On model selection, validation and reconstruction in the context of physical cosmology

Arman Shafieloo,

Korea Astronomy and Space Science Institute (KASI)

University of Science and Technology (UST)

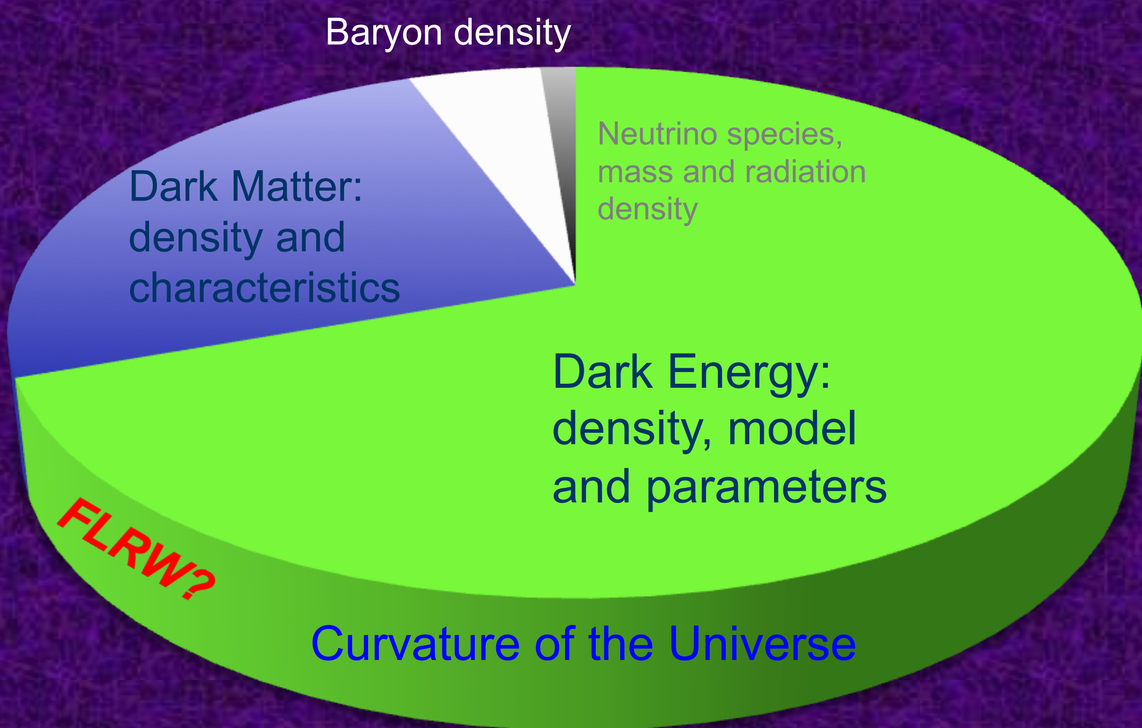
The 2nd Shanghai Assembly on Cosmology and Structure Fromation

30 October-3 November 2023,

SJTU-TDLI, Shanghai

Era of Precision Cosmology

We try to reconstruct and understand the dynamics of the universe and properties of its constituents using various measurements and statistical techniques. Phenomenological and then theoretical works can follow to place constraints on suggested 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 Rate of Expansion

What do we do?

- There are various reconstruction approaches, parametric and non-parametric.

- There have been many phenomenological and theoretical models proposed (recently, to alleviate tensions).

Reconstruction → Phenomenology → Theory

- There have been continuous attempts looking for systematics in various data.

- These models/reconstructions can be very different.

How do we compare them?

Consistency of a proposed model and the data:

Frequentist Approach:

Assuming a proposed model, the probability of the observed data must not be insignificant. Best is to do large number of careful simulations based on a well defined covariance error-matrix.

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

Why things are more complicated than what we think…

Likelihood

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

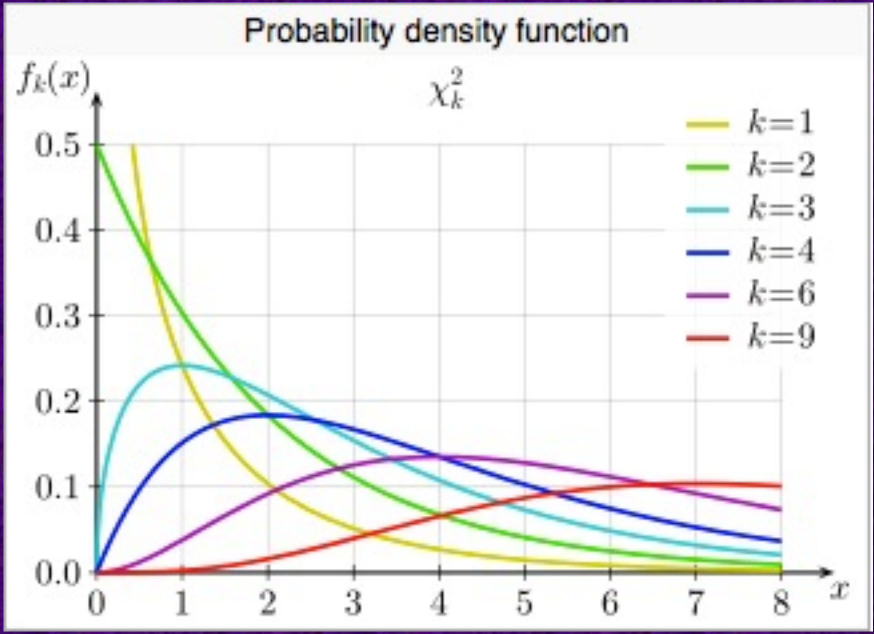
$$\chi^2 = \sum_i^N (\mu_i^t - \mu_i^e)^T Cov^{-1} (\mu_i^t - \mu_i^e)$$

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

When data is uncorrelated

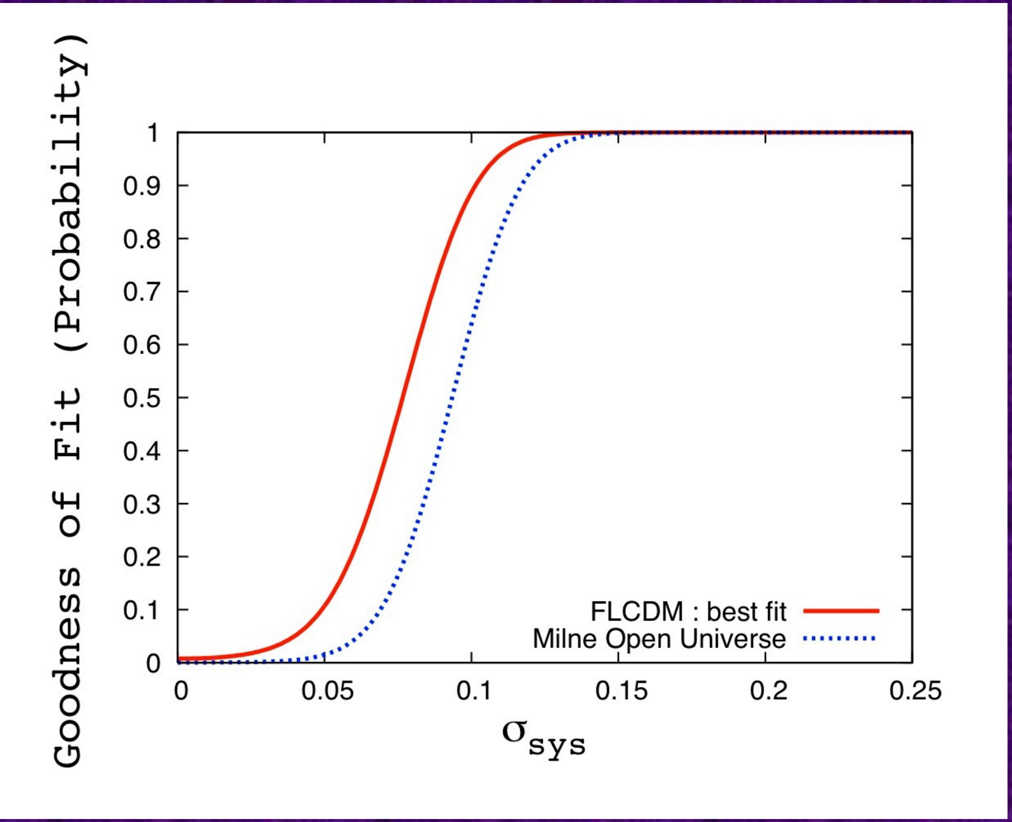
$$P(\chi^2; N) = \frac{2^{-N/2}}{\Gamma(N/2)} \chi^{N-2} e^{-\chi^2/2}$$

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



What if the exact form of the error matrix is not known?

Point 1



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

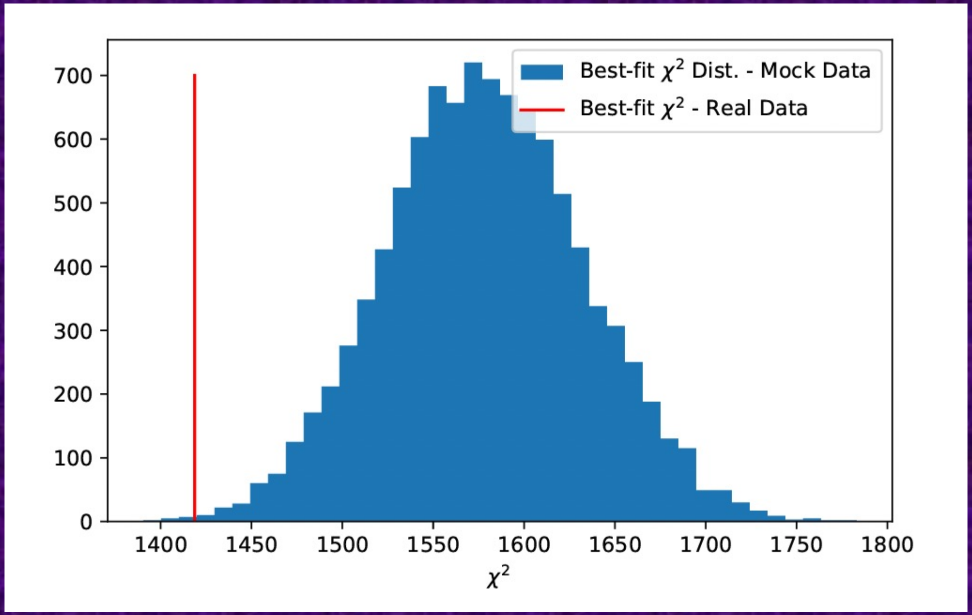
$$\sigma_i^2 = \sigma_{i(data)}^2 + \sigma_{(sys)}^2$$

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

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

e.g. *The case of Type Ia supernovae*

This can still happen!



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

$$\sigma_i^2 = \sigma_{i(data)}^2 + \sigma_{(sys)}^2$$

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

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

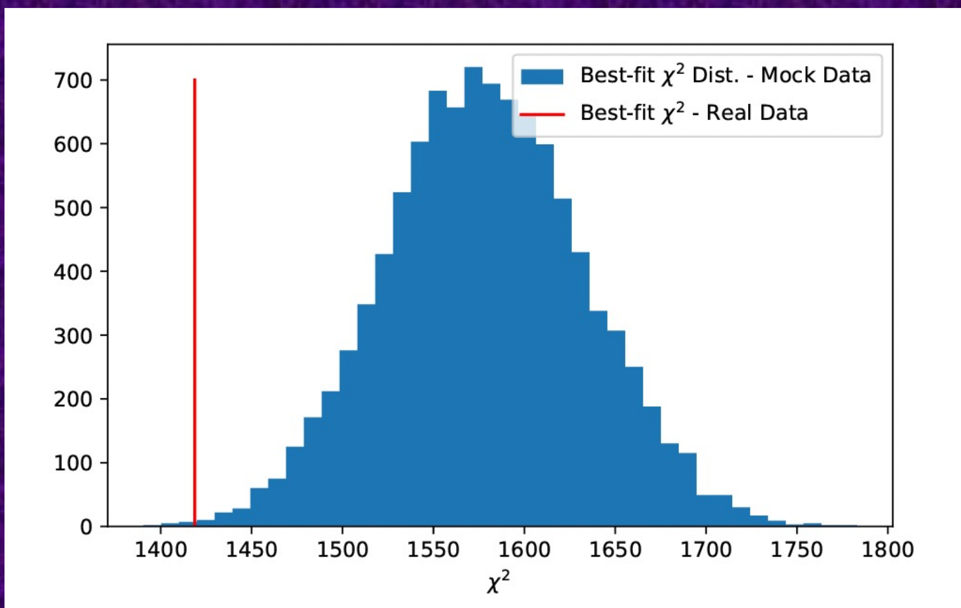
Pantheon+ data

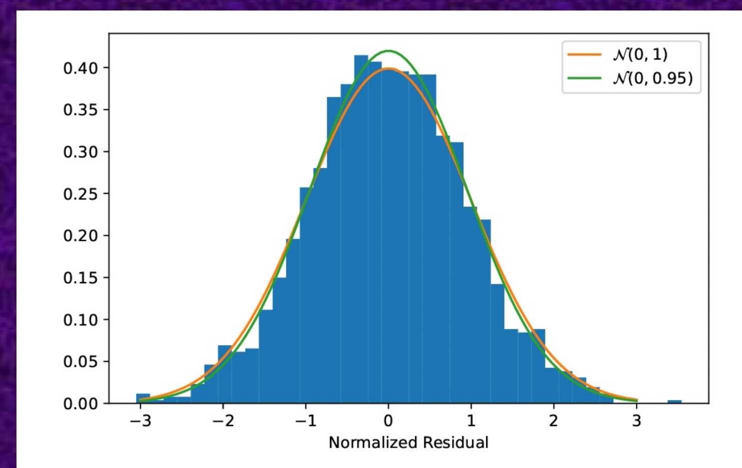
Keeley, Shafieloo, L'Huillier, arXiv:2212.07917

This can still happen!

Point 1

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





Pantheon+ data

Keeley, Shafieloo, L'Huillier, arXiv:2212.07917

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

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.

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

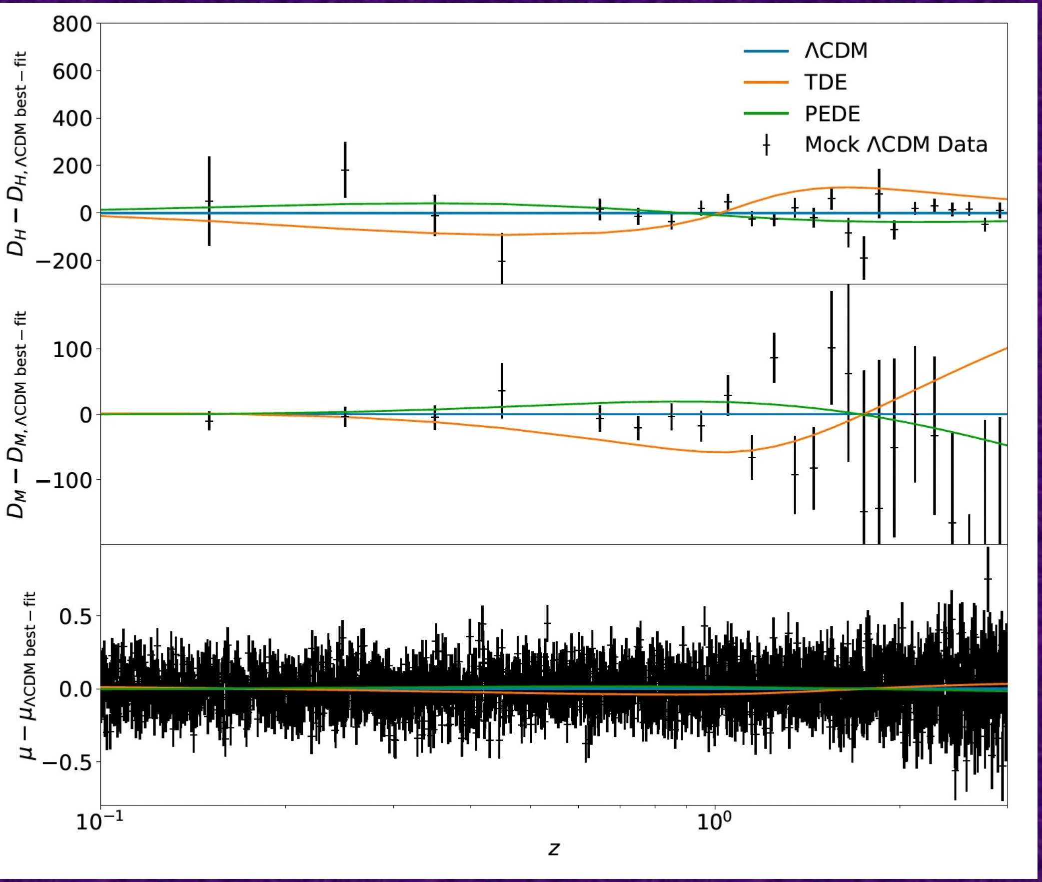
In likelihood estimation:

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

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

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

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



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

Model fitting has Bayesian essence since we assume that we are considering a correct model.

Point 2

Bayesian Evidence and Model Selection

- Bayesian evidence: Integral of (likelihood)x(prior) over the parameter space: $Z = \int L(\theta)\pi(\theta)d\theta$

- Bayes factor: Ratio of the evidence of the two models:

$$\Delta \log Z = \log Z(\mathrm{M}_1) - \log Z(\mathrm{M}_2)$$

Supports Model 1 over Model 2 when $\Delta \log Z$ have a positive value

Jeffreys scale Z_i/Z_j	Kass-Rafferty scale Z_i/Z_j	Interpretation
1 to 3.2	1 to 3	Not worth mentioning
3.2 to 10	3 to 20	Positive
10 to 100	20 to 150	Strong
> 100	>150	Very Strong

Bayesian Evidence and Model Selection

- Bayesian evidence: Integral of (likelihood)x(prior) over the parameter space: $Z = \int L(\theta)\pi(\theta)d\theta$

- Bayes factor: Ratio of the evidence of the two models:
$$\Delta \log Z = \log Z(\mathrm{M}_1) - \log Z(\mathrm{M}_2)$$

Supports Model 1 over Model 2 when $\Delta \log Z$ have a positive value

Jeffreys scale Z_i/Z_j	Kass-Rafferty scale Z_i/Z_j	Interpretation
1 to 3.2	1 to 3	Not worth mentioning
3.2 to 10	3 to 20	Positive
10 to 100	20 to 150	Strong
> 100	>150	Very Strong

Point 3

How reliable are these scales?

(Present)

Standard Model of Cosmology

Universe is Flat

Let's solve Hubble tension with evolving DE!

Universe is Isotropic

Universe is Homogeneous

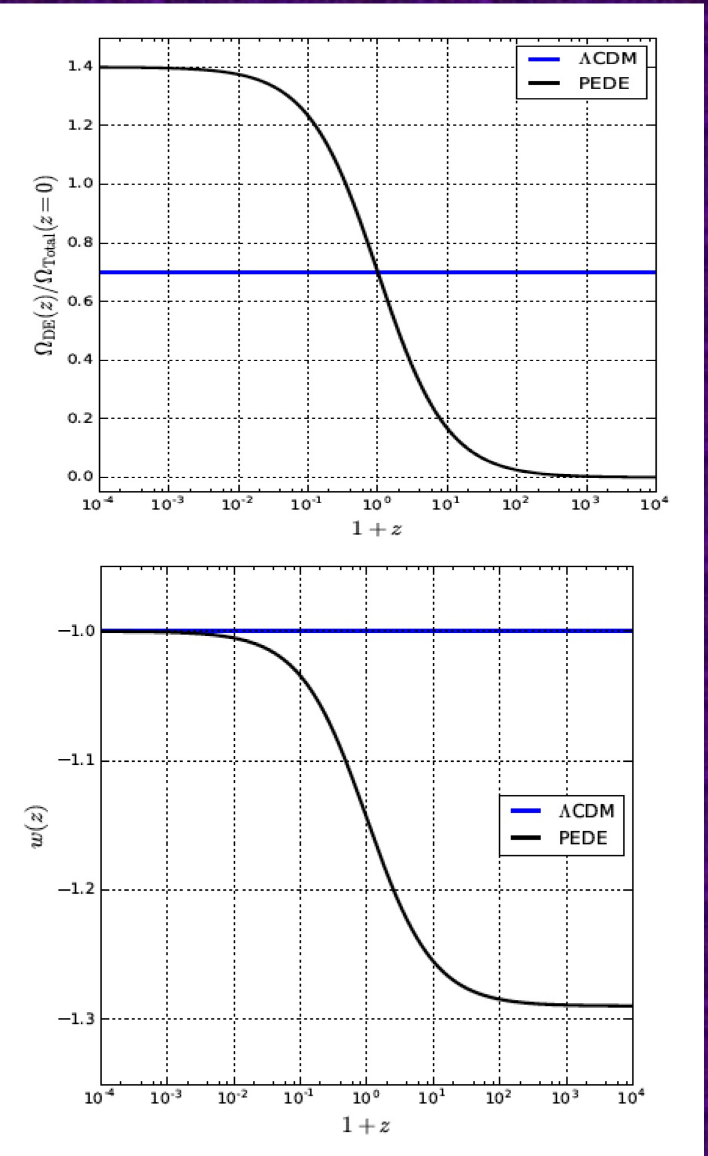
Dark Energy is Lambda (w=-1)

Power-Law primordial spectrum (n_s=const)

Dark Matter is cold

All within framework of FLRW

Phenomenologically Emergent Dark Energy (PEDE)



No Dark Energy in the past and it acts as an emergent phenomena:

Allows lower rate of expansion in the past and higher rate of expansion at late times

$$\Omega_{\rm DE}(z) = \Omega_{\rm DE,0} \times [1 - \tanh(\log_{10}(1+z))]$$

$$\begin{aligned} w(z) &= -\frac{1}{3\ln 10} \times \frac{1 - \tanh^2[\log_{10}(1+z)]}{1 - \tanh[\log_{10}(1+z)]} - 1 \\ &= -\frac{1}{3\ln 10} \times (1 + \tanh[\log_{10}(1+z)]) - 1. \end{aligned}$$

Li and Shafieloo, ApJ Lett 2019

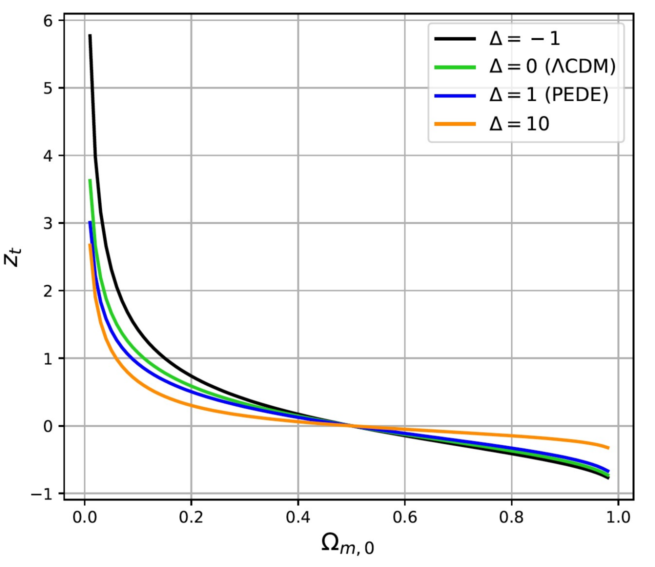
Generalized Emergent Dark Energy (GEDE)

$$\widetilde{\Omega}_{\mathrm{DE}}(z) = \Omega_{\mathrm{DE},0}\frac{1-\tanh\left(\Delta\times\log_{10}(\frac{1+z}{1+z_t})\right)}{1+\tanh\left(\Delta\times\log_{10}(1+z_t)\right)}$$

$$w(z) = -\frac{\Delta}{3\ln 10}\times\left(1+\tanh\left(\Delta\times\log_{10}(\frac{1+z}{1+z_t})\right)\right)-1.$$

-Has one degree of freedom for DE sector

-LCDM and PEDE are both included at special limits



$$\Omega_{\mathrm{DE}}(z_t) = \Omega_{m,0}(1+z_t)^3$$

$\Delta = 0$ LCDM

$\Delta = 1$ PEDE

Li and Shafieloo, ApJ 2020 (arXiv:2001.05103)

Generalized Emergent Dark Energy (GEDE)

Data	$\ln B_{ij}$
Planck 2018	2.9
Planck 2018+BAO	0.8
Planck 2018+R19	12.1
Planck 2018+BAO+R19	7.9
Planck 2018+JLA	−0.2
Planck 2018+Pantheon	−0.9
Planck 2018+BAO+JLA+R19	6.1
Planck 2018+BAO+Pantheon+R19	5.8

Full analysis using various combination of the data

$\Delta \log Z$	Evidence against M_1
0 to 1	Negligible
1 to 3	Positive
3 to 5	Strong
> 5	Very strong

Model Comparison:
Bayesian evidence analysis in strong support of emergent dark energy

W. Yang, et al, PRD 2021 [arXiv:2103.03815]

Generalized Emergent Dark Energy (GEDE)

Data	ln B_{ij}
Planck 2018	2.9
Planck 2018+BAO	0.8
Planck 2018+R19	12.1
Planck 2018+BAO+R19	7.9
Planck 2018+JLA	−0.2
Planck 2018+Pantheon	−0.9
Planck 2018+BAO+JLA+R19	6.1
Planck 2018+BAO+Pantheon+R19	5.8

Full analysis using various combination of the data

Current tensions allow us to find models statistically better (?) than LCDM but are all tensions resolved?

No!

Model Comparison:
Bayesian evidence analysis in strong support of emergent dark energy

True for any successful evolving DE model!

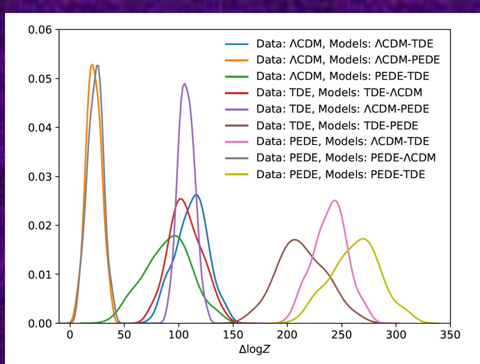
W. Yang,et al, PRD 2021 [arXiv:2103.03815]

Distribution of Bayesian Evidence:

- Be cautious about Jeffery's scale!

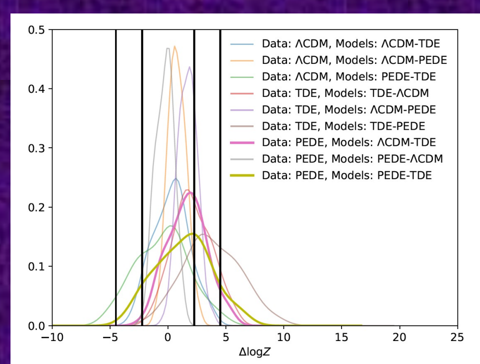
Jeffreys scale Z_i/Z_j	Kass-Rafferty scale Z_i/Z_j	Interpretation
1 to 3.2	1 to 3	Not worth mentioning
3.2 to 10	3 to 20	Positive
10 to 100	20 to 150	Strong
> 100	>150	Very Strong

Distribution of Bayes factors can greatly depend on the models and the data!



Data with OK quality

Data with OK quality



Data with worse quality

On The Distribution of Bayesian Evidences

Ryan E. Keeley,[1,2]⋆ Arman Shafieloo,[2,3]†
[1]Department of Physics, University of California Merced, 5200 North Lake Road, Merced, CA 95343, USA
[2]Korea Astronomy and Space Science Institute (KASI), 776 Daedeok-daero, Yuseong-gu, Daejeon 34055, Korea
[3]KASI Campus, University of Science and Technology, 217 Gajeong-ro, Yuseong-gu, Daejeon 34113, Korea

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT
We look at the distribution of the Bayesian evidence for mock realizations of supernova and baryon acoustic oscillation data. The ratios of Bayesian evidences of different models are often used to perform model selection. The significance of these Bayes factors are then interpreted using scales such as the Jeffreys or Kass & Raftery scale. First, we demonstrate how to use the evidence itself to validate the model, that is to say how well a model fits the data, regardless of how well other models perform. The basic idea is that if, for some real dataset a model's evidence lies outside the distribution of evidences that result when the same fiducial model that generates the datasets is used for the analysis, then the model in question is robustly ruled out. Further, we show how to assess the significance of a hypothetically computed Bayes factor. We show that the range of the distribution of Bayes factors can greatly depend on the models in question and also the number of data points in the dataset. Thus, we have demonstrated that the significance of Bayes factors needs to be calculated for each unique dataset.

Key words: dark energy – cosmological parameters – methods: statistical

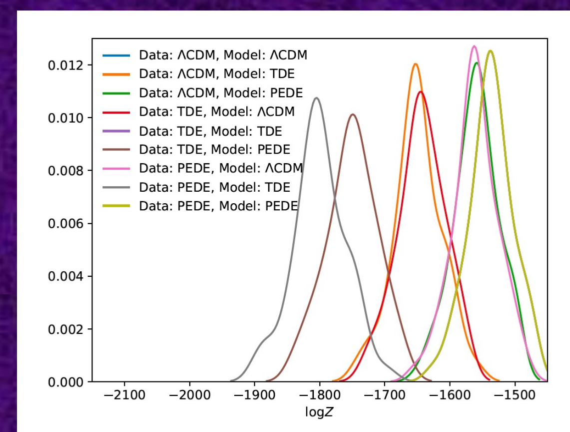
Keeley and Shafieloo, MNRAS 2022

Bayes Factor:

Point 3

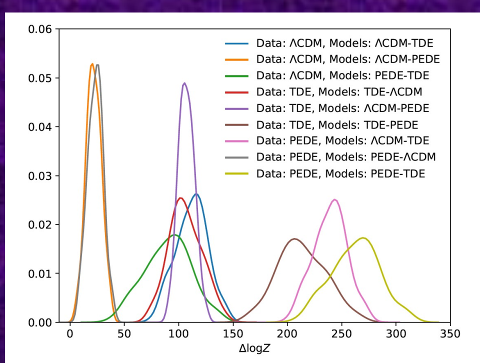
- Be cautious about Jeffery's scale!

Jeffreys scale Z_i/Z_j	Kass-Rafferty scale Z_i/Z_j	Interpretation
1 to 3.2	1 to 3	Not worth mentioning
3.2 to 10	3 to 20	Positive
10 to 100	20 to 150	Strong
> 100	>150	Very Strong



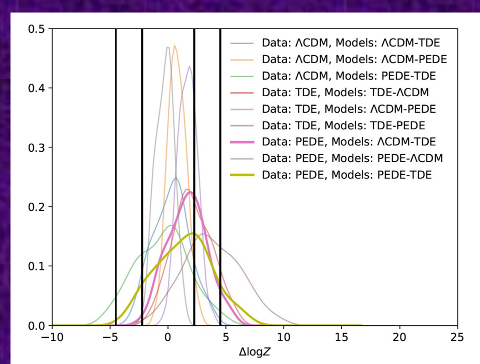
Keeley and Shafieloo, MNRAS 2022

Distribution of Bayes factors can greatly depend on the models and the data!



Data with OK quality

Data with OK quality



Data with worse quality

See also:

Starkman et al, arXiv:0811.2415
Jenkins & Peacock, MNRAS 2011;
Nesseris & Garcia-Bellido, JCAP 2013;
Joachimi et al., A&A 2021;

Model Validation

Bayesian evidence approach is solid but only can find the better model among the candidates (or less wrong model/ranking models)

Importance of Model Validation

$\Delta \log Z > 3$	PEDE consistent	PEDE ruled-out
ΛCDM consistent	6	994
ΛCDM ruled-out	0	0
$\Delta \log Z > 5$	PEDE consistent	PEDE ruled-out
ΛCDM consistent	89	911
ΛCDM ruled-out	0	0

Conventional Bayesian Evidence Approach

→When true model is unknown, finding a statistical anchor is not trivial

→One can attempt using reliable non-parametric reconstructions

Both models are wrong!

Point 4

Koo, Keeley, Shafieloo, L'Huillier, JCAP 2022

Iterative Smoothing Method

- The non-parametric method to reconstruct the distance modulus and expansion history of the universe
  Shafieloo et al. 2006, 2018; Shafieloo. 2007; Shafieloo & Clarkson 2010

- Starts from initial guess of distance modulus, but generates model-independent reconstruction of distance modulus with lower χ^2 value after numerous iterations

$$\hat{\mu}_{n+1}(z) = \hat{\mu}_n(z) + \frac{\boldsymbol{\delta\mu_n}^T \cdot \mathbf{C}^{-1} \cdot \boldsymbol{W}(z)}{\mathbf{1}^T \cdot \mathbf{C}^{-1} \cdot \boldsymbol{W}(z)}$$ ($\mathbf{C}$: Covariance matrix of the data)

$$\mathbf{1}^T = (1, \cdots, 1), \boldsymbol{W}_i(z) = \exp\left(-\frac{\ln^2\left(\frac{1+z}{1+z_i}\right)}{2\Delta^2}\right), \boldsymbol{\delta\mu_n}|_i = \mu_i - \hat{\mu}_n(z_i)$$ ($\boldsymbol{\Delta}$: Smoothing width)

$$\chi_n^2 = \boldsymbol{\delta\mu_n}^T \cdot \mathbf{C}^{-1} \cdot \boldsymbol{\delta\mu_n}$$

- Derive the **likelihood distribution** function $P(\Delta\chi^2)$ (for a large number of data realizations), where $\Delta\chi^2 = \chi^2_{\text{smooth}} - \chi^2_{\text{best-fit}}$, when the true model is assumed
  Koo et al. 2021, JCAP, 03, 034

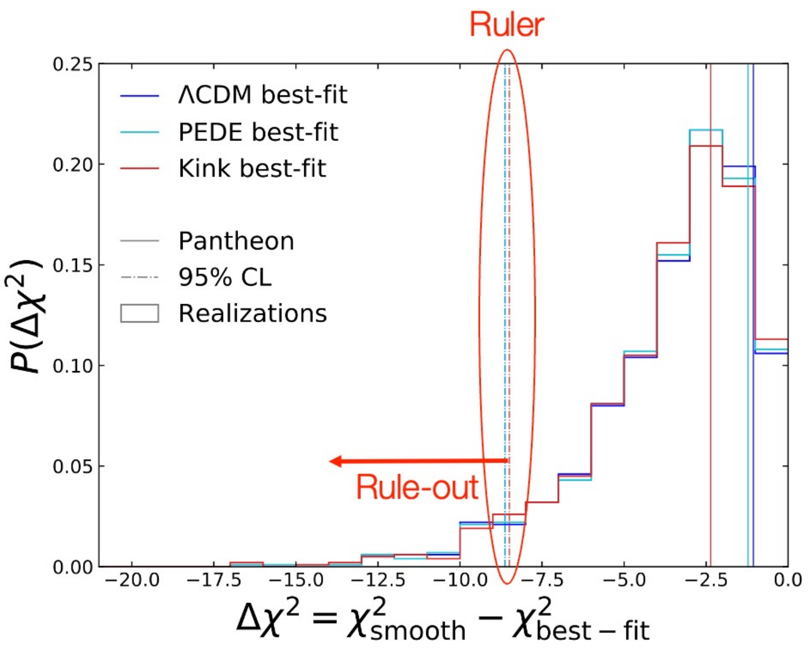
- χ^2_{smooth}: χ^2 of the converged reconstruction using smoothing method

- $\chi^2_{\text{best-fit}}$: Best-fit χ^2 of the correct model fits the data

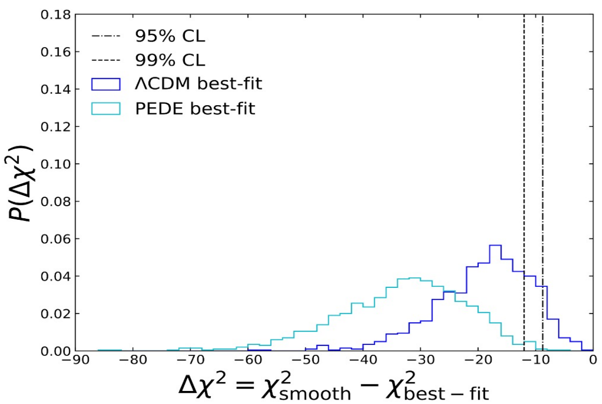
Testing Models based on Likelihood Distribution

- $P(\Delta\chi^2)$ have no dependence on the true model and depends only on the covariance matrix of the data
  → One $\Delta\chi^2$ for given confidence (Ruler)
  Koo et al. 2021, JCAP, 03, 034

- The model being tested is ruled out if the $\Delta\chi^2$ value is lower than the ruler



- Likelihood distributions exclude both models



95% CL	PEDE consistent	PEDE ruled-out
ΛCDM consistent	2	82
ΛCDM ruled-out	0	916
99% CL	PEDE consistent	PEDE ruled-out
ΛCDM consistent	14	193
ΛCDM ruled-out	0	793

Non-parametric reconstruction and Model Validation

Model Validation

Bayesian evidence approach is solid but only can find the better model among the candidates (or less wrong model/ranking models)

Importance of Model Validation

One can design robust statistical approaches for model validation

$\Delta \log Z > 3$	PEDE consistent	PEDE ruled-out
ΛCDM consistent	6	994
ΛCDM ruled-out	0	0
$\Delta \log Z > 5$	PEDE consistent	PEDE ruled-out
ΛCDM consistent	89	911
ΛCDM ruled-out	0	0

Conventional Bayesian Evidence Approach

95% CL	PEDE consistent	PEDE ruled-out
ΛCDM consistent	2	82
ΛCDM ruled-out	0	916
99% CL	PEDE consistent	PEDE ruled-out
ΛCDM consistent	14	193
ΛCDM ruled-out	0	793

Iterative smoothing validation approach

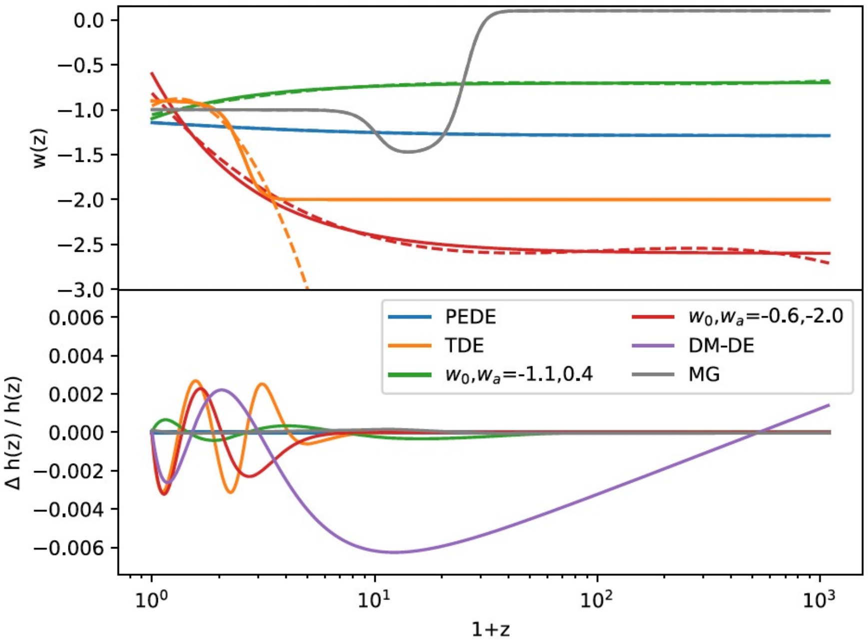
Both models are wrong!

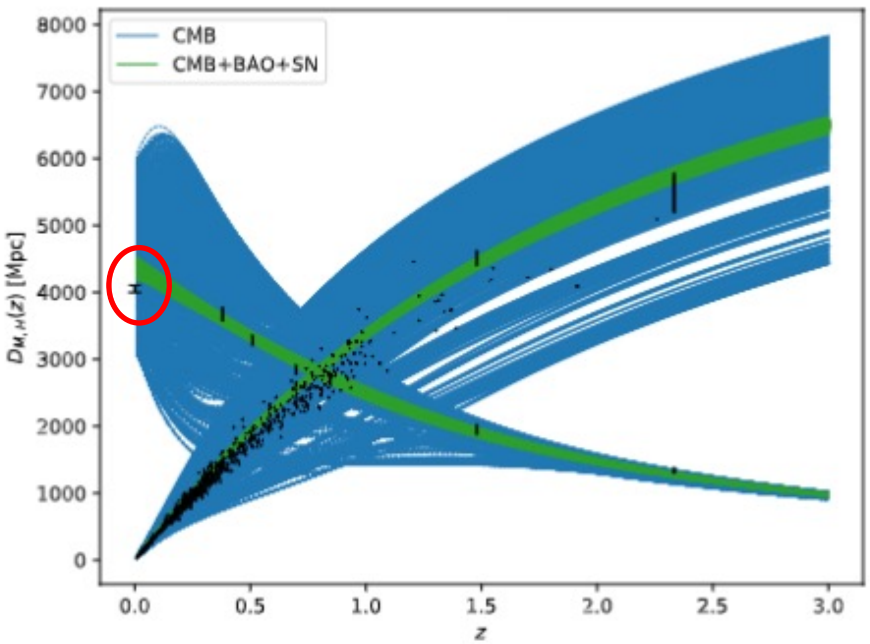
Point 4

Koo, Keeley, Shafieloo, L'Huillier, JCAP 2022

Application of model validation

Ruling Out New Physics at Low Redshift as a solution to the H0 Tension



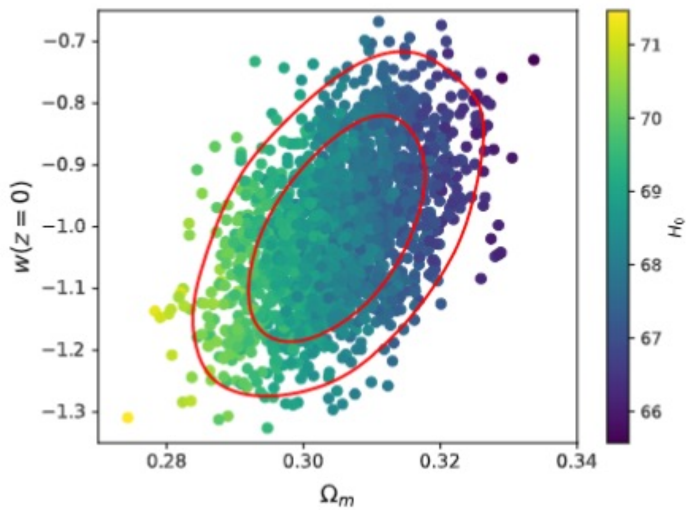


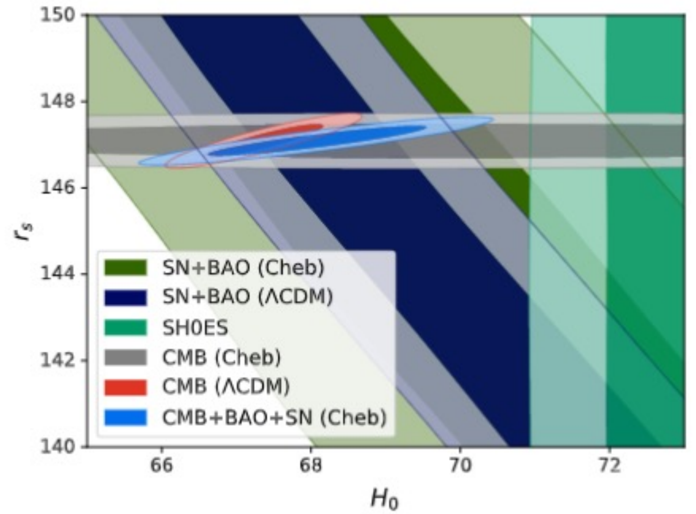
Exploring an **extensive** physical space with Crossing functions for validation (Chebyshev polynomials)

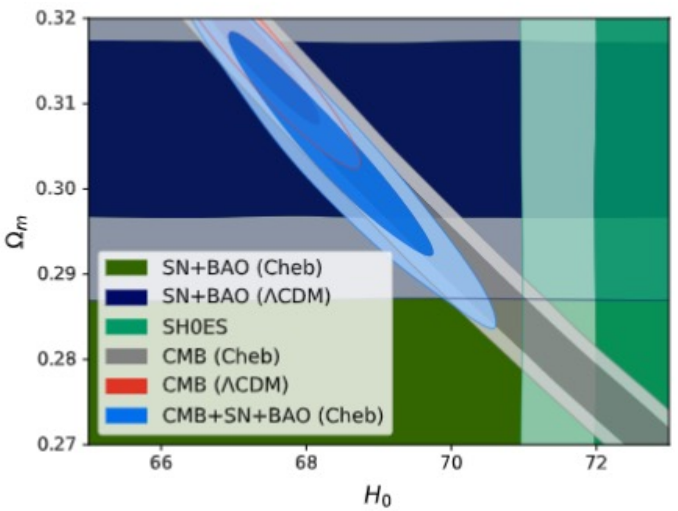
Keeley and Shafieloo, Phys. Rev. Lett, 2023

Application of model validation

Ruling Out New Physics at Low Redshift as a solution to the H0 Tension





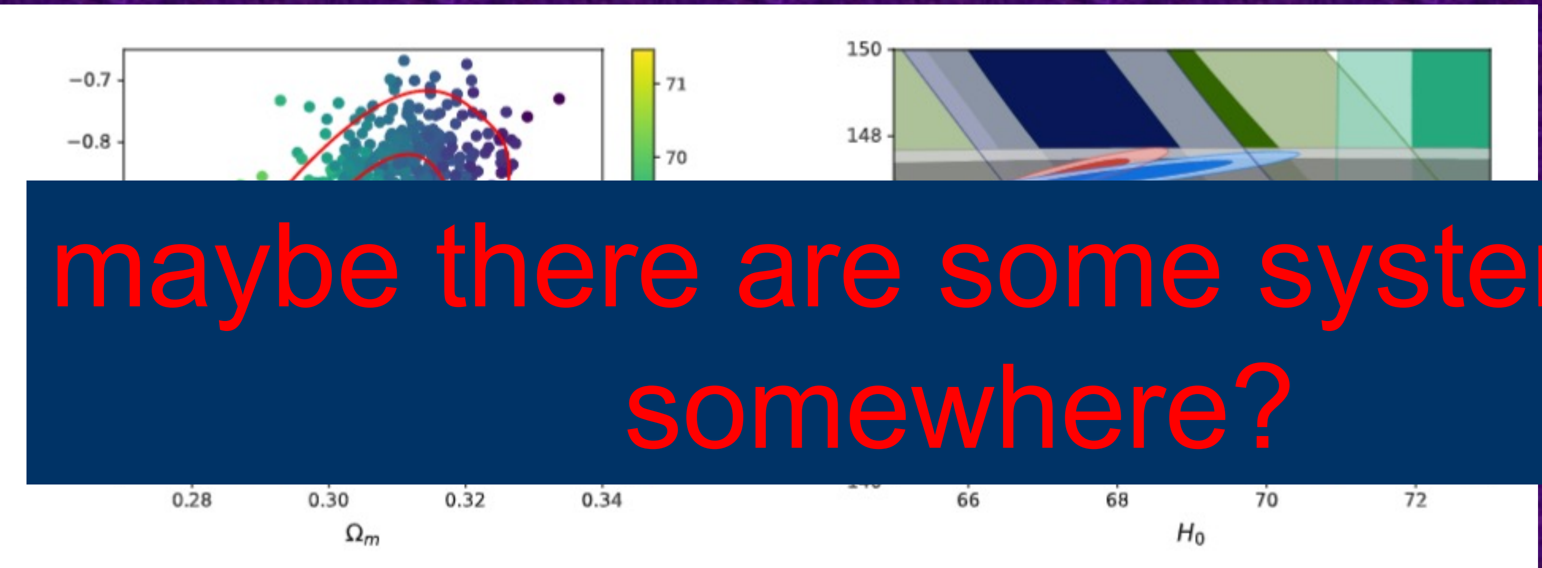


Even in such extensive physical space, inference on H0 is not consistent with SH0ES.

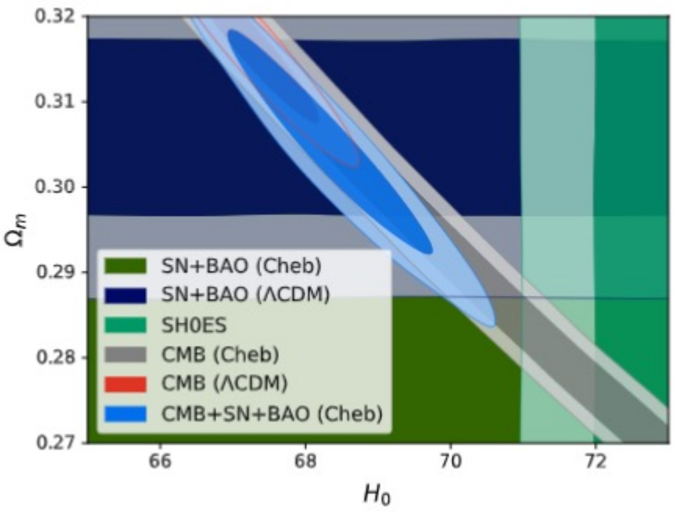
Keeley and Shafieloo, Phys. Rev. Lett, 2023

Application of model validation

Isn’t it suspicious that nothing works?!



maybe there are some systematics somewhere?



Even in such extensive physical space, inference on H0 is not consistent with SH0ES.

Validation of a large number of models can hints towards systematic

(Present)

Lets talk about tensions again…

Standard Model of Cosmology

Universe is Flat

On Importance of non-parametric reconstruction

Universe is Isotropic

Universe is Homogeneous

Dark Energy is Lambda (w=-1)

Power-Law primordial spectrum (n_s=const)

Dark Matter is cold

When we don't know what to look for!

All within framework of FLRW

Point 5

Let's Reconstruction Leads the way!

Point 5

Model Independent Reconstruction of Primordial Spectrum

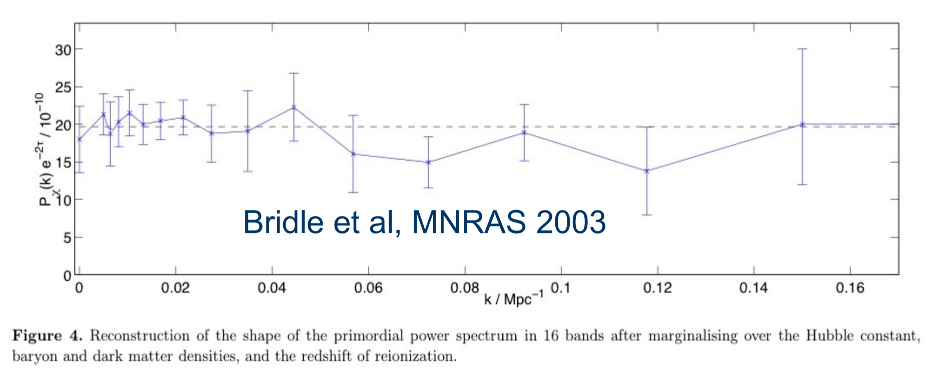
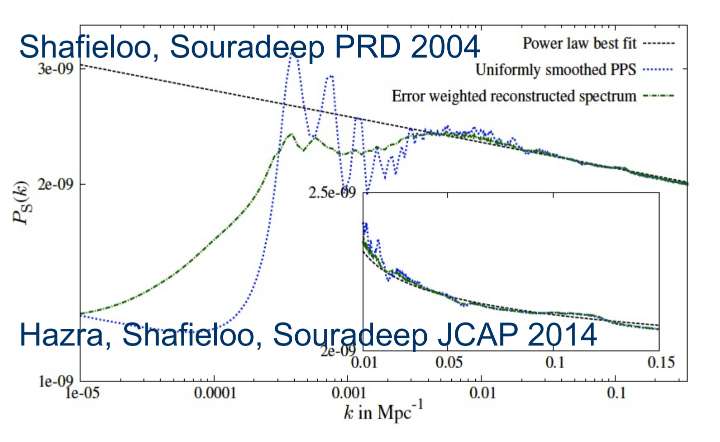
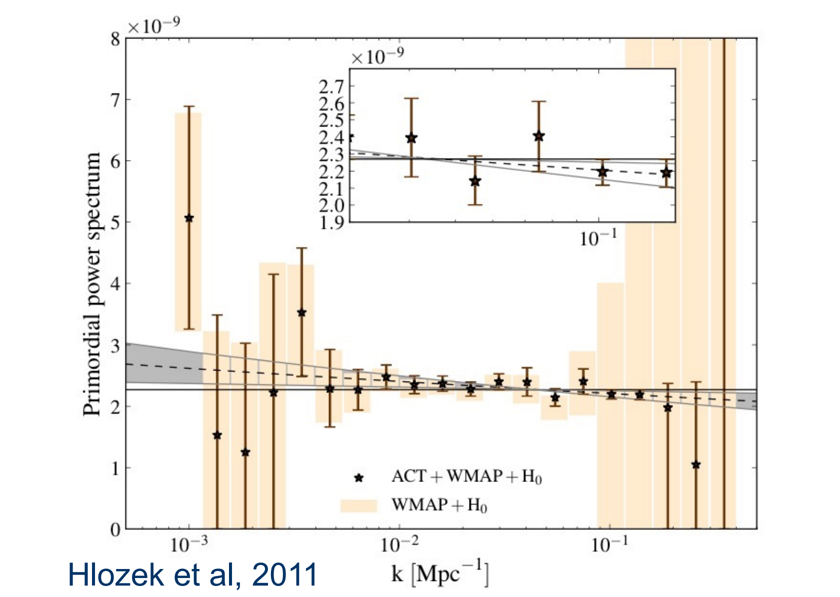
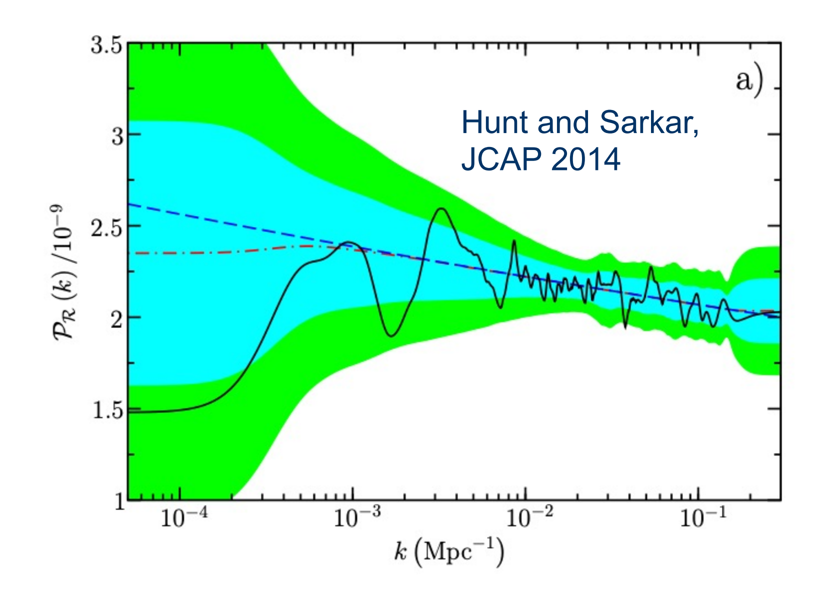


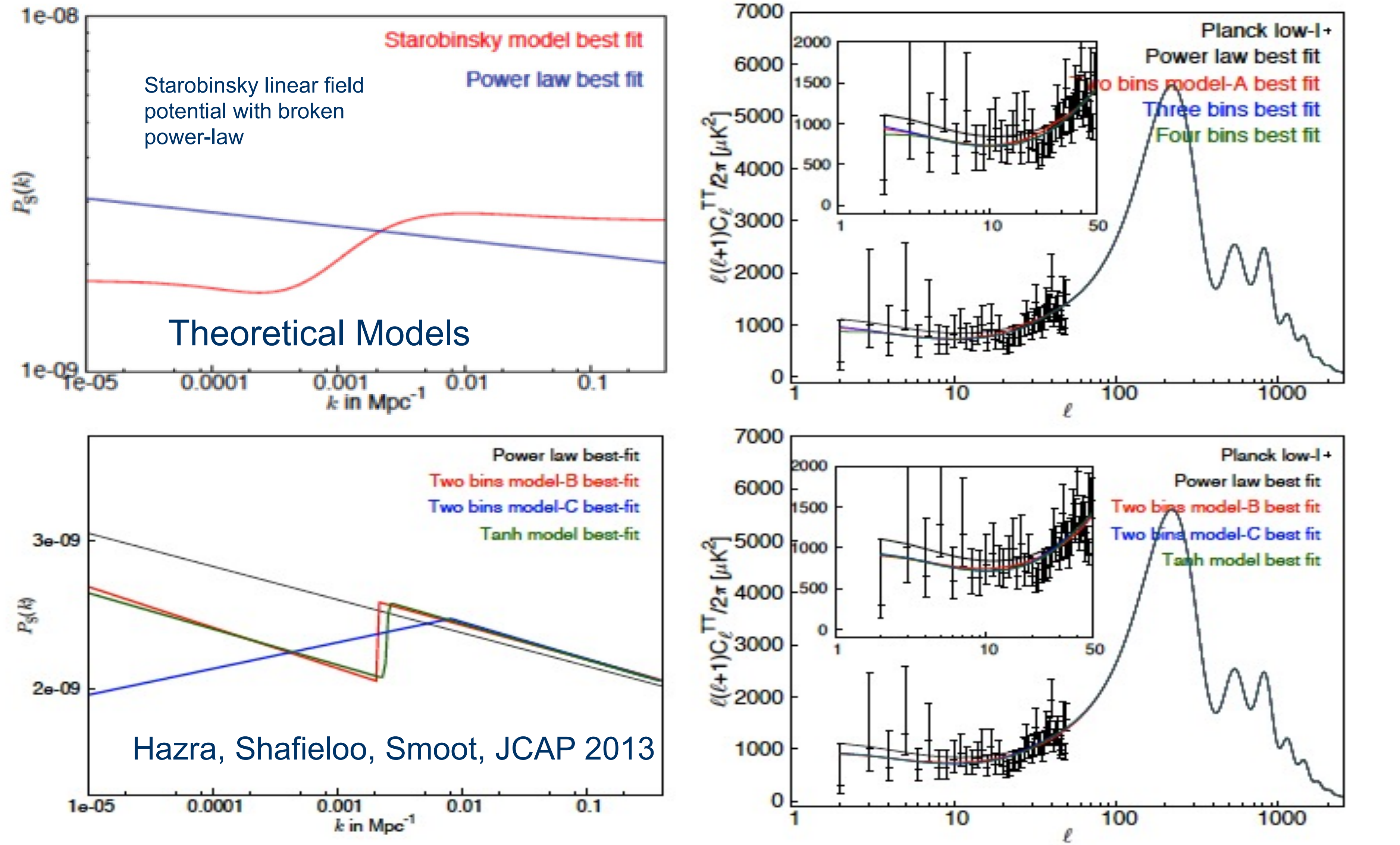
Figure 4. Reconstruction of the shape of the primordial power spectrum in 16 bands after marginalising over the Hubble constant, baryon and dark matter densities, and the redshift of reionization.







Beyond Power-Law: there are some other models consistent to the data.

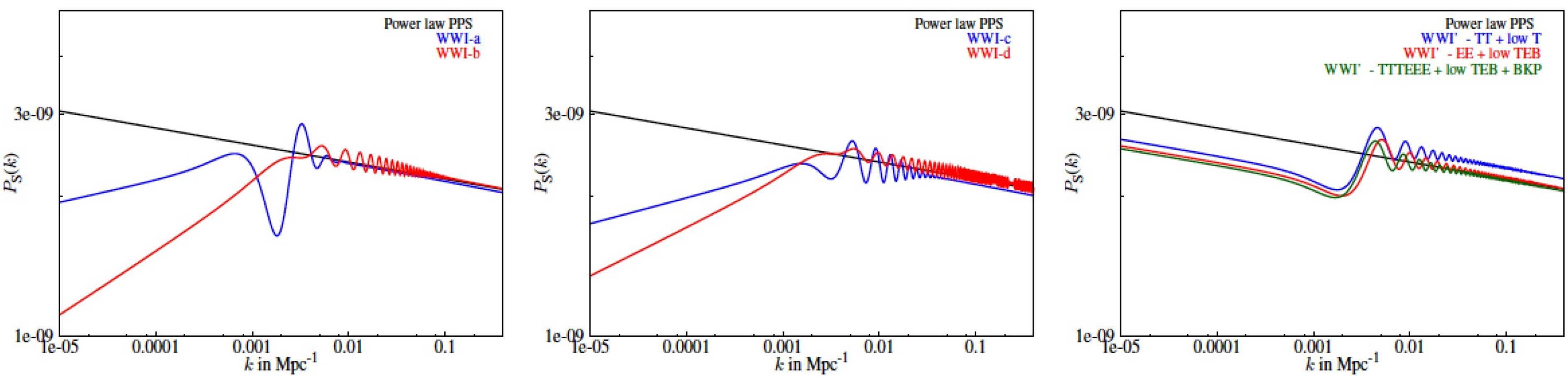


Individual likelihoods comparison						
Individual likelihood	Baseline	WWI-a $\Delta_{\mathrm{DOF}} = 4$	WWI-b $\Delta_{\mathrm{DOF}} = 4$	WWI-c $\Delta_{\mathrm{DOF}} = 4$	WWI-d $\Delta_{\mathrm{DOF}} = 4$	WWI′ $\Delta_{\mathrm{DOF}} = 2$
TT	761.1	762	761.9	762.8	762.8	762.4
lowT	15.4	8.2	13.4	12.1	13	10.2
Total	778.1	772.1 (-6)	777 (-1.1)	777 (-1.1)	778.4 (0.3)	775 (-3.1)
EE	751.2	748.8	747.2	748.6	750.2	746.8
lowTEB	10493.6	10490	10495.6	10492.4	10495.7	10492.2
Total	11248.8	11241.8 (-7)	11246.2 (-2.6)	11244.5 (-4.3)	11249.3 (0.5)	11242.3 (-6.5)
TTTEEE	2431.7	2432.7	2422.6	2427.8	2421.7	2426.5
lowTEB	10497	10490.8	10495.1	10493.4	10495.3	10492.7
Total	12935.6	12929.5 (-6.1)	12924.2 (-11.4)	12927.6 (-8)	12923.4 (-12.2)	12925.2 (-10.4)
TT	764.5	763.6	762.2	764.4	762.9	762.8
EE	753.9	754.8	750.5	750.8	750.8	751
TE	932	933.4	928.7	929.2	927	928.8
lowTEB	10498.4	10490.4	10495.8	10493.7	10495.6	10492.4
BKP	41.6	42	42	42.6	41.8	42.9
Total	12997	12991 (-6)	12985.9 (-11.1)	12987.2 (-9.8)	12985 (-12)	12985.1 (-11.9)
TTTEEE	2431.7	2432.8	2421.4	2426.7	2421	2425.7
lowTEB	10498.5	10490.5	10495.5	10493.6	10495.8	10492.6
BKP	41.6	42	42.7	42	41.9	42.5
Total	12978.3	12971.3 (-7)	12967.3 (-11)	12968.6 (-9.7)	12965 (-13.3)	12968.6 (-9.7)
TT (bin1)	8402.1	8404.1	8403.9	8405.2	8402.1	8401.9
lowT	15.4	8.3	13.3	11.9	13.2	10.3
Total	8419.6	8414.7 (-4.9)	8419.5 (-0.1)	8419.8 (0.2)	8418.1 (-1.5)	8414.4 (-5.2)
TTTEEE (bin1)	24158.2	24158.6	24149	24155	24148.4	24151.5
lowTEB	10497.6	10490.3	10493.4	10493.6	10495.3	10492.7
Total	34661.9	34655.3 (-6.6)	34650.5 (-11.4)	34654.4 (-7.5)	34649.5 (-12.4)	34650.6 (-11.3)

Beyond Power-Law: ***there are some other models consistent to the data.***

Whipped Inflation

Hazra, Shafieloo, Smoot, JCAP 2013
Hazra, Shafieloo, Smoot, Starobinsky, JCAP 2014A
Hazra, Shafieloo, Smoot, Starobinsky, JCAP 2014B
Hazra, Shafieloo, Smoot, Starobinsky, Phys. Rev. Lett 2014
Hazra, Shafieloo, Smoot, Starobinsky, JCAP 2016
Hazra et al, JCAP 2018
Debono, Hazra, Shafieloo, Smoot, Starobinsky, MNRAS 2020
Hazra, Paoletti, Debono, Shafieloo, Smoot, Starobinsky, JCAP 2021



Forms of PPS and Effects on the Background Cosmology

- Flat Lambda Cold Dark Matter Universe (LCDM) with power–law form of the primordial spectrum
- It has 6 main parameters.

Ω_b

Ω_m

H_0

τ

A_s

n_s

1

2

G(l,k)

$$C_l = \sum G(l,k)P(k)$$

2

3

1

$$C_l^{obs}$$

$$P(k) = A_s[\frac{k}{k_*}]^{n_s-1}$$

Forms of PPS and Effects on the Background Cosmology

- Cosmological parameter estimation with free form primordial power spectrum

Ω_b

Ω_m

H_0

τ

$\cancel{A_s}$

$\cancel{n_s}$

1

G(l,k)

3

2

$$C_l = \sum G(l,k)P(k)$$

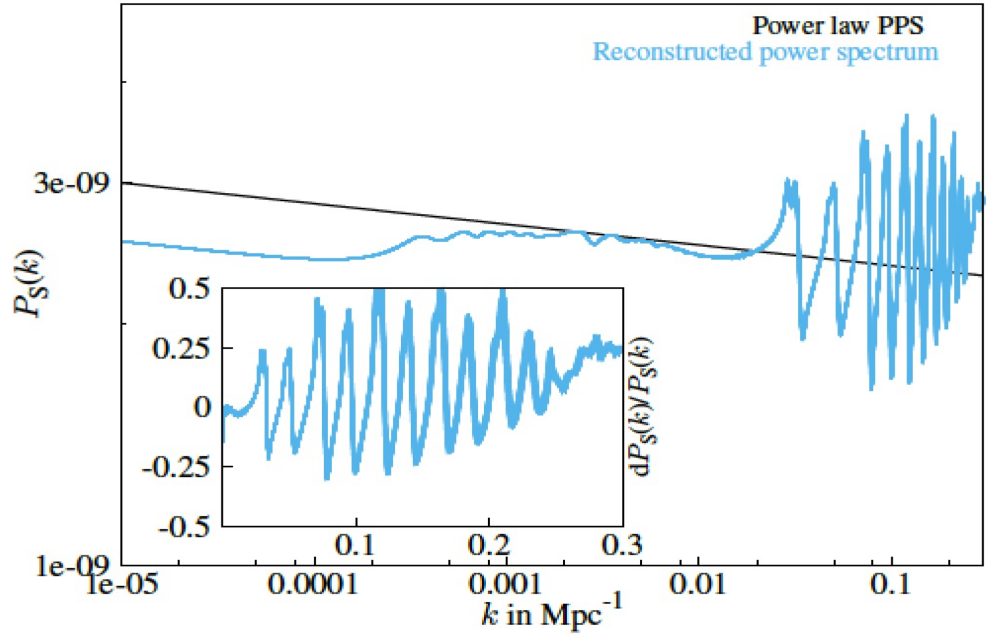
3

P(k)

4

C_l^{obs}

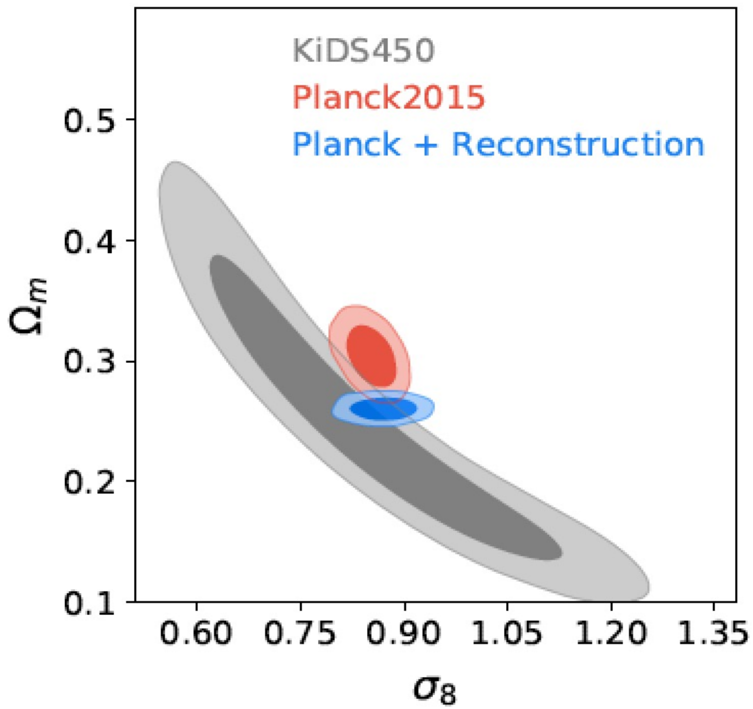
2

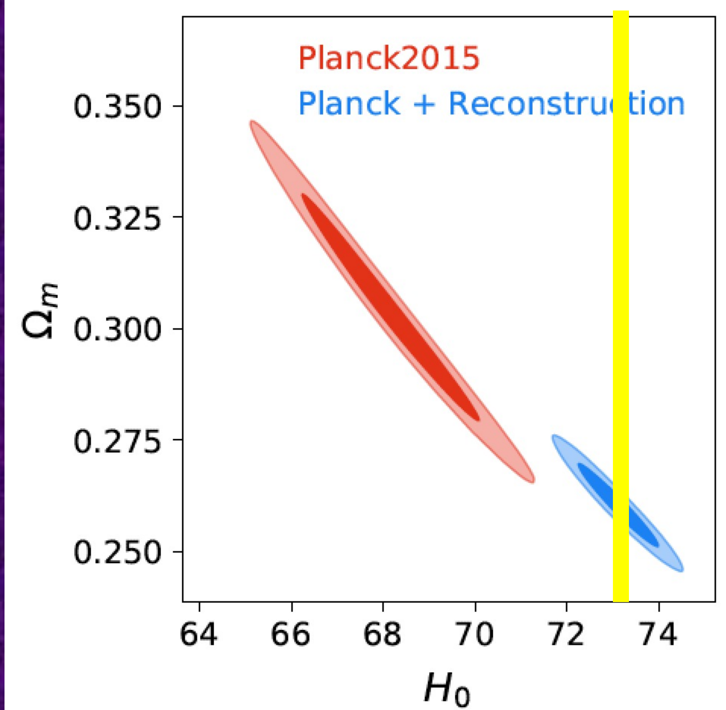


Hazra, Shafieloo, Souradeep, JCAP 2019
Keeley et al, MNRAS 2020

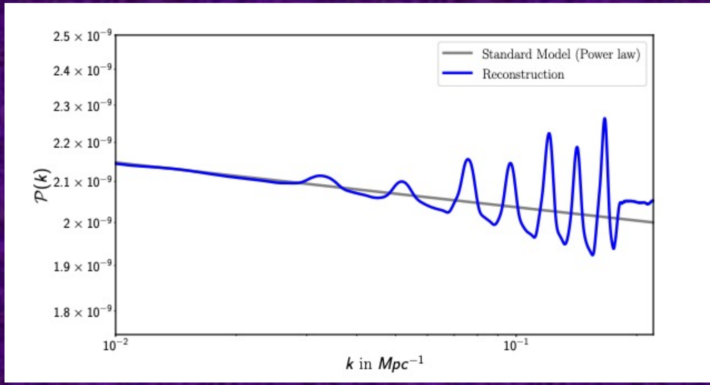
Background Cosmological Parameters and PPS

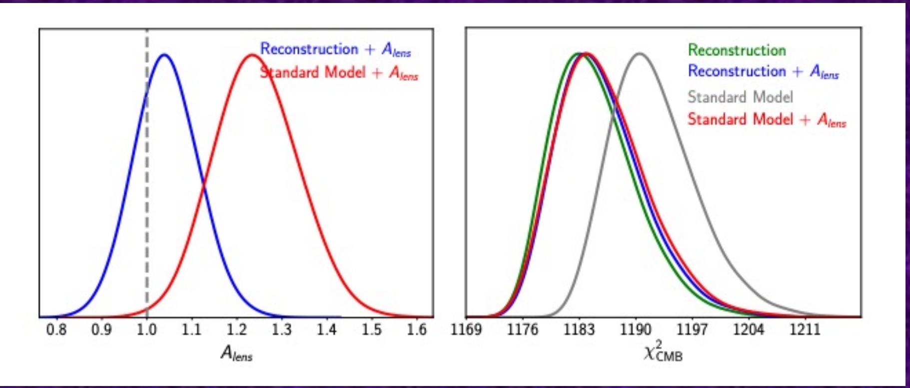
We use the reconstructed PPS for parameter estimation, similar to what we do with PL.



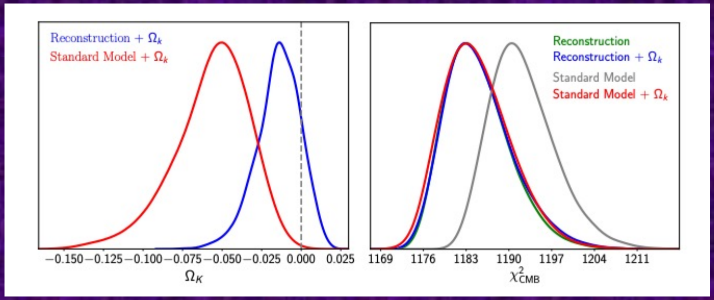


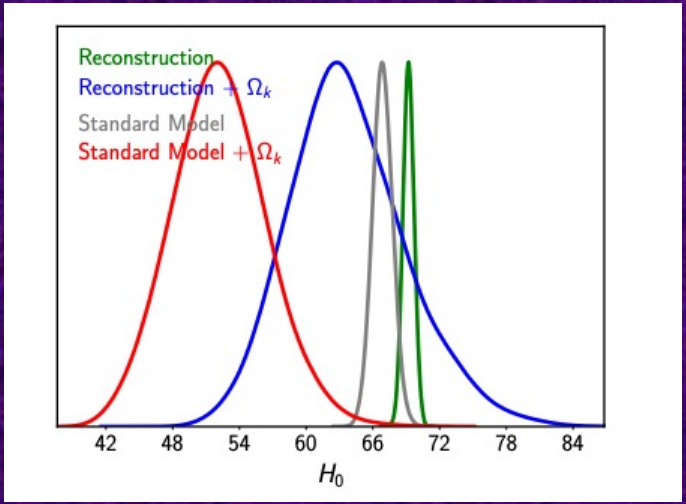
One spectrum to cure them all: looking for signature from early Universe to solve major anomalies and tensions in cosmology





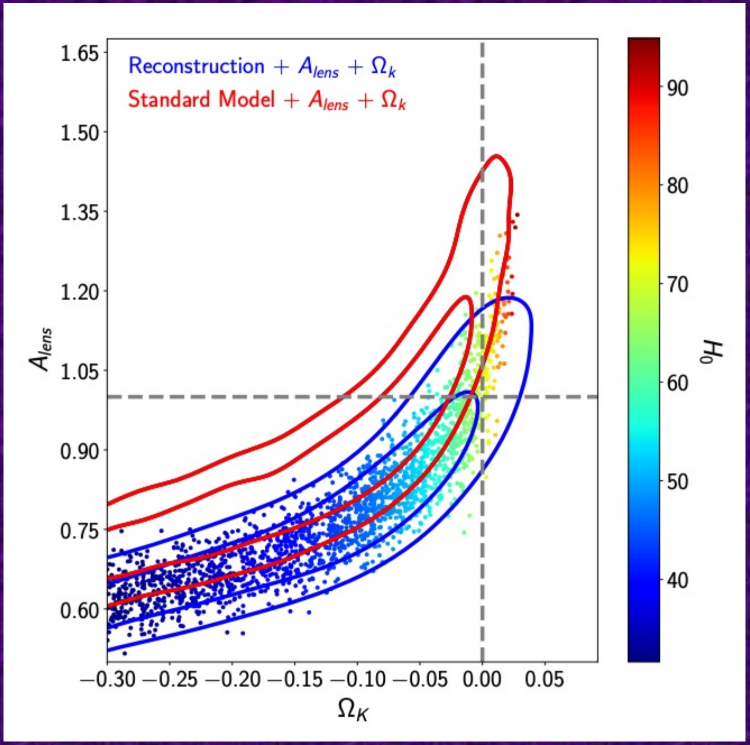
Curvature and A_lens anomalies

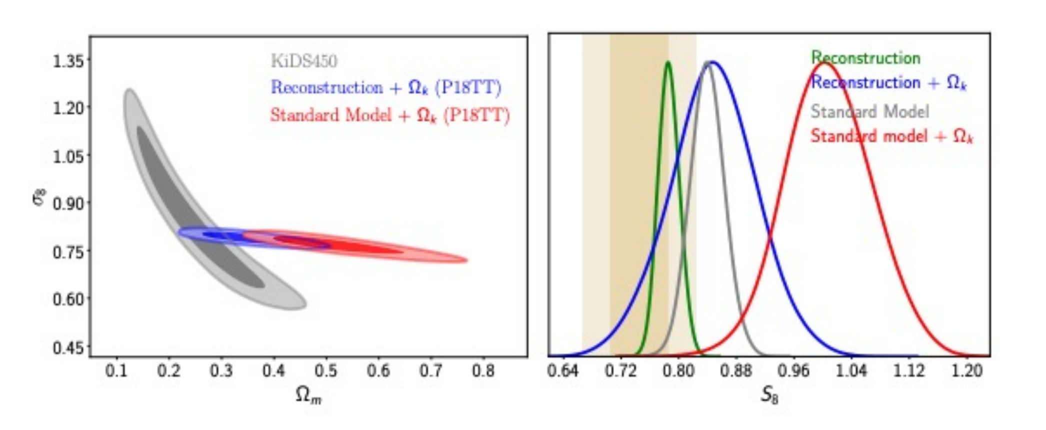




Hazra, Antony, Shafieloo : JCAP 2022

One spectrum to cure them all: Signature from early Universe solves major anomalies and tensions in cosmology





Hazra, Antony, Shafieloo :JCAP 2022

Addressing Majour Anomalies and tensions

Now we know what to look for!

Point 5

Reconstruction → Phenomenology → Theory

See Antony, Finelli, Hazra, Shafieloo, Phys Rev Lett 2023, for theoretical implication

Current Status

Open problem. Many tensions and hints for various systematics

Many theoretical/phenomenological models are proposed to ease the tensions. None is convincing so far (none can pass all validation tests).

Not possible to resolve all problems with minimal modification of the standard model. This has helped the standard model to survive so far.

Model independent consistency test between various data is essential to rule out systematics.

Point 6

Looking for systematics

Model independent consistency test between various data is essential to rule out systematics.

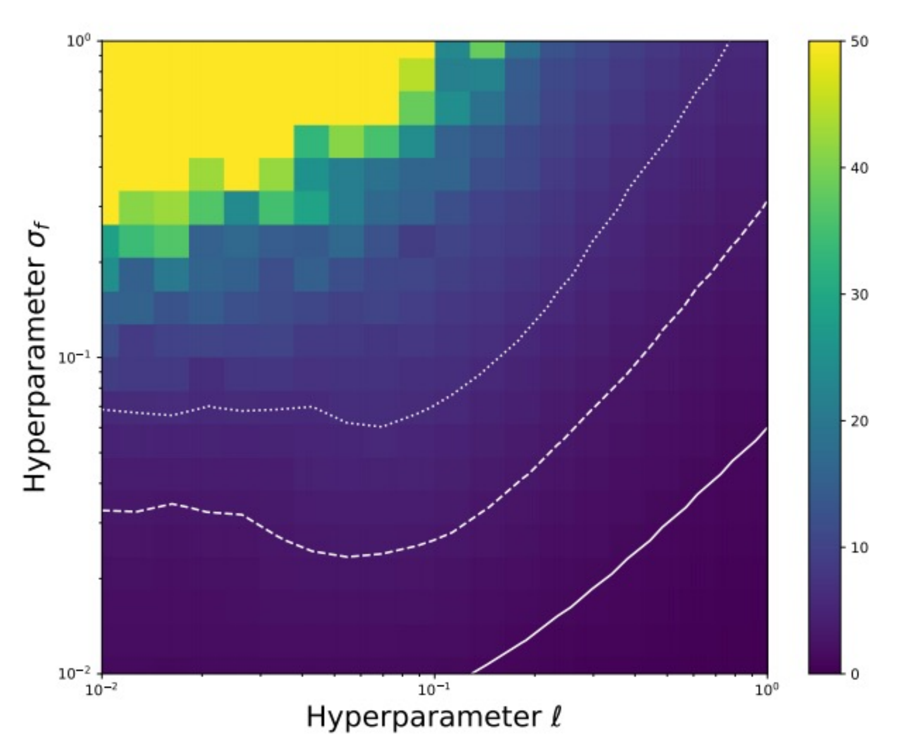
GP for Falsification

Shafieloo, Kim, Linder, PRD 2012
Shafieloo, Kim, Linder, PRD 2013
Hwang et al, JCAP 2023

Consistency of SDSS BAO and Pantheon SN Ia data
Keeley, Shafieloo, Zhao,…, MNRAS 2021
[arXiv:2010.03234] [SDSS IV paper]

$H0rd$ = 10040 ± 140 km/s and
Ωk = 0.02 ± 0.20

Point 6



Future Perspective

High possibilities for systematics in different data

Need for independent measurements

Two key questions:

Power-Law Primordial Power Spectrum?
Lambda Dark Energy?

Tip of the Red Giant Branch

Future Perspective

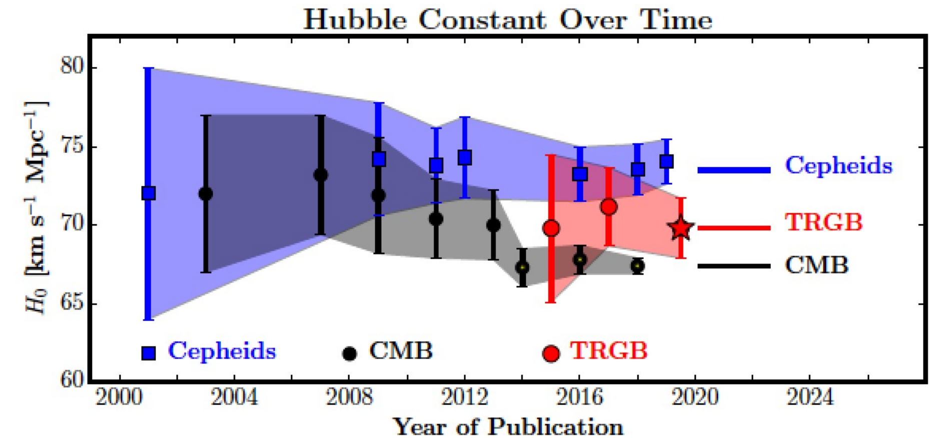
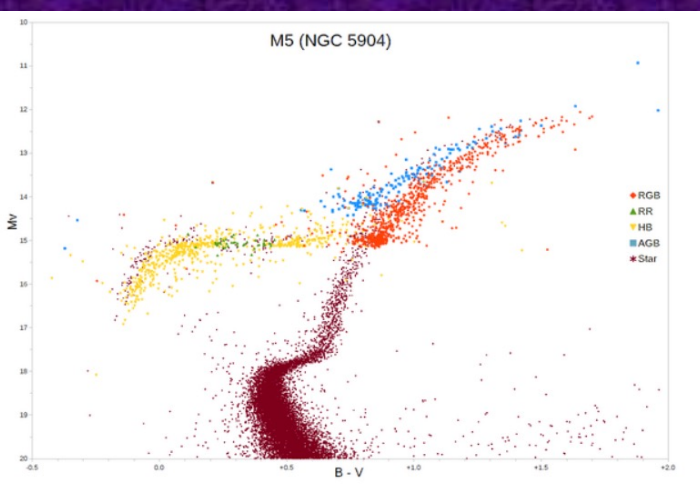


Figure 17. A plot of H_0 values as a function of time. The points and shaded region in black are those determined from measurements of the CMB; those in blue are Cepheid calibrations of the local value of H_0; and the red points are TRGB calibrations. The red star is the best-fit value obtained in this paper. Error bars are 1σ.

Freedman et al, arXiv:1907.05922



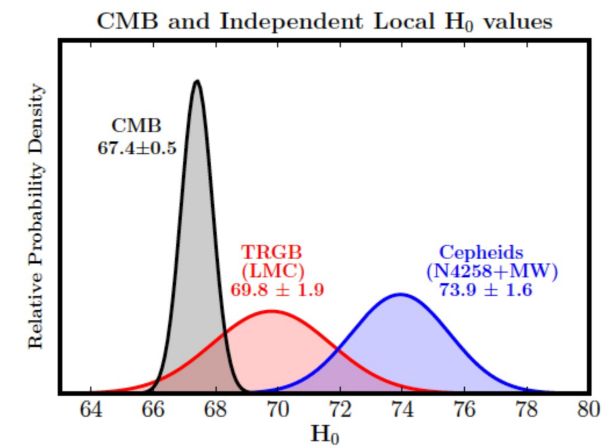
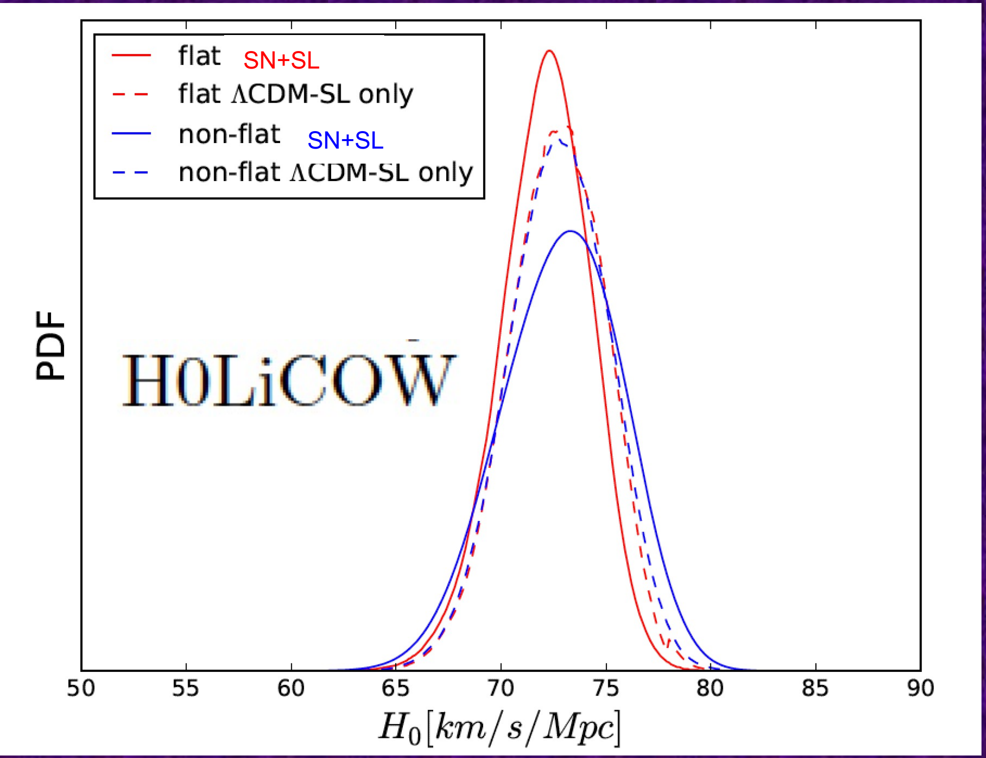


Figure 18. Completely independent calibrations of H_0. Shown in red is the probability density function based on our LMC CCHP TRGB calibration of CSP-I SNe Ia; in blue is the Cepheid calibration of H_0 (Riess et al. 2016), using the Milky Way parallaxes and the maser distance to NGC 4258 as anchors (excluding the LMC). The Planck value of H_0 is shown in black.

Cosmology with Strong Lens Systems: Has become already competative!



Liao, Shafieloo, Keeley, Linder, ApJ Letters 2019

H0 from Strongly Lensed systems

$$H_0 = 72.8^{+1.6}_{-1.7}\ \mathrm{km/s/Mpc}$$

2.3% model-independent measurment of Hubble constant

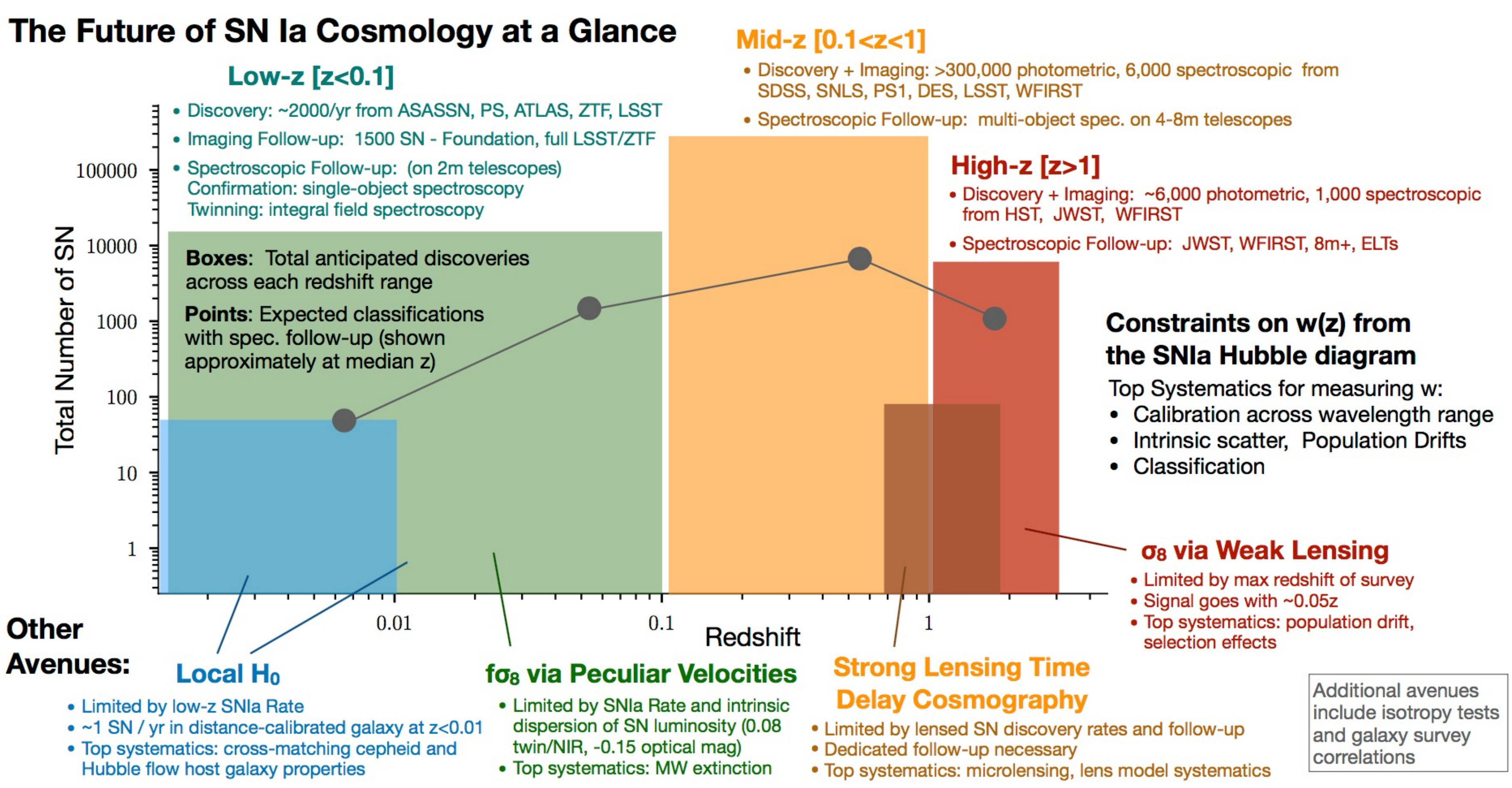
Liao, Shafieloo, Keeley, Linder, ApJ Letters 2020

H0LiCOW I. H0 Lenses in COSMOGRAIL's Wellspring

Suyu et al. MNRAS 2017

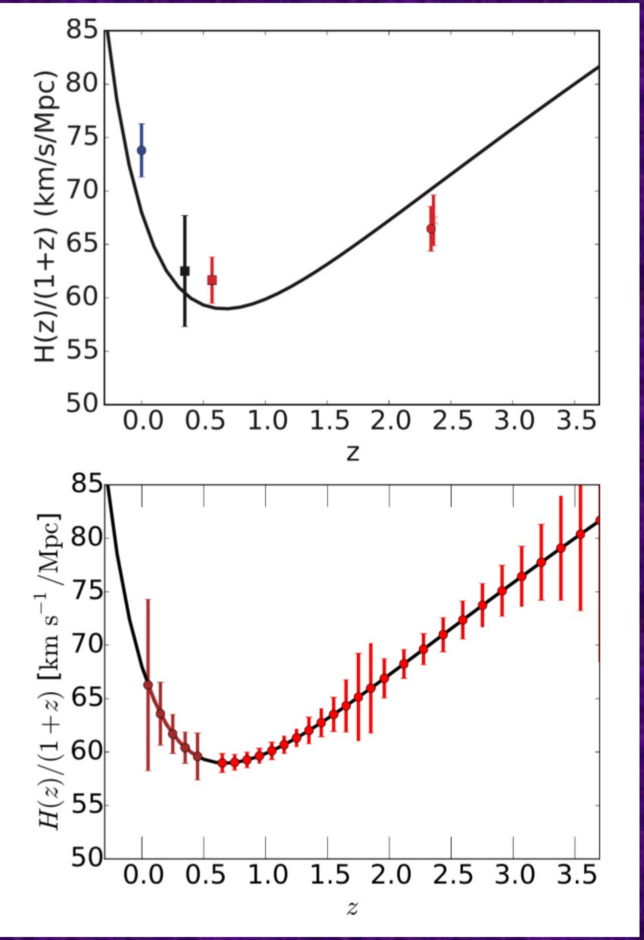
Order	Name	z_L	z_S
1	RXJ1131-1231	0.295	0.654
2	HE 0435-1223	0.4546	1.693
3	B1608+656	0.6304	1.394
4	SDSS 1206+4332	0.745	1.789

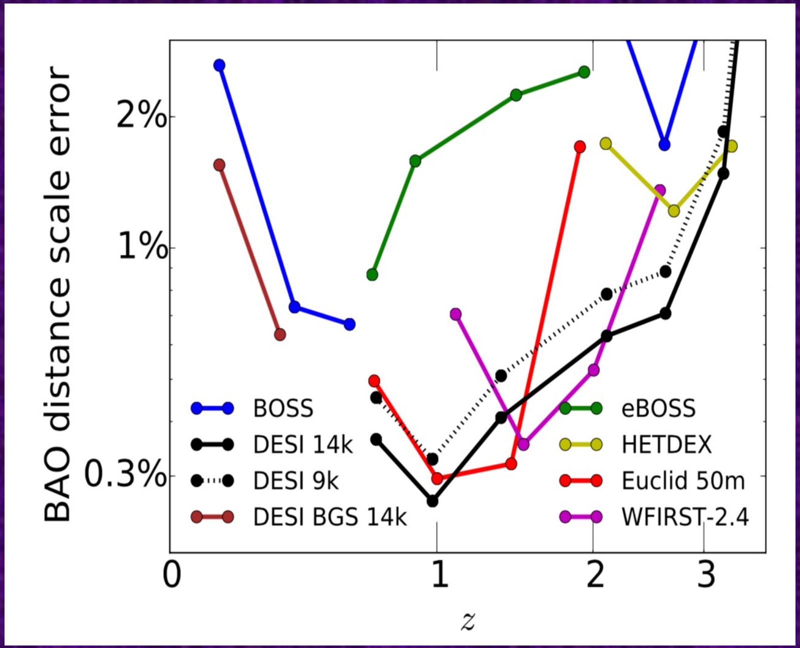
Future perspective (late universe, SN Ia)

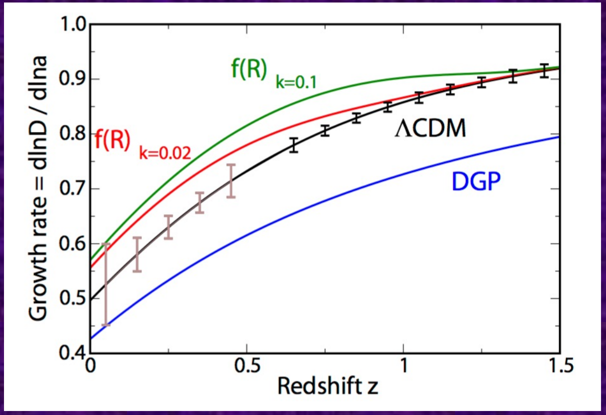


Scolnic,et al, arXiv:1903.05128

Future perspective (late universe; BAO & RSD)

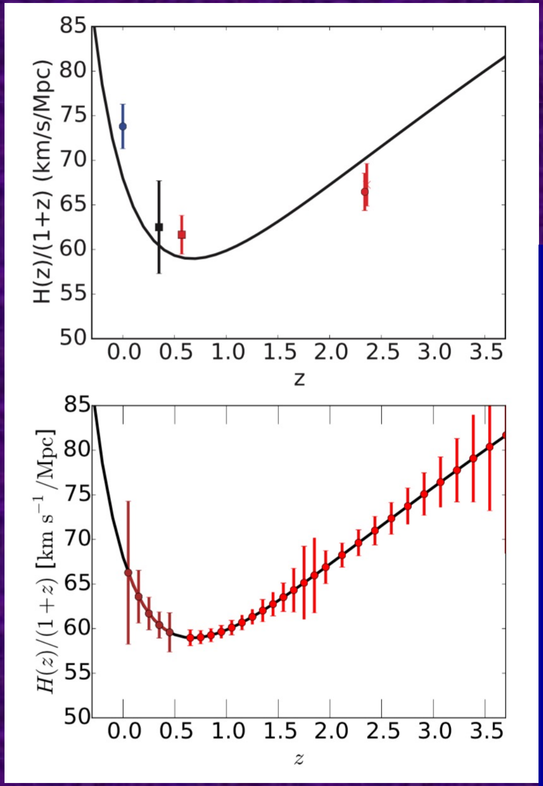






Aghamousa et al, [arXiv:1611.00036]
DESI Collaboration

Future perspective (late universe; BAO, RSD)



DESI Y1 data will be released soon and it will be **better than all existing LSS data combined**

arXiv:2306.06307 **DESI SV**
arXiv:2306.06308 **DESI EDR**

Aghamousa et al, [arXiv:1611.00036] DESI Collaboration

Future perspective [G-Waves and Standard Sirens]

Astro2020

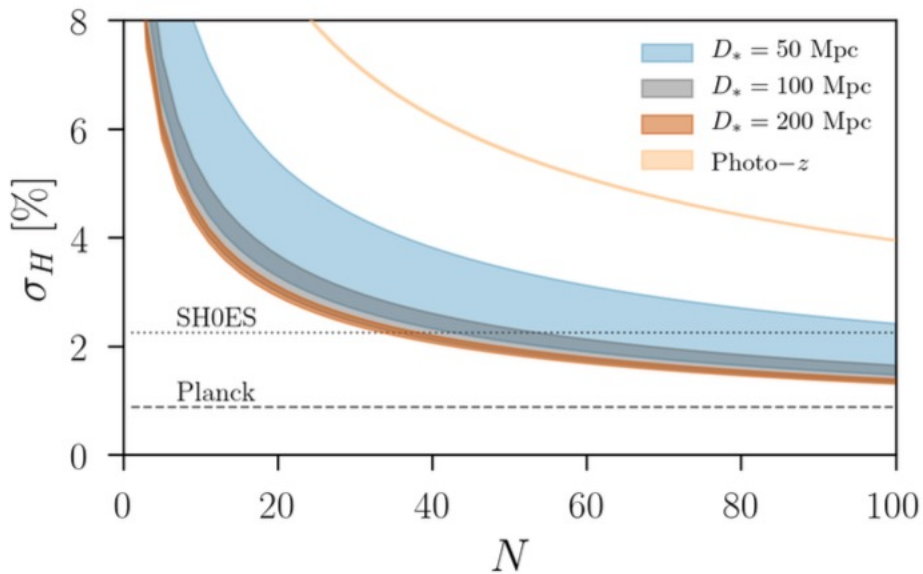


Figure 1: Hubble constant uncertainty (1σ) as a function of combined GW events with associated EM counterpart. The shaded regions show the impact of the peculiar velocity uncertainty between 100 and 400 $\mathrm{km\,s^{-1}}$ for different distance reaches D_*. The latest results from standard candles (SH0ES, [13]) and CMB (*Planck*, [14]) are also shown.

Palmese et al,
arXiv:1903.04730

Future Perspective (primordial)

$$C_\ell^{TT} = \int \frac{dk}{k} P(k) \quad G_\ell^{TT}(k)$$

$$C_\ell^{EE} = \int \frac{dk}{k} P(k) \quad G_\ell^{EE}(k)$$

$$C_\ell^{BB} = \int \frac{dk}{k} P_t(k) \quad G_\ell^{BB}(k)$$

$$C_\ell^{TE} = \int \frac{dk}{k} P(k) \quad G_\ell^{TE}(k)$$

Full picture

Complete reconstruction analysis with polarization data

Searching for correlations!

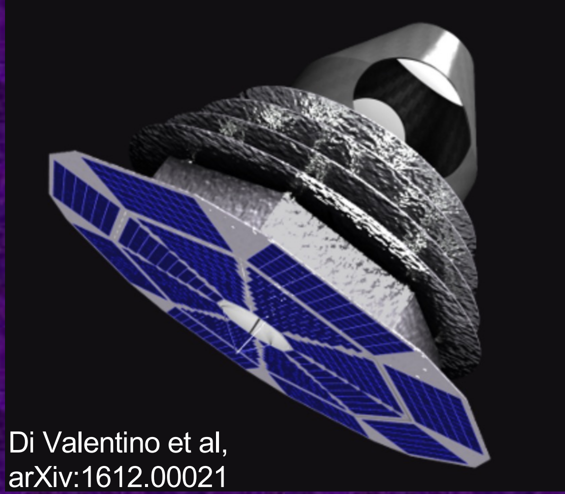
$$P_S(k), P_T(k), P_{iso}(k)$$

Primordial power spectra from Early universe

$$G_\ell^{TT}(k), G_\ell^{EE}(k), G_\ell^{BB}(k), G_\ell^{TE}(k)$$

Post recombination Radiative transport kernels in a given cosmology

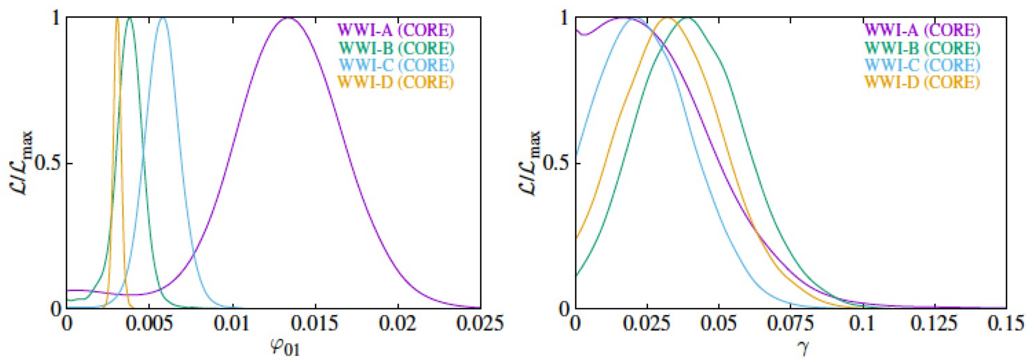
Features with Future of CMB (S4)



Di Valentino et al,
arXiv:1612.00021

With Cosmic Origins Explorer (CORE)-like survey specification

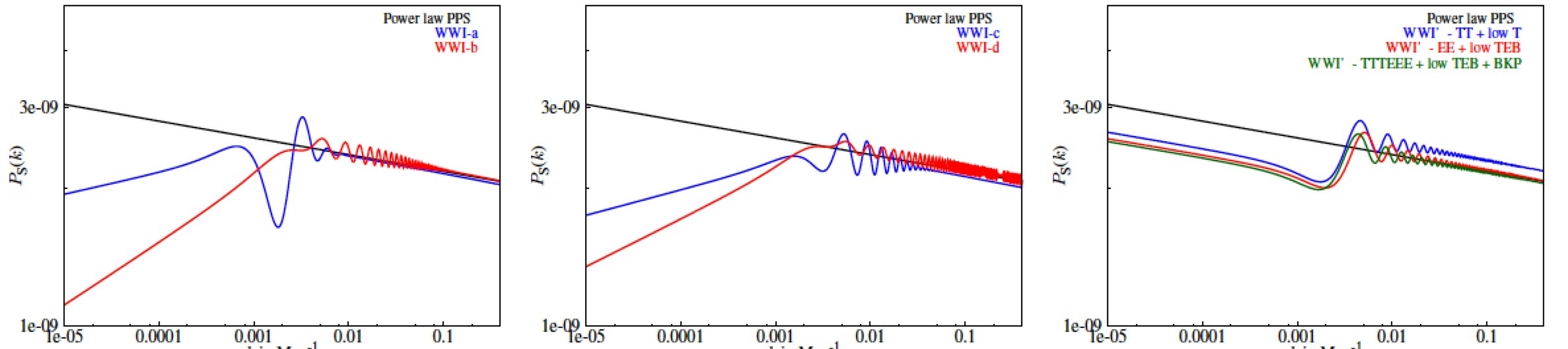
Wiggles **Suppression**



Distinguishing EU models

Hazra et al, JCAP 2018
Debono et al, MNRAS 2020

- Large scale suppressions can not be detected with high significance
- Some of the intermediate and small scale oscillations can be detected, if present



Future Perspective

From 2D to 3D

Using LSS data to test early universe scenarios

1. We need to estimate matter power spectrum but we observe galaxies. Hence we have to model linear clustering bias and estimate its parameters accurately and precisely to connect the observables to theory. Bias modeling would be different for different surveys and susceptible to systematics.

1. Does power spectrum (or bi-spectrum, etc) necessarily contains all the information in 3D data of LSS? Can't reducing dimensionality of the data wash out some information?

Data will be hugely better…but we have to be careful!

Cosmology vs Systematics

- With higher quality of the data the role of systematics will become more and more prominent.
- Higher precision may cost us uncontrollable bias if we make wrong assumptions.

Point 6

What we should be worried about!

Conclusion

- Many statistical tools are not used appropriately in cosmology and astrophysics and results can be strong but invalid conclusions.
- H0 tension (and some others) seems remaining persistent in the context of the LCDM model. This can open ways for competitive alternatives (GEDE?, EDE, features in PPS?) but we should not over-sell these models.
- ***Tensions are not resolved with minimal extensions of the standard model and there is no clear resolution.*** It is highly possible, from statistical point of view, that there are ***systematics*** in some of the data and we might need ***new physics*** too. ***It can be a combination of both!*** New independent measurements and observations can help to clear things up.
- With higher quality data, the effect of systematics and wrong assumptions are much more prominent in introducing substantial inaccuracies. This is a real challenge to avoid making big fake discoveries.

Conclusion

- Standard Model of Cosmology fits different data pretty well ***individually*** but there are tensions fitting different combinations of the data.

- H0 tension (and some others) seems remaining persistent in the context of the LCDM model. This can open ways for competitive alternatives ***(GEDE?, EDE, features in PPS?)***.

- ***Tensions are not resolved with minimal extensions of the standard model and there is no low redshift resolution.*** It is highly possible that there are ***systematics*** in some of the data and we might need ***new physics*** too. ***It can be a combination of both!*** New independent measurements and observations can help to clear things up.

- ***First target*** can be testing different aspects of the standard model. If it is not **'*Lambda*' dark energy or '*power-law*' primordial spectrum** then we can look further. It is possible to focus the power of the data for the purpose of the falsification. ***Next generation of astronomical observations***, (DESI, Euclid, LSST, WFIRST, SKA(?), etc) will make it much more clear about the status of the concordance model in 2020s.