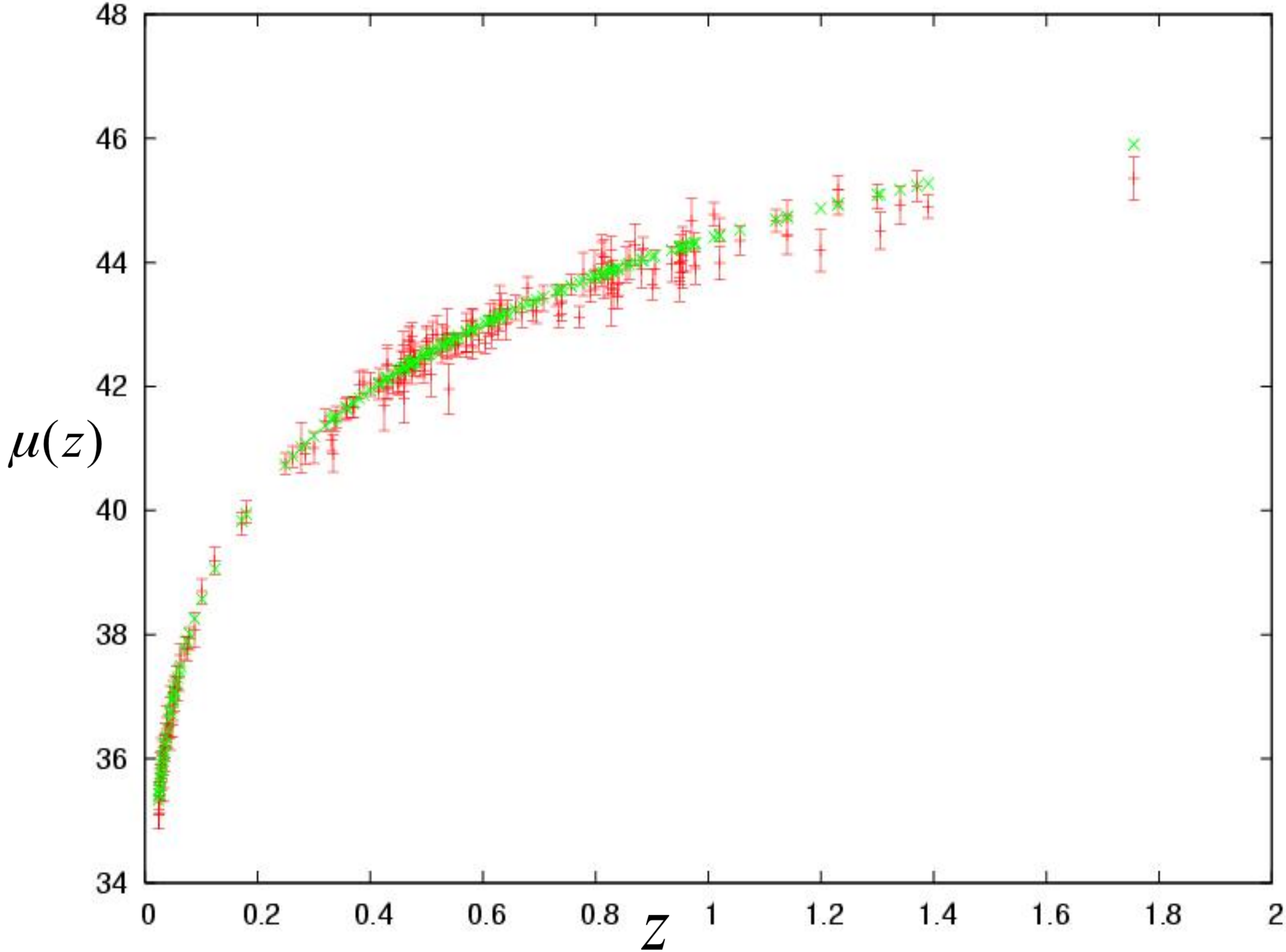
Oops! Which one has a better fit?
NOT EASY ANYMORE!
$\mu(z)$
Distance Modules
z
46
44
42
40
38
36
34
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8



WMAP9 + hyper-parameters best fit [Till quadratic order]
Planck best fit
WMAP9 best fit
Planck 100 X 100 GHz
Planck 143 X 143 GHz
Planck 143 X 217 GHz
Planck 217 X 217 GHz
Which one has a better fit?
Planck low-l
$\ell(\ell+1)C_\ell^{TT}/2\pi\ [\mu K^2]$
ℓ

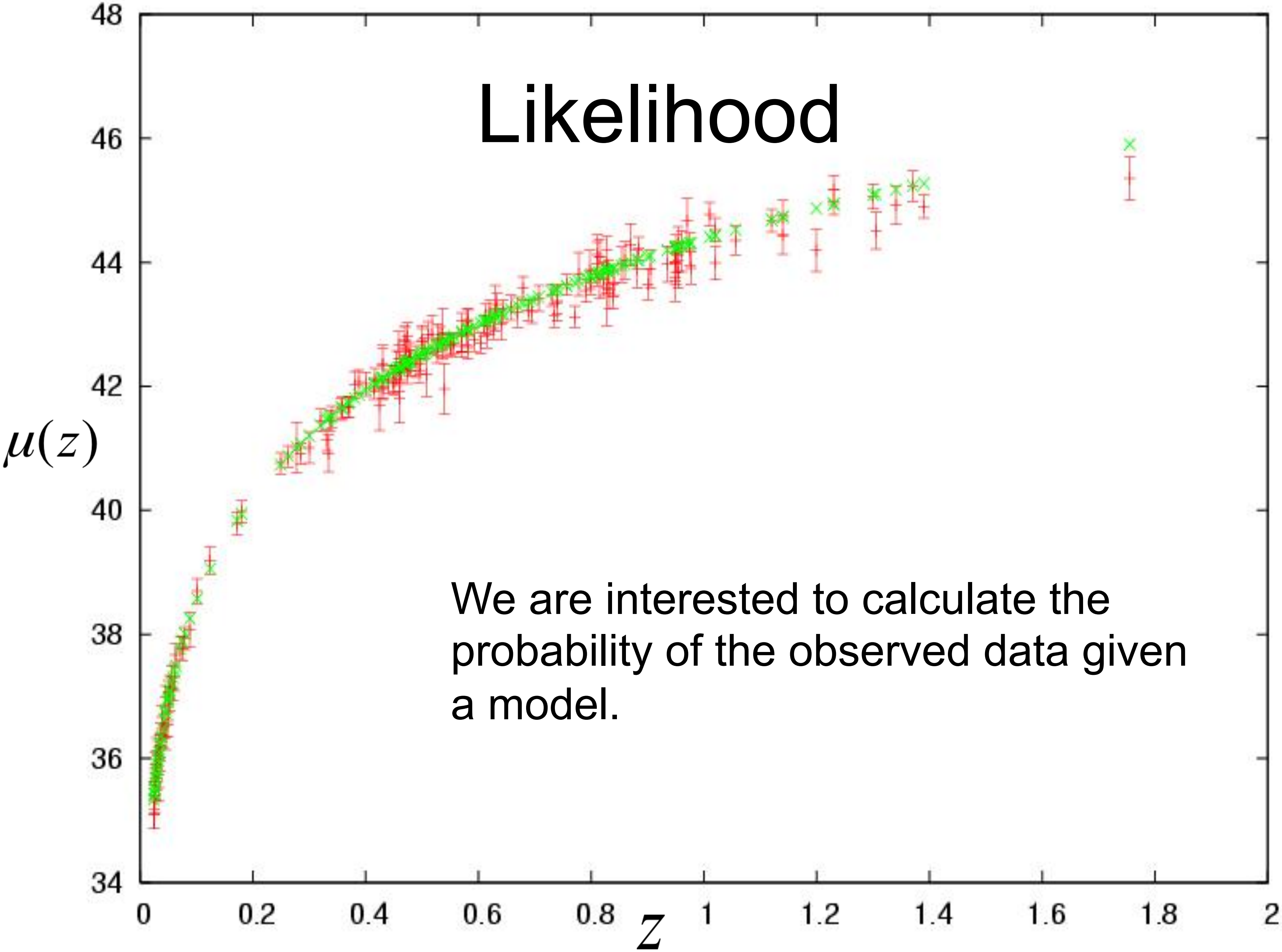


Which one has a better fit?
In fact we can rule one of these models with 99.9% confidence! Here comes the power of advanced statistical methods of data analysis.
WMAP9 + hyper-parameters best fit [Till quadratic order]
Planck best fit
WMAP9 best fit
Planck 100 X 100 GHz
Planck 143 X 143 GHz
Planck 143 X 217 GHz
Planck 217 X 217 GHz
$\ell(\ell+1)C_\ell^{TT}/2\pi\ [\mu K^2]$
ℓ
0
1000
2000
3000
4000
5000
6000
7000
500
1000
1500
2000
2500
10
50
D. Hazra & A. Shafieloo, in prep



48
46
44
42
40
38
36
34
μ(z)
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
z

Likelihood



We are interested to calculate the probability of the observed data given a model.

Likelihood

Minimization of reduced Chi square or effective Chi square is the most common approach in cosmology (and many other fields of science) to do parameter estimation and also model selection.

$$\chi^2 = \sum_i^N \frac{(\mu_i^t - \mu_i^e)^2}{\sigma_i^2},$$

Likelihood

We are interested to calculate the probability of the observed data given the model.

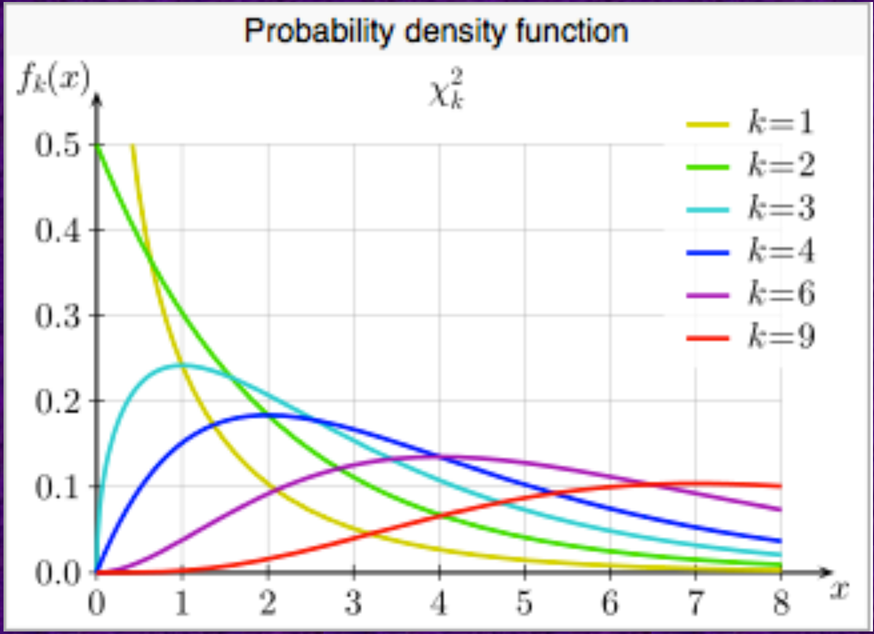
$$\chi^2 = \sum_i^N (\mu_i^t - \mu_i^e)^T Cov^{-1} (\mu_i^t - \mu_i^e)$$

$$\chi^2 = \sum_i^N \frac{(\mu_i^t - \mu_i^e)^2}{\sigma_i^2},$$

When data is uncorrelated

$$P(\chi^2; N) = \frac{2^{-N/2}}{\Gamma(N/2)} \chi^{N-2} e^{-\chi^2/2}$$

$$Prob(\chi^2; N) = \int_{\chi^2}^{\infty} P(\chi^2; N) d\chi'^2.$$



Likelihood and Model Fitting

When number of data points is more than ~30 one can use relative chi square for likelihood analysis and N, number of free parameters of the fitting function, will become the degrees of freedom.

$$\chi^2 = \sum_i^N \frac{(\mu_i^t - \mu_i^e)^2}{\sigma_i^2},$$

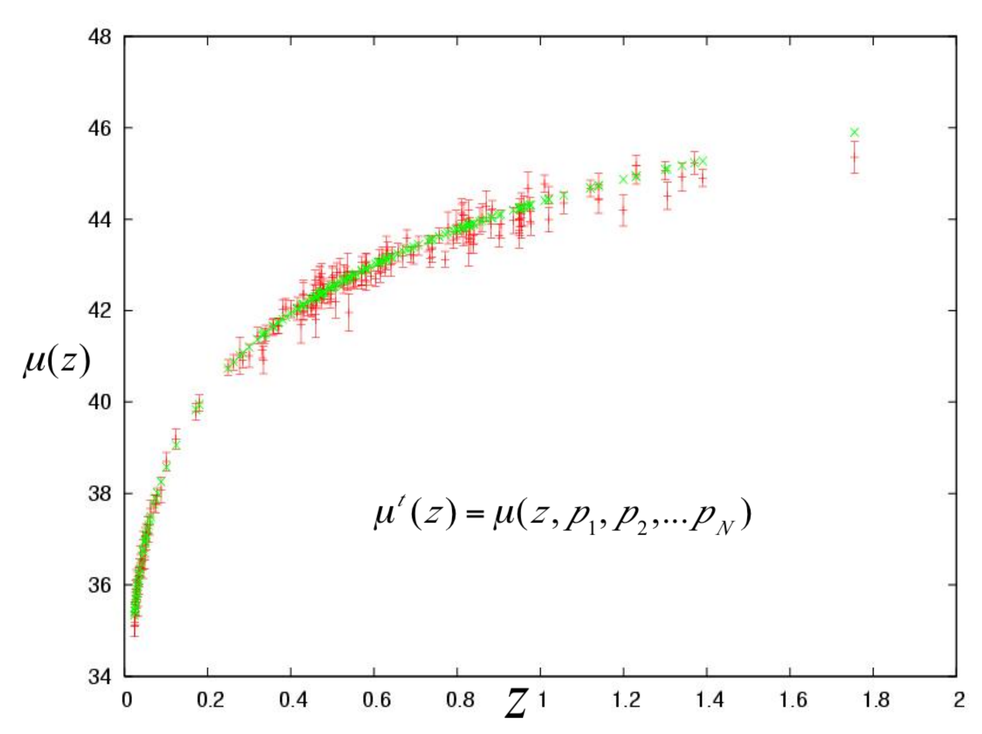
In likelihood estimation:

$$\chi^2 \longrightarrow \Delta\chi^2$$

$$\Delta\chi^2 = \chi^2 - \chi^2_{best}$$

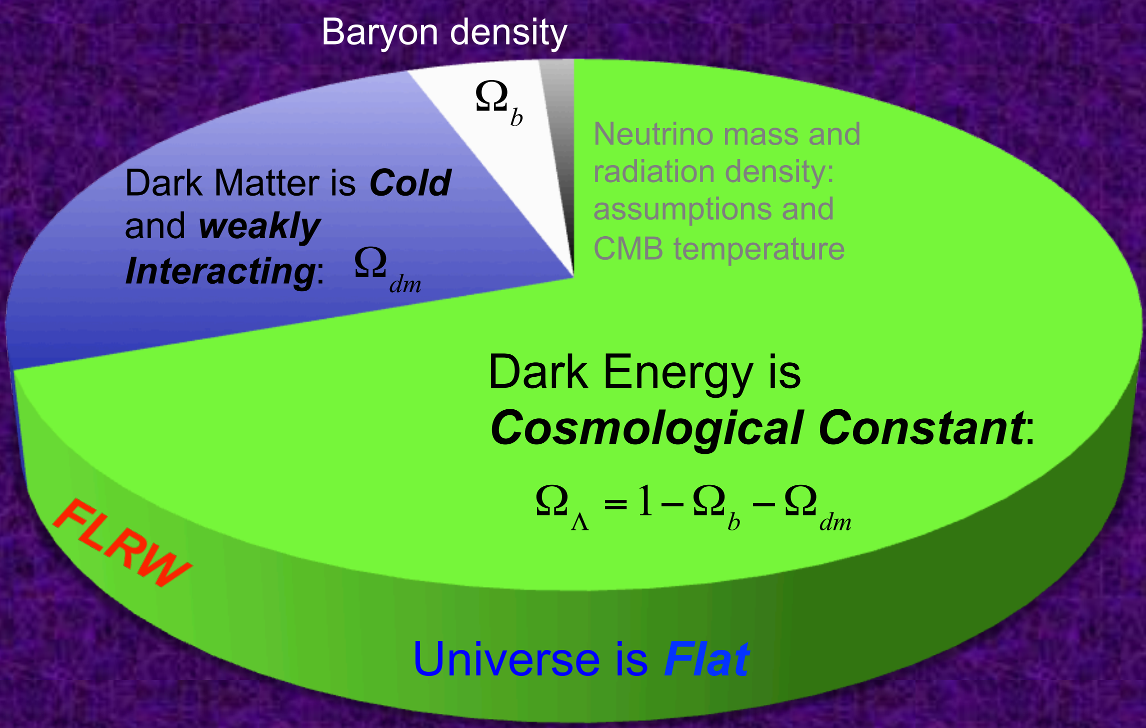
$$P(\chi^2; N) = \frac{2^{-N/2}}{\Gamma(N/2)} \chi^{N-2} e^{-\chi^2/2}$$

$$Prob(\chi^2; N) = \int_{\chi^2}^{\infty} P(\chi^2; N) d\chi'^2.$$



Standard Model of Cosmology

Using measurements and statistical techniques to place sharp constraints on parameters of the standard cosmological model.



Initial Conditions: Form of the Primordial Spectrum is ***Power-law***

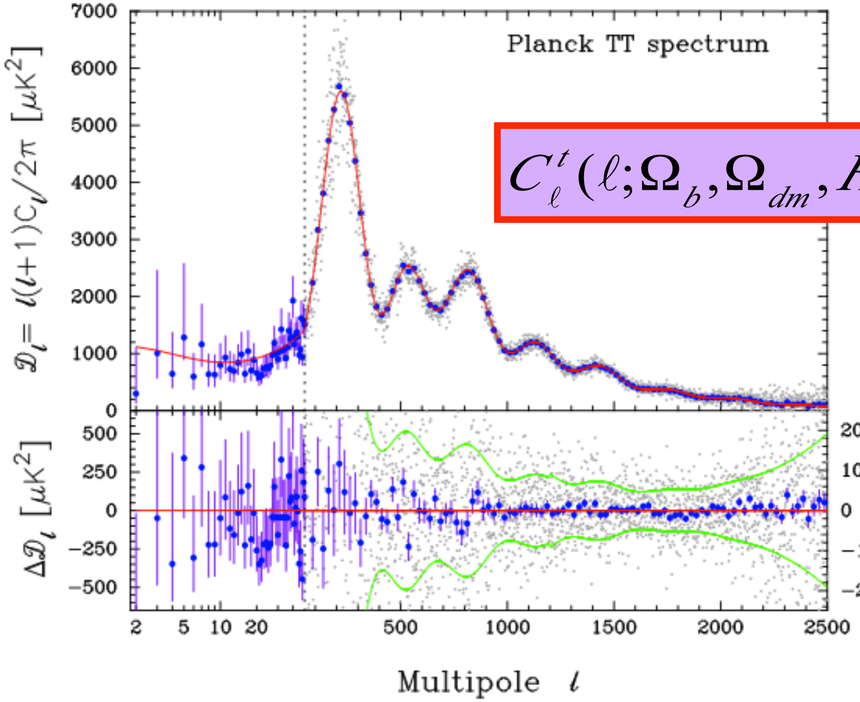
n_s, A_s

Epoch of reionization

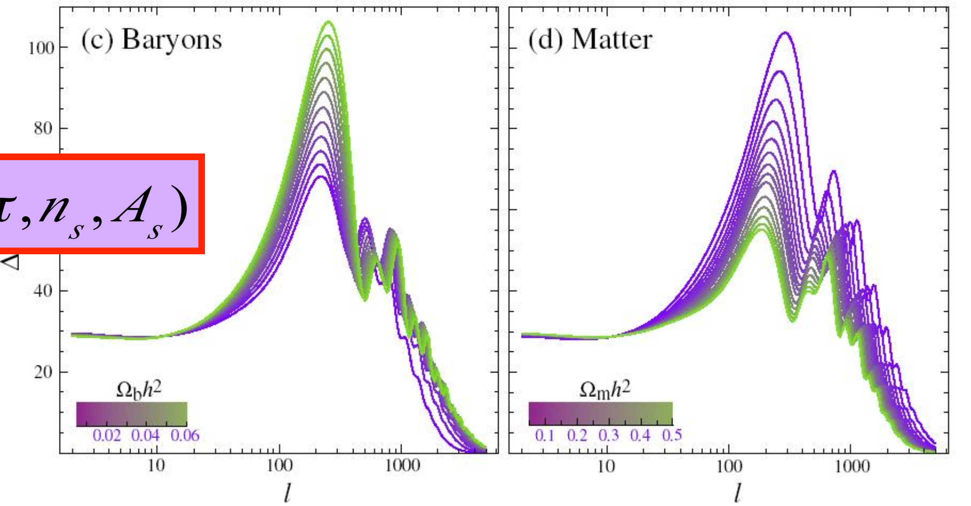
τ

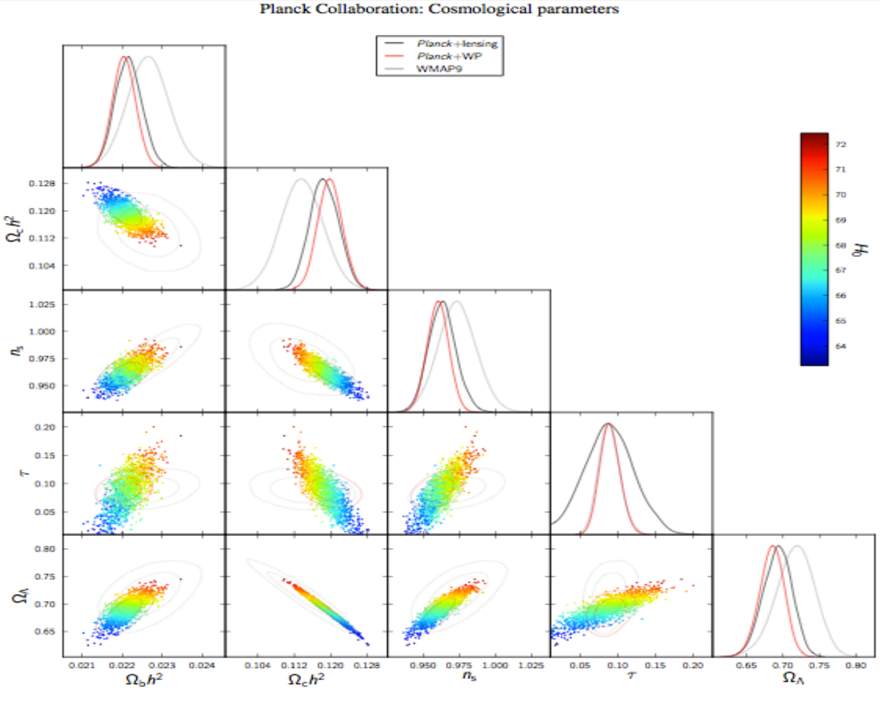
Hubble Parameter and the Rate of Expansion

H_0

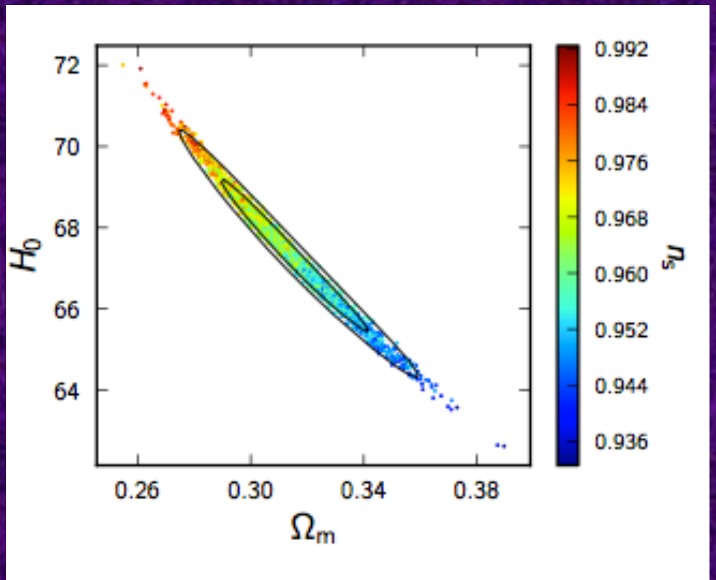


$$C_\ell^t(\ell;\Omega_b,\Omega_{dm},H_0,\tau,n_s,A_s)$$





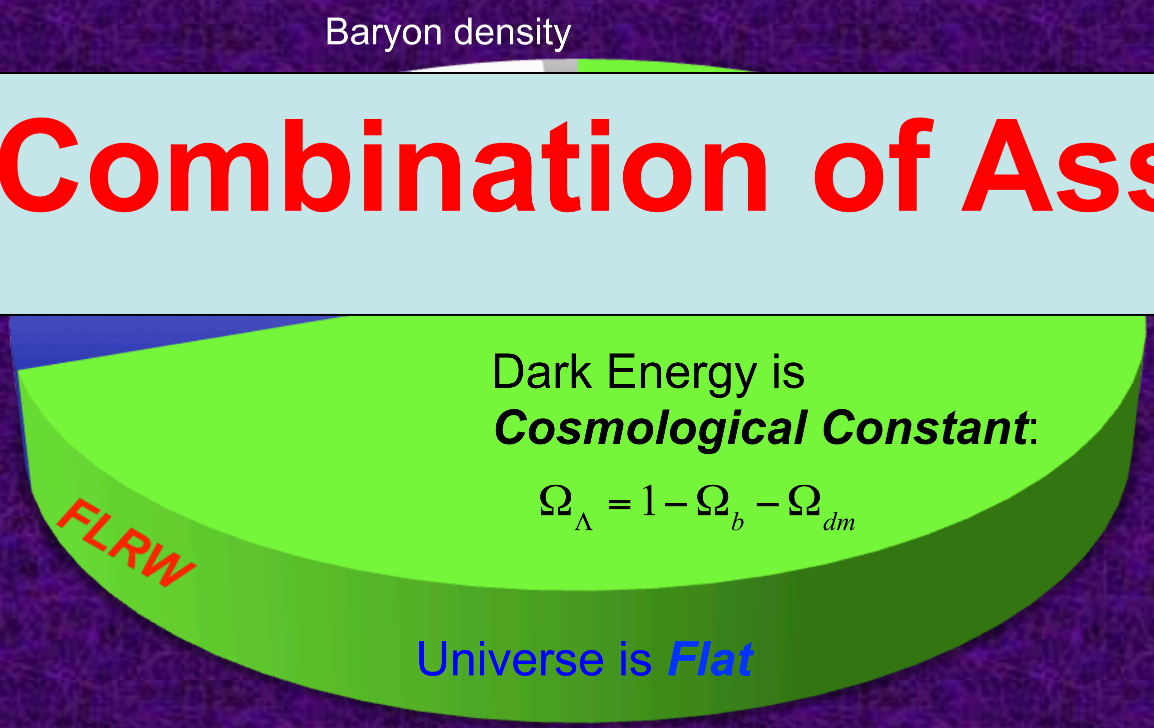
$$\chi^2 = \sum_{\ell=2}^{N}(C_\ell^t - C_\ell^e)^T Cov^{-1}(C_\ell^t - C_\ell^e)$$



Standard Model of Cosmology

Using measurements and statistical techniques to place sharp constraints on parameters of the standard cosmological model.

Combination of Assumptions



Epoch of reionization

τ

Hubble Parameter and the Rate of Expansion

H_0

Consistency of a model and the data:

Frequentist Approach:

Assuming a proposed model, the probability of the observed data must not be insignificant.

Bayesian Approach:

Priors and simplicity of the proposed model also matters (in model comparison)

Chi square analysis plays a crucial role in calculation of the likelihood in both approaches

Bayesian Analysis

- Bayesian approach provides the means to incorporate ***prior*** knowledge in data analysis.
- Bayes's law states that the posterior probability is proportional to the product of the likelihood and the prior probability.

Posterior probability and the priors:

Likelihood

Prior probability

$$p(x|d) = \frac{p(d|x)p(x)}{p(d)}$$

Posterior probability

Normalization factor

Posterior probability and the priors:

Likelihood

Prior probability

$$p(x|d) = \frac{p(d|x)p(x)}{p(d)}$$

Posterior probability

Normalization factor

Model fitting has Bayesian essence since we assume that we are considering a correct model. What if we are wrong?

Reconstruction

To find cosmological quantities and parameters there are two general approaches:

1. Parametric methods

   Easy to confront with cosmological observations to put constrains on the parameters, but the results are highly biased by the assumed models and parametric forms.

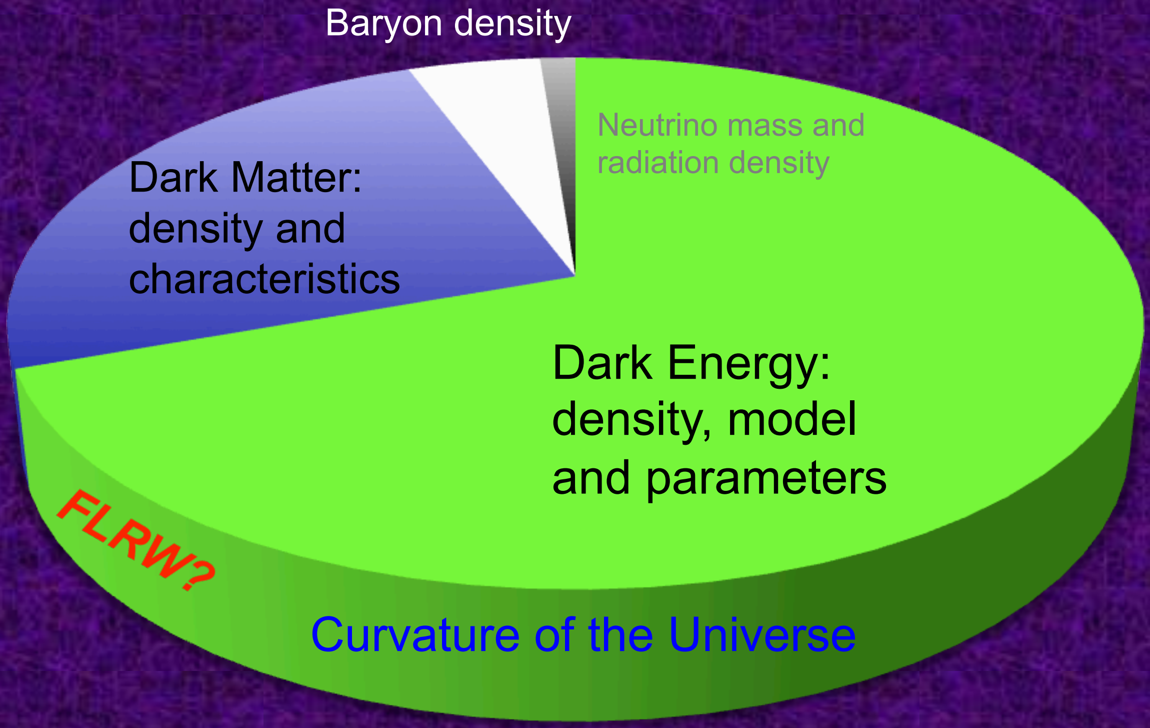
2. Non Parametric methods

   Difficult to apply on the raw data, but the results will less biased and more reliable and independent of theoretical models or parametric forms.

.

Era of Precision Cosmology

Combining new measurements and using statistical techniques to place sharp constraints on cosmological models and their parameters.



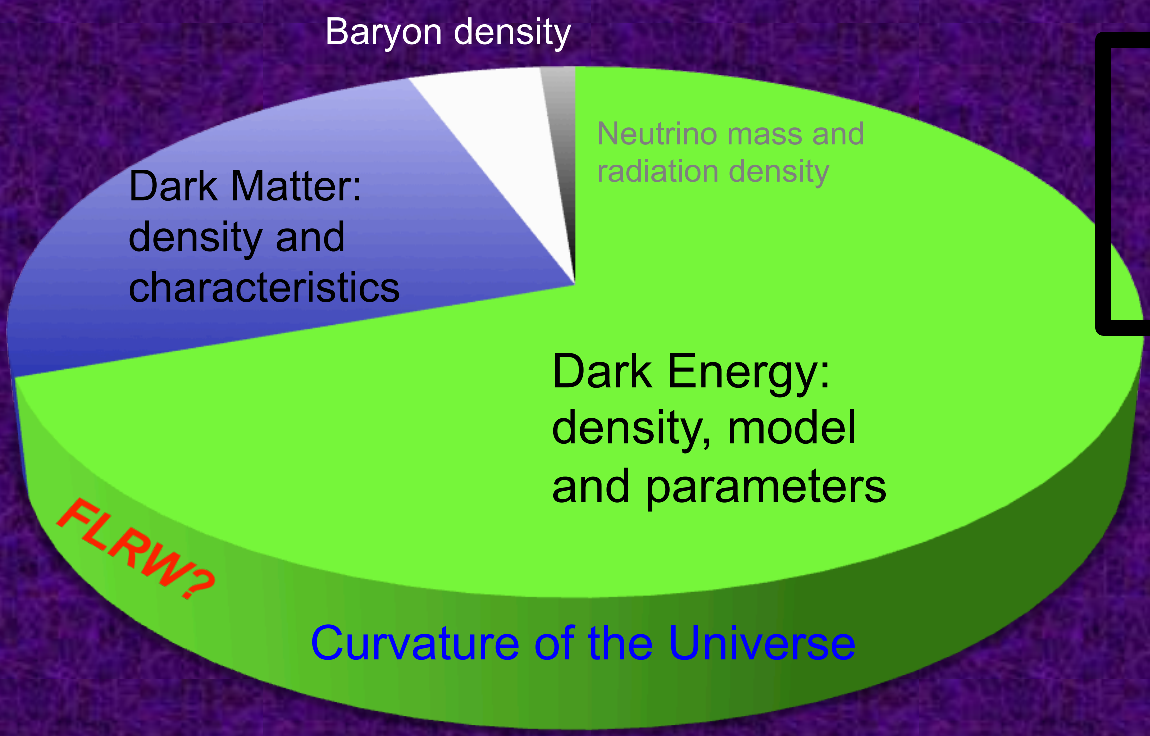
Initial Conditions:
Form of the Primordial Spectrum and Model of Inflation and its Parameters

Epoch of reionization

Hubble Parameter and the Current Rate of Expansion

Era of Precision Cosmology

Combining new measurements and using statistical techniques to place sharp constraints on cosmological models and their parameters.



Initial Conditions:
Form of the Primordial Spectrum and Model of Inflation and its Parameters

Epoch of reionization

Hubble Parameter and the Current Rate of Expansion

Parameter estimation within a cosmological framework

Harisson-Zel'dovich (HZ)

WMAP Cosmological Parameters Model: lcdm+ns=1 Data: wmap	
$10^2\Omega_b h^2$	$2.405^{+0.046}_{-0.047}$
$\Delta^2_{\mathcal{R}}(k=0.002/\mathrm{Mpc})$	$(23.1 \pm 1.2) \times 10^{-10}$
h	0.778 ± 0.032
H_0	77.8 ± 3.2 km/s/Mpc
$\Omega_b h^2$	$0.02405^{+0.00046}_{-0.00047}$
Ω_Λ	0.788 ± 0.031
Ω_m	0.212 ± 0.031
$\Omega_m h^2$	$0.1271^{+0.0086}_{-0.0087}$
σ_8	$0.796^{+0.053}_{-0.054}$
A_{SZ}	$0.92^{+0.63}_{-0.61}$
t_0	13.353 ± 0.096 Gyr
τ	0.141 ± 0.029
θ_A	0.5986 ± 0.0017 °
z_r	14.6 ± 2.0

Tables from NASA - LAMBDA website

Power-Law (PL)

WMAP Cosmological Parameters Model: lcdm Data: wmap	
$10^2\Omega_b h^2$	2.229 ± 0.073
$\Delta^2_{\mathcal{R}}(k=0.002/\mathrm{Mpc})$	$(23.5 \pm 1.3) \times 10^{-10}$
h	$0.732^{+0.031}_{-0.032}$
H_0	$73.2^{+3.1}_{-3.2}$ km/s/Mpc
$\log(10^{10} A_s)$	3.156 ± 0.056
$n_s(0.002)$	0.958 ± 0.016
$\Omega_b h^2$	0.02229 ± 0.00073
$\Omega_c h^2$	$0.1054^{+0.0078}_{-0.0077}$
Ω_Λ	0.759 ± 0.034
Ω_m	0.241 ± 0.034
$\Omega_m h^2$	$0.1277^{+0.0080}_{-0.0079}$
σ_8	$0.761^{+0.049}_{-0.048}$
τ	0.089 ± 0.030
θ_A	0.5952 ± 0.0021 °
z_r	$11.0^{+2.6}_{-2.5}$

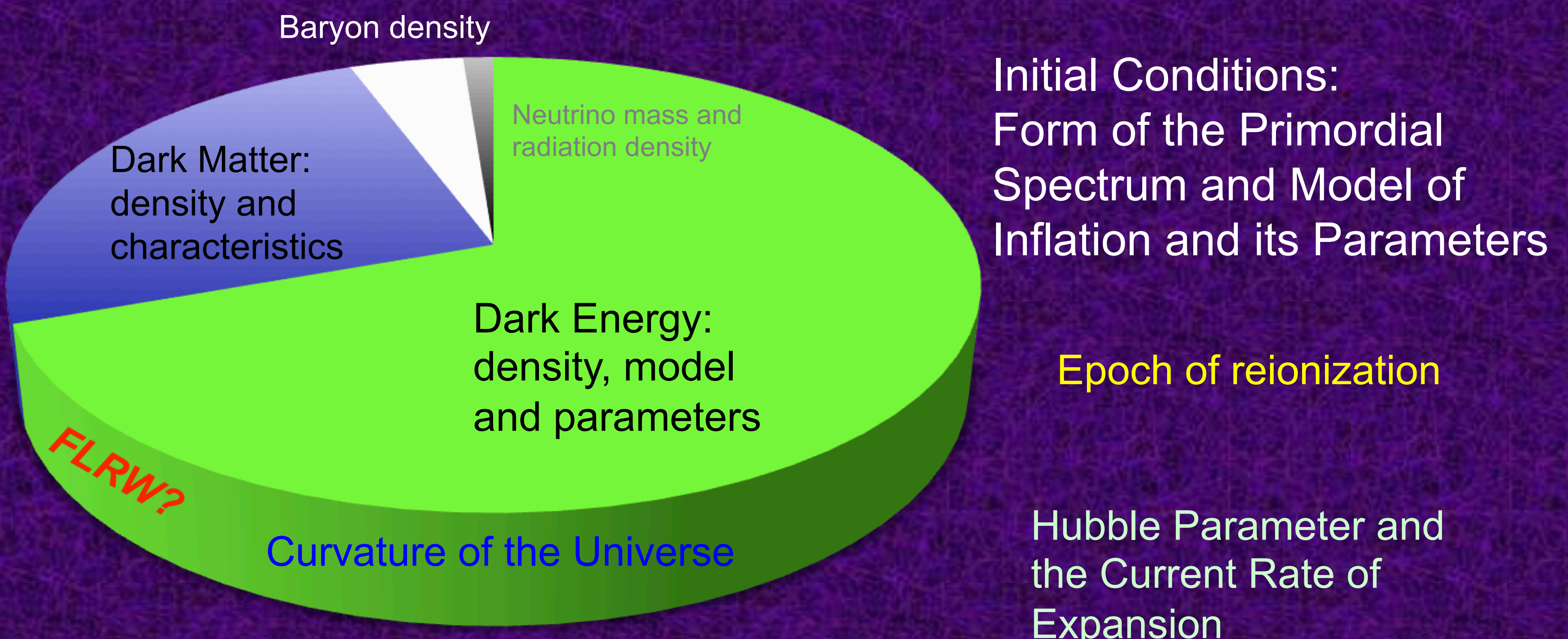
PL with Running (RN)

WMAP Cosmological Parameters Model: lcdm+run Data: wmap	
$10^2\Omega_b h^2$	2.10 ± 0.10
$\Delta^2_{\mathcal{R}}(k=0.002/\mathrm{Mpc})$	$(23.9 \pm 1.3) \times 10^{-10}$
$dn_s/d\ln k$	$-0.055^{+0.030}_{-0.031}$
h	$0.681^{+0.042}_{-0.041}$
H_0	$68.1^{+4.2}_{-4.1}$ km/s/Mpc
$n_s(0.002)$	$1.050^{+0.059}_{-0.058}$
$\Omega_b h^2$	0.0210 ± 0.0010
Ω_Λ	$0.703^{+0.056}_{-0.055}$
Ω_m	$0.297^{+0.055}_{-0.056}$
$\Omega_m h^2$	$0.1350^{+0.0099}_{-0.0097}$
σ_8	$0.771^{+0.051}_{-0.050}$
A_{SZ}	$1.06^{+0.62}_{-0.65}$
t_0	13.97 ± 0.20 Gyr
τ	0.101 ± 0.031
θ_A	0.5940 ± 0.0021 °
z_r	12.8 ± 2.8

Functional parameterizations affect estimation of cosmological parameters

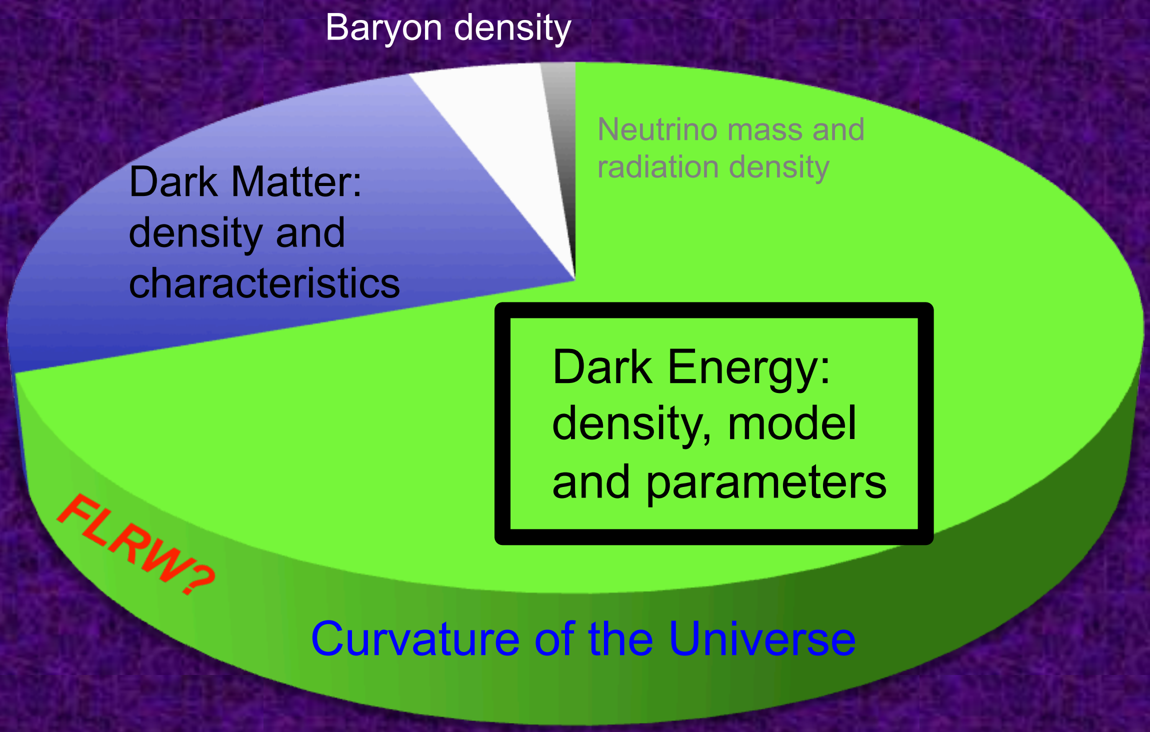
Era of Precision Cosmology

Combining new measurements and using statistical techniques to place sharp constraints on cosmological models and their parameters.



Era of Precision Cosmology

Combining new measurements and using statistical techniques to place sharp constraints on cosmological models and their parameters.



Initial Conditions:
Form of the Primordial Spectrum and Model of Inflation and its Parameters

Epoch of reionization

Hubble Parameter and the Current Rate of Expansion

Probes of Dark Energy

- ***Standard candles***:
  measure luminosity distance.

$$F = \frac{L}{4\pi d_L^2}$$

Supernovae Ia as Standardized Candles

- ***Standard rulers***:
  measure angular diameter distance.

$$\Delta\theta = \frac{\Delta\chi}{d_A(z)}$$

BAO as standard ruler

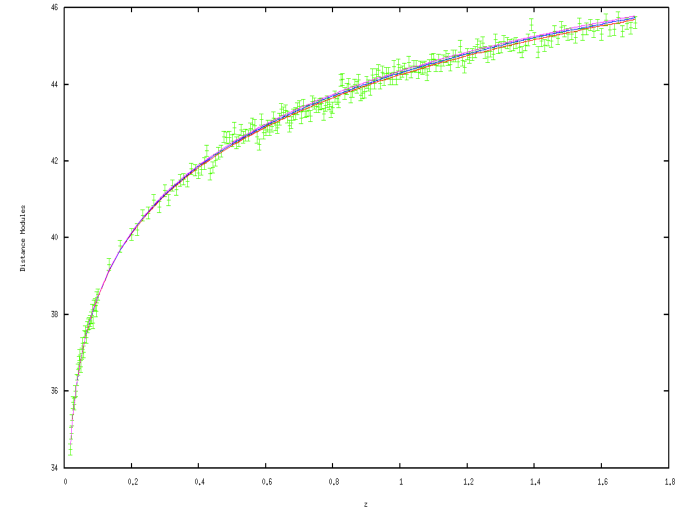
- ***Growth of fluctuations***:
  testing modified gravity or to distinguish between physical and geometrical models of Dark Energy.

$$\ddot{\delta} + 2H\dot{\delta} - 4\pi G_{\rm eff}\rho\delta = 0,$$

Dark Energy Models

- Cosmological Constant
- Quintessence and k-essence (scalar fields)
- Exotic matter (Chaplygin gas, phantom, etc.)
- Braneworlds (higher-dimensional theories)
- Modified Gravity
- ……

But which one is really responsible for the acceleration of the expanding universe?!



Dark Energy Parameterizations

$$F = \frac{L}{4\pi d_L^2}$$

Supernovae Ia as Standardized Candles

$$\Delta\theta = \frac{\Delta\chi}{d_A(z)}$$

BAO as standard ruler

$$d_L(z) = (1+z)\int_0^z \frac{dz'}{H(z')}$$

$$d_A(z) = (1+z)^{-1}\int_0^z \frac{dz'}{H(z')}$$

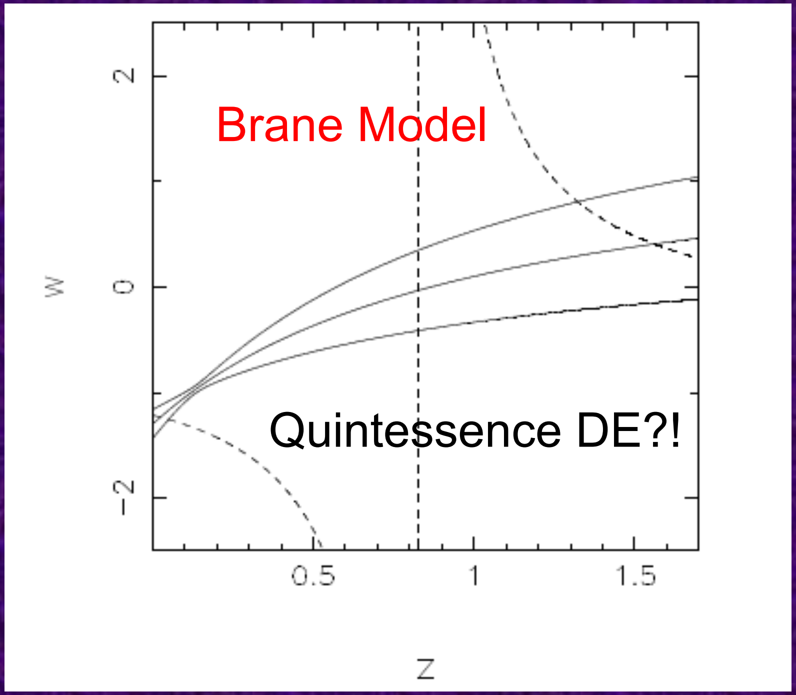
$$\frac{H^2(z)}{H^2_0} = \left[\Omega_{0M}(1+z)^3 + (1-\Omega_{0M})X(z)\right]$$

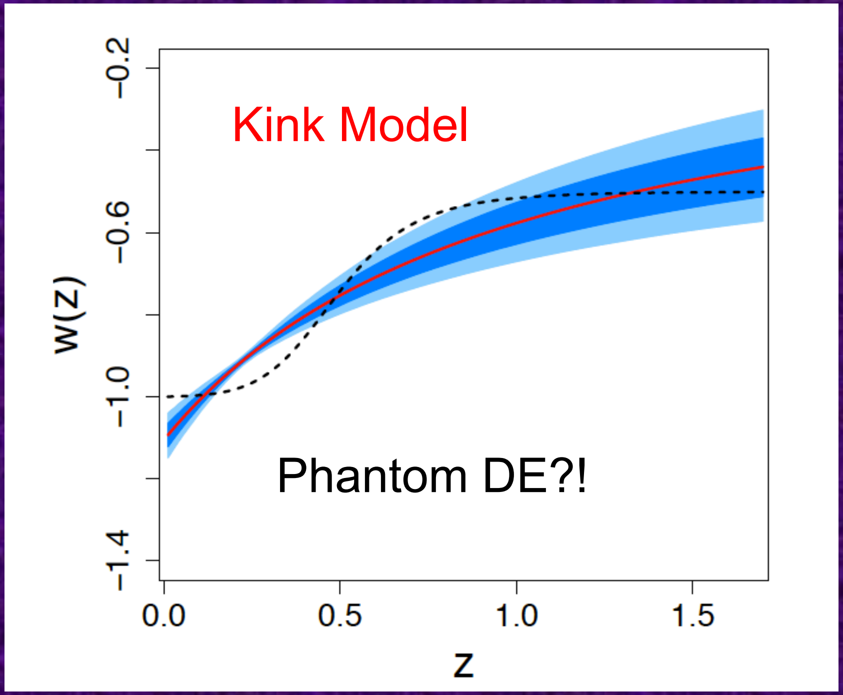
1. Fitting functions for d_l(z)
2. Fitting functions for DE density
3. Fitting functions for EOS

Most general form

$$\frac{H^2(z)}{H^2_0} = \left[\Omega_{0M}(1+z)^3 + (1-\Omega_{0M})\exp[\int 3(1+w(z))\frac{dz}{1+z}]\right]$$

Problems of Dark Energy Parameterizations (model fitting)





Shafieloo, Alam, Sahni & Starobinsky, MNRAS 2006

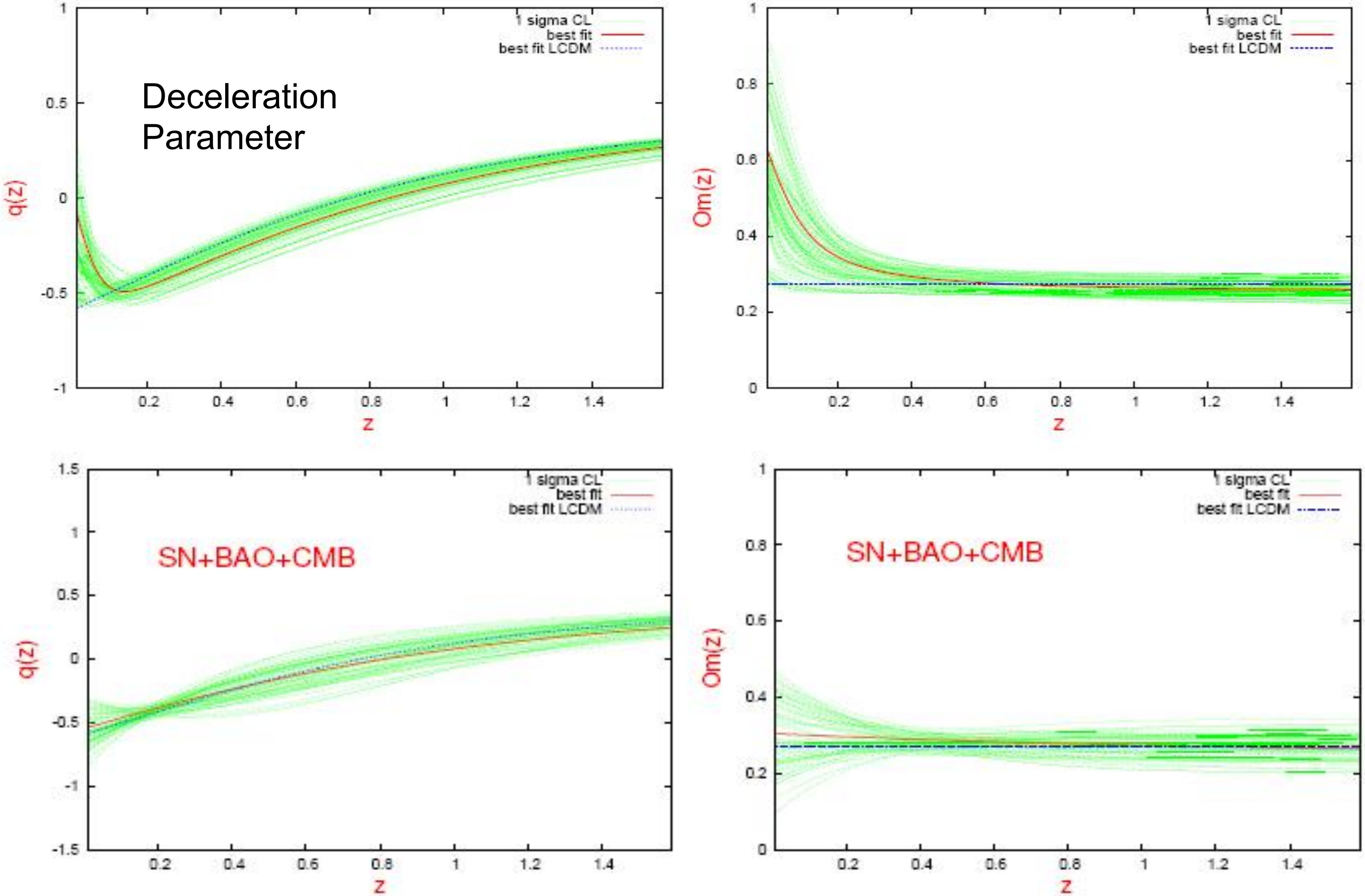
Holsclaw et al, PRD 2011

$$w(z) = w_0 - w_a \frac{z}{1+z}.$$

Chevallier-Polarski-Linder ansatz (CPL)..

Same data being used. Which one is correct?

Problems of model fitting



Non Parametric methods of Reconstruction

Usually involves binning and smoothing

$$F = \frac{L}{4\pi d_L^2}$$

$$d_L(z) = (1+z)\int_0^z \frac{dz'}{H(z')}$$

$$H(z) = \left[\frac{d}{dz}\left(\frac{d_L(z)}{1+z}\right)\right]^{-1}$$

$$\frac{H^2(z)}{H^2{}_0} = \left[\Omega_{0M}(1+z)^3 + (1-\Omega_{0M})\exp[\int 3\left(1+w(z)\right)\frac{dz}{1+z}]\right]$$

$$\omega_{DE} = \frac{\left(\frac{2(1+z)}{3}\frac{H'}{H}\right)-1}{1-\left(\frac{H_0}{H}\right)^2\Omega_{0M}(1+z)^3}$$

The Method of Smoothing

error-sensitive

$$\ln d_L(z,\Delta) = \ln d_L(z)^g + N(z)\sum_i [\ln d_L(z_i) - \ln d_L(z_i)^g]\exp\left[-\frac{\ln^2(\frac{1+z_i}{1+z})}{2\Delta^2}\right](\frac{1}{\sigma_{\ln d_L(z_i)}})^2$$

$$N(z)^{-1} = \sum_i \exp\left[-\frac{\ln^2(\frac{1+z_i}{1+z})}{2\Delta^2}\right](\frac{1}{\sigma_{\ln d_L(z_i)}})^2$$

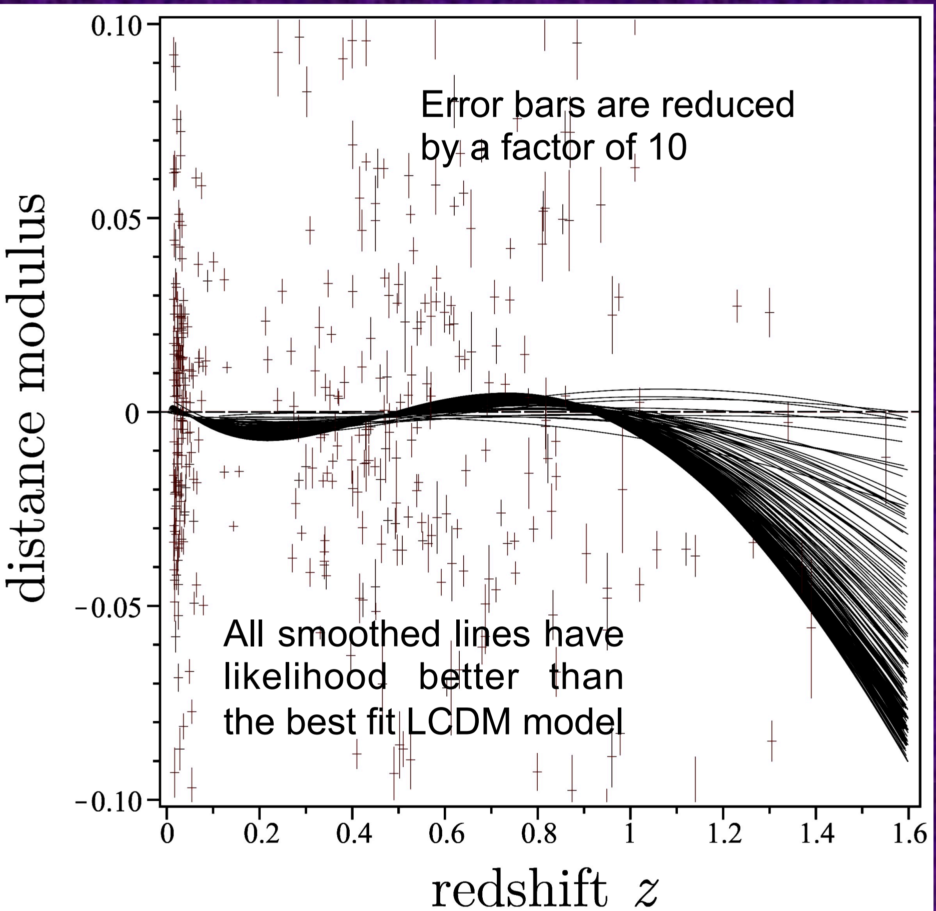
Smoothing function (error-sensitive)

$$H(z) = \left[\frac{d}{dz}\left(\frac{d_L(z)}{1+z}\right)\right]^{-1}$$

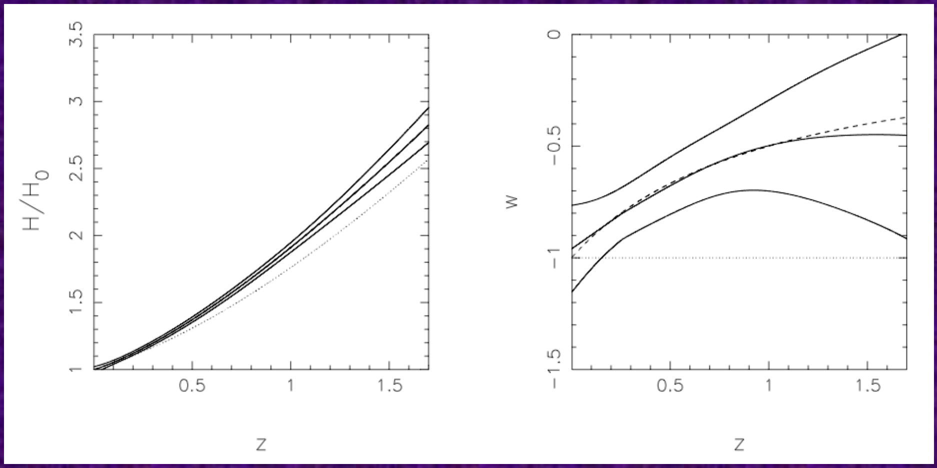
$$\omega_{DE} = \frac{(\frac{2(1+z)}{3}\frac{H'}{H}) - 1}{1-(\frac{H_0}{H})^2\Omega_{0M}(1+z)^3}$$

Smoothing Method

A. Shafieloo, U. Alam, V. Sahni, A. Starobinsky, MNRAS (2006)
A. Shafieloo, MNRAS (2007)
A. Shafieloo & C. Clarkson PRD (2010)
A. Shafieloo JCAP (2012)



- ➔Iterative approach with lognormal smoothing kernel
- ➔Error sensitive
- ➔Independent of the initial guess



$$H(z) = \left[\frac{d}{dz} \left(\frac{d_L(z)}{1+z} \right) \right]^{-1}$$

$$\omega_{DE} = \frac{\left(\frac{2(1+z)}{3} \frac{H'}{H} \right) - 1}{1 - \left(\frac{H_0}{H} \right)^2 \Omega_{0M} (1+z)^3}$$

Gaussian Process

- ➔Efficient in statistical modeling of stochastic variables
- ➔Derivatives of Gaussian Processes are Gaussian Processes
- ➔Provides us with all covariance matrices

A. Shafieloo. A. Kim & E. Linder, PRD 2012, PRD 2013

Data

Mean Function

$$\begin{bmatrix} \mathbf{y} \\ \mathbf{f} \\ \mathbf{f}' \\ \mathbf{f}'' \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mathbf{m}(\mathbf{Z}) \\ \mathbf{m}(\mathbf{Z_1}) \\ \mathbf{m}'(\mathbf{Z_1}) \\ \mathbf{m}''(\mathbf{Z_1}) \end{bmatrix}, \begin{bmatrix} \Sigma_{00}(Z,Z) & \Sigma_{00}(Z,Z_1) & \Sigma_{01}(Z,Z_1) & \Sigma_{02}(Z,Z_1) \\ \Sigma_{00}(Z_1,Z) & \Sigma_{00}(Z_1,Z_1) & \Sigma_{01}(Z_1,Z_1) & \Sigma_{02}(Z_1,Z_1) \\ \Sigma_{10}(Z_1,Z) & \Sigma_{10}(Z_1,Z_1) & \Sigma_{11}(Z_1,Z_1) & \Sigma_{12}(Z_1,Z_1) \\ \Sigma_{20}(Z_1,Z) & \Sigma_{20}(Z_1,Z_1) & \Sigma_{21}(Z_1,Z_1) & \Sigma_{22}(Z_1,Z_1) \end{bmatrix} \right),$$

$$\Sigma_{\alpha\beta} = \frac{d^{(\alpha+\beta)}K}{dz_i^\alpha dz_j^\beta},$$

$$\begin{bmatrix} \overline{\mathbf{f}} \\ \overline{\mathbf{f}'} \\ \overline{\mathbf{f}''} \end{bmatrix} = \begin{bmatrix} \mathbf{m}(\mathbf{Z_1}) \\ \mathbf{m}'(\mathbf{Z_1}) \\ \mathbf{m}''(\mathbf{Z_1}) \end{bmatrix} + \begin{bmatrix} \Sigma_{00}(Z_1,Z) \\ \Sigma_{10}(Z_1,Z) \\ \Sigma_{20}(Z_1,Z) \end{bmatrix} \Sigma_{00}^{-1}(Z,Z)\,\mathbf{y}$$

Kernel ➔

$$k(z,z') = \sigma_f^2 \exp\left(-\frac{|z-z'|^2}{2l^2} \right),$$

GP Hyper-parameters

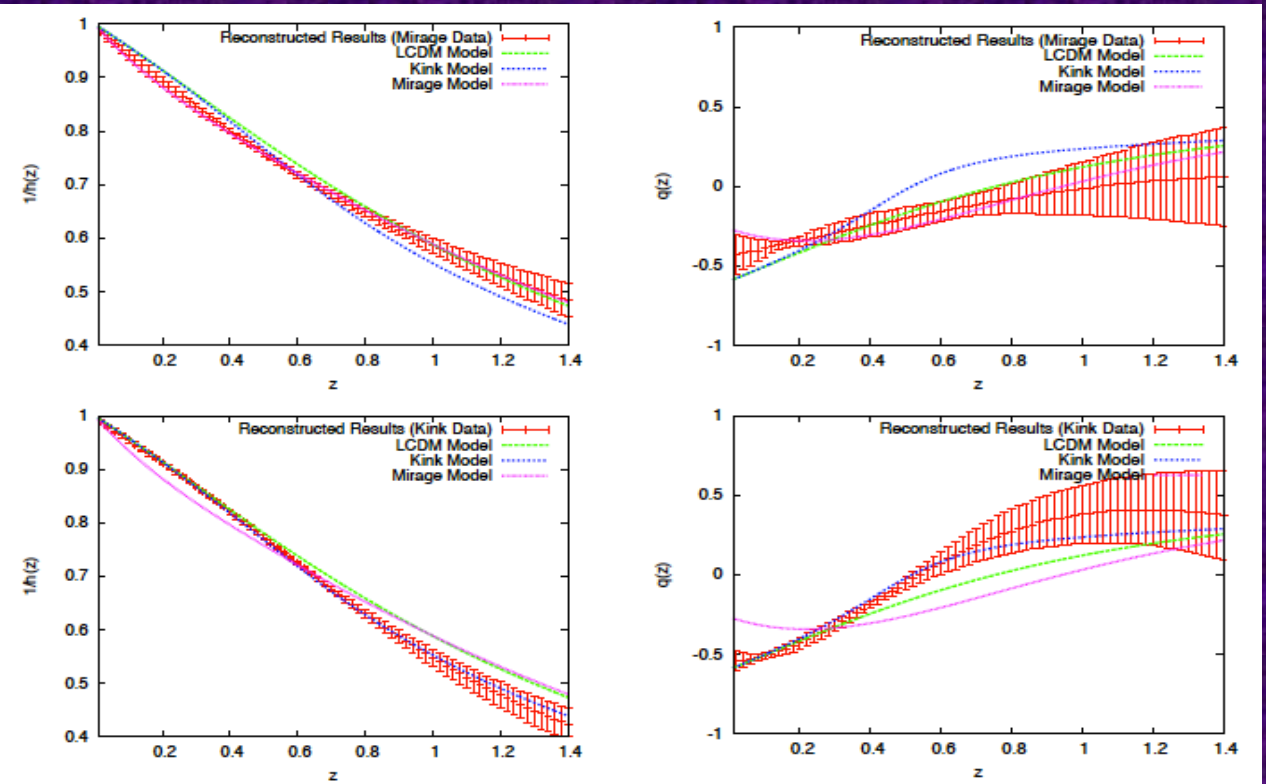
$$\mathrm{Cov}\left(\begin{bmatrix} \mathbf{f} \\ \mathbf{f}' \\ \mathbf{f}'' \end{bmatrix} \right) = \begin{bmatrix} \Sigma_{00}(Z_1,Z_1) & \Sigma_{01}(Z_1,Z_1) & \Sigma_{02}(Z_1,Z_1) \\ \Sigma_{10}(Z_1,Z_1) & \Sigma_{11}(Z_1,Z_1) & \Sigma_{12}(Z_1,Z_1) \\ \Sigma_{20}(Z_1,Z_1) & \Sigma_{21}(Z_1,Z_1) & \Sigma_{22}(Z_1,Z_1) \end{bmatrix} - \begin{bmatrix} \Sigma_{00}(Z_1,Z) \\ \Sigma_{10}(Z_1,Z) \\ \Sigma_{20}(Z_1,Z) \end{bmatrix} \Sigma_{00}^{-1}(Z,Z)\left[\Sigma_{00}(Z,Z_1), \Sigma_{01}(Z,Z_1), \Sigma_{02}(Z,Z_1)\right].$$

$$2\ln p(y|f) = -y^T\Sigma_{00}(Z,Z)^{-1}y - \ln\det\Sigma_{00}(Z,Z) - n\ln(2\pi),$$

GP Likelihood

Gaussian Process

- ➔Efficient in statistical modeling of stochastic variables
- ➔Derivatives of Gaussian Processes are Gaussian Processes (we can derive H(z) directly and q(z) indirectly)
- ➔Provides us with all covariance matrices



$$H(z) = \left[\frac{d}{dz}\left(\frac{d_L(z)}{1+z}\right)\right]^{-1}$$

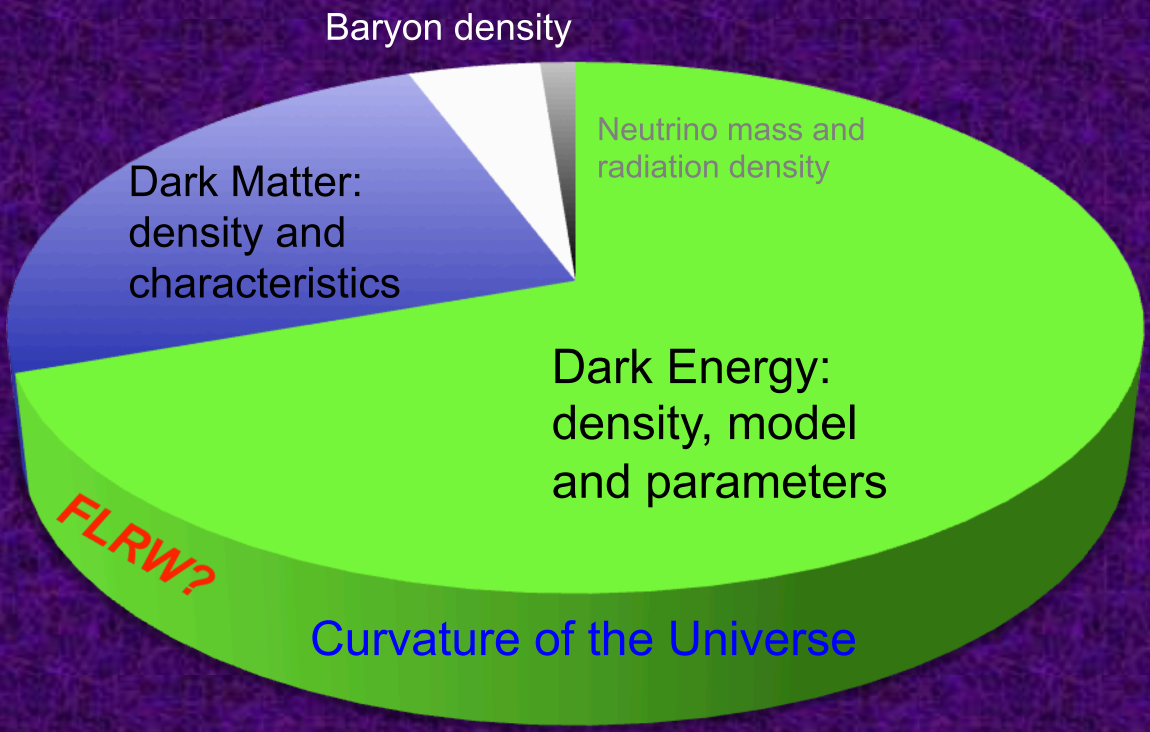
$$q(z) = (1+z)\frac{H'(z)}{H(z)} - 1$$

A. Shafieloo. A. Kim & E. Linder, PRD 2012

Reconstruction & Falsification

Reconstruction: Understanding the behavior

Falsification: Testing the Consistency



Initial Conditions:
Form of the Primordial Spectrum and Model of Inflation and its Parameters

Epoch of reionization

Hubble Parameter and the Rate of Expansion

Falsification, signal detection and systematics:

Testing deviations from an assumed model

(without comparing different models)

Null tests:

Falsifying a hypothesis using special statistical characteristics.

Using modeling of Stochastic variables:

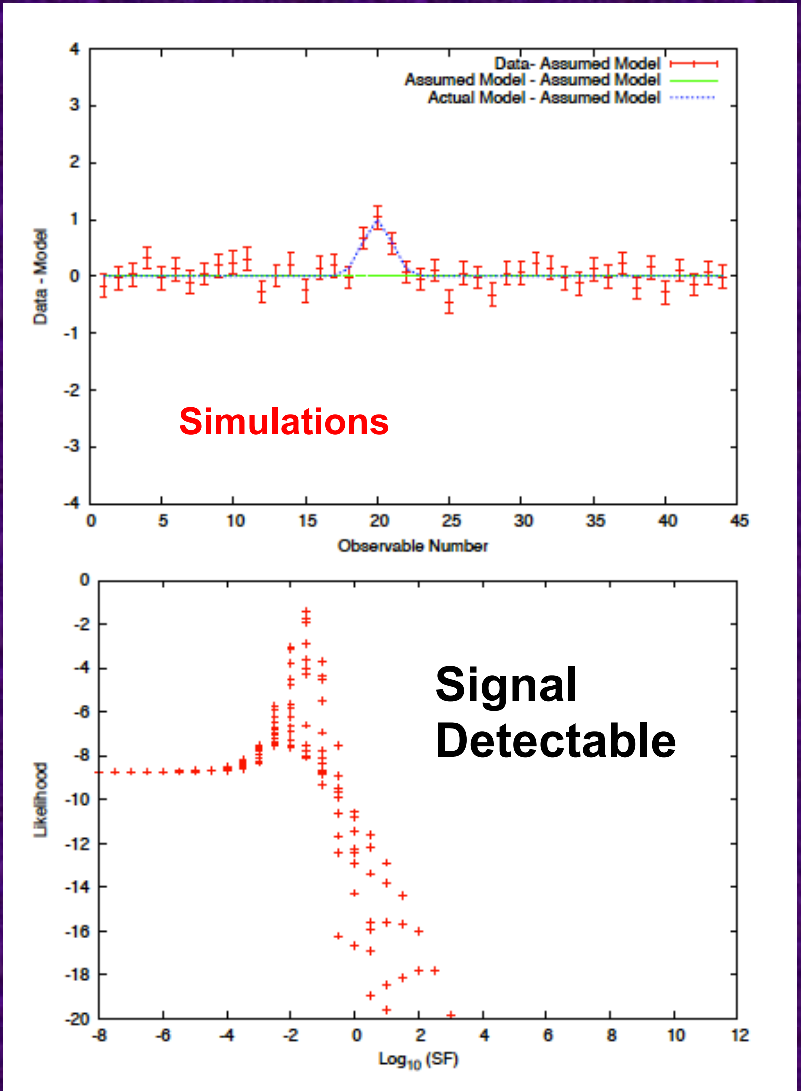
Gaussian Processes:

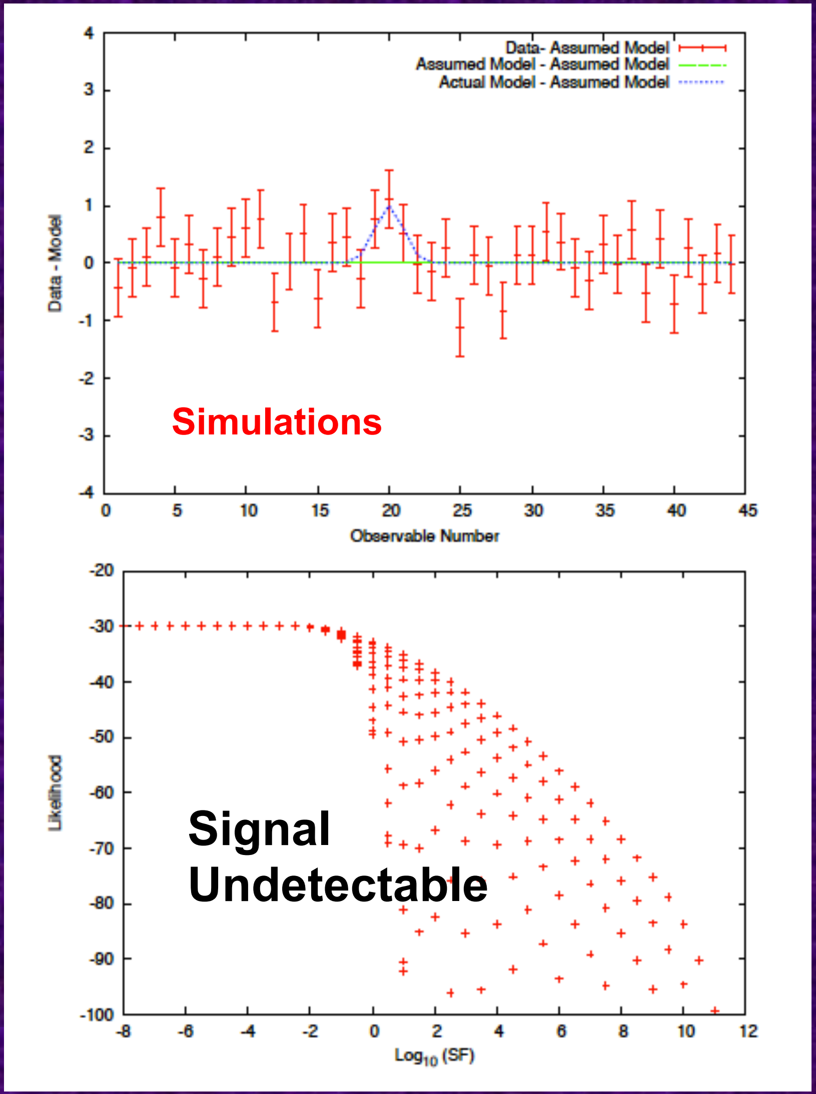
Modeling of the data around a mean function searching for likely features by looking at the the likelihood space of the hyperparameters.

Bayesian Interpretation of Crossing Statistic:

Comparing a model with its own possible variations.

Signal Detection: detection of the features in the residuals

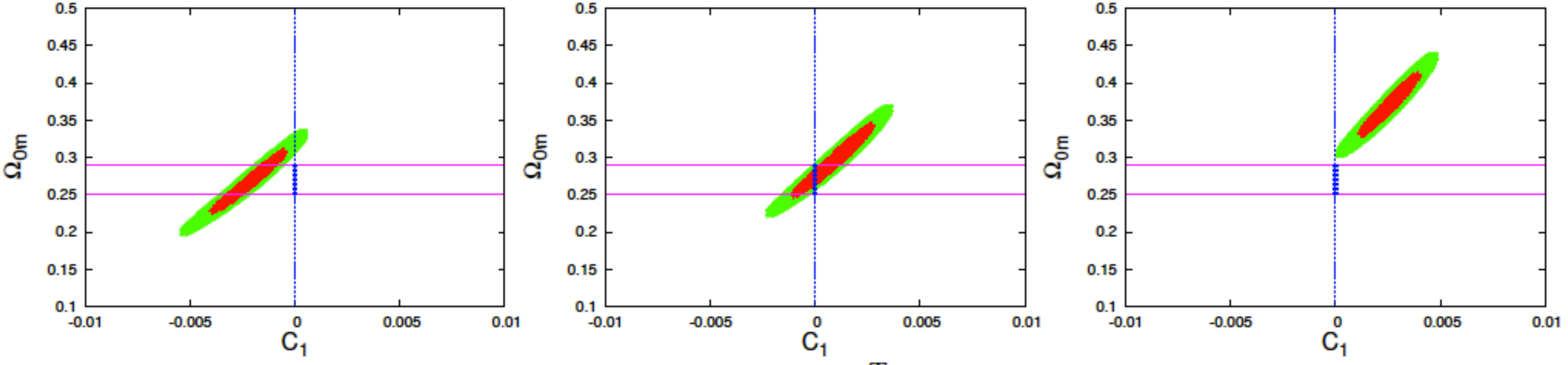




Signal Detection: Crossing Statistic (Bayesian Interpretation)

Theoretical model Crossing function

$$\mu_M^{T_N}(z) = \mu_M(p_i, z) \times T_N(C_1, \ldots, C_N, z)$$



$$T_I(C_1, z) = 1 + C_1\left(\frac{z}{z_{max}}\right)$$

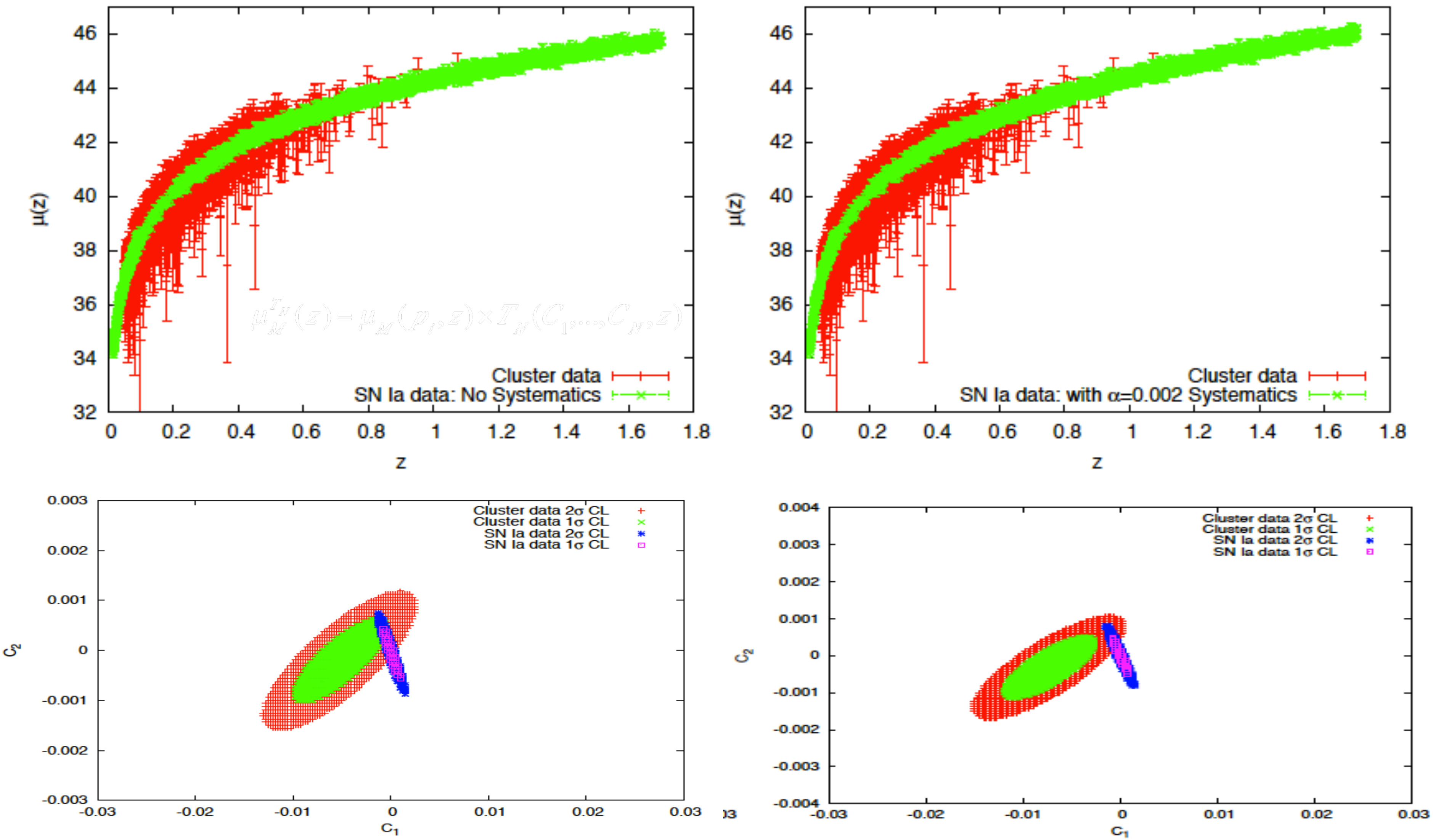
Chebishev Polynomials as Crossing Functions

$$T_{II}(C_1, C_2, z) = 1 + C_1\left(\frac{z}{z_{max}}\right) + C_2\left[2\left(\frac{z}{z_{max}}\right)^2 - 1\right],$$

A. Shafieloo. JCAP 2012 (b)

Looking for systematics

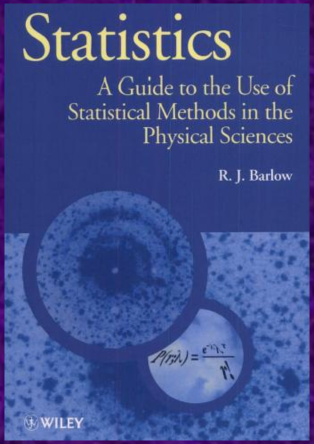
Modeling different data independent of theoretical assumptions looking for systematics



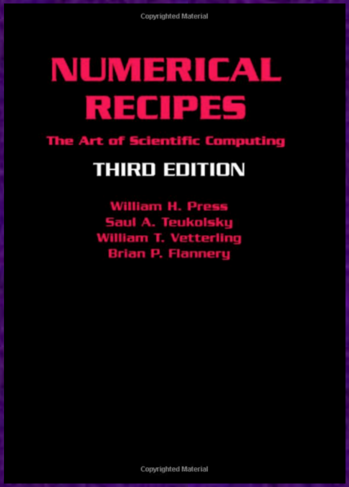
Summary

- Cosmological Data is very very expensive so it is important to extract the most possible amount of information from it.
- Dealing with the data is not easy. One should be very careful not to misinterpret the data. There have been many discoveries which have been wrong and ends to embarrassment.
- About 96% of the universe is still missing. We do not know what is Dark matter and what is dark energy. Actual model of the early universe is not known. We should be careful when we model any of these.
- Parametric and Non-Parametric, at the same time, frequentist and bayesian approaches are all useful and each has some advantages. Best is to combine and use them in an ***appropriate way***.

What to read:



- **Statistics**, R. J. Barlow, John Wiley & Sons Ltd, (1989)



- **Numerical Recipes**, William H. Press, Saul A. Teukolsky, William T. Vetterling, Brian P. Flannery, Cambridge University Press (2007)

- Arxiv:0712.3028, L. Verde
  Arxiv:0911.3105, L. Verde