Prior effects and the cosmological tensions

Elisa G. M. Ferreira

Kavli IPMU & University of Sao Paulo

2nd Shanghai Assembly on Cosmology and Structure Formation

03/November/2023

In this talk

Statistics is in the core of precision cosmology and any of its discoveries - parameter inference, model selection and validation

Important to correctly interpret the observations with the ***correct statistical tool*** and using the valid assumptions

Bayesian vs frequentist

Particularly important with the increase of high quality data, where systematics will play important role: stage IV (even stage III), cosmological tensions, …

In this talk:

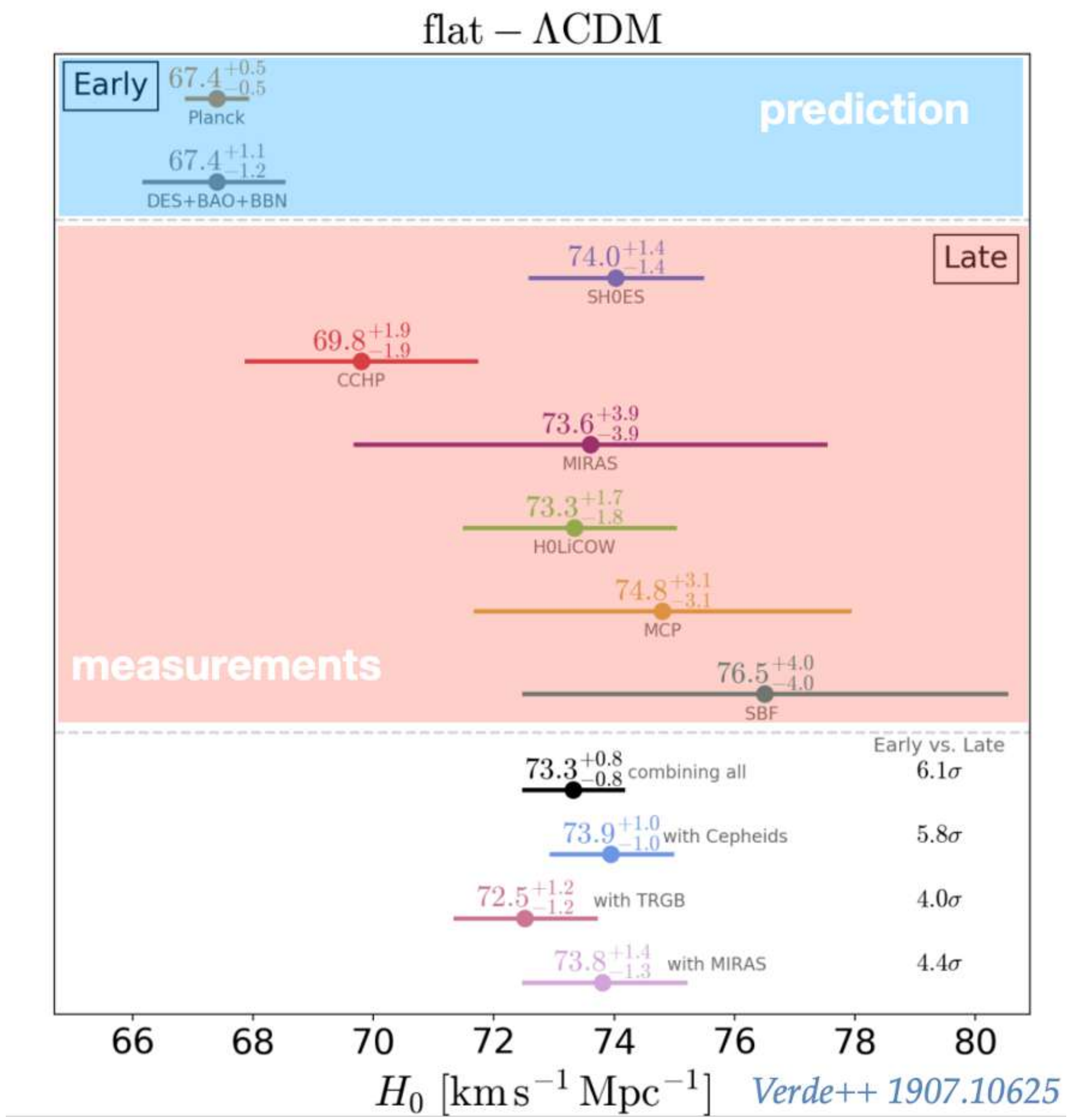
MCMC effect: Prior volume effect ⟶ when large, bias in the result and what to do

➡ *Profile likelihood*

Hubble tension

Indirect

Direct



→ Depends on the cosmological model

New SHOES result (*Riess et al 2021*):

$$H_0 = 73.04 \pm 1.04\,\mathrm{km/s/Mpc}$$

5σ tension with Planck!

↓

$$H_0 = 67.71 \pm 0.40\,\mathrm{km/s/Mpc}$$

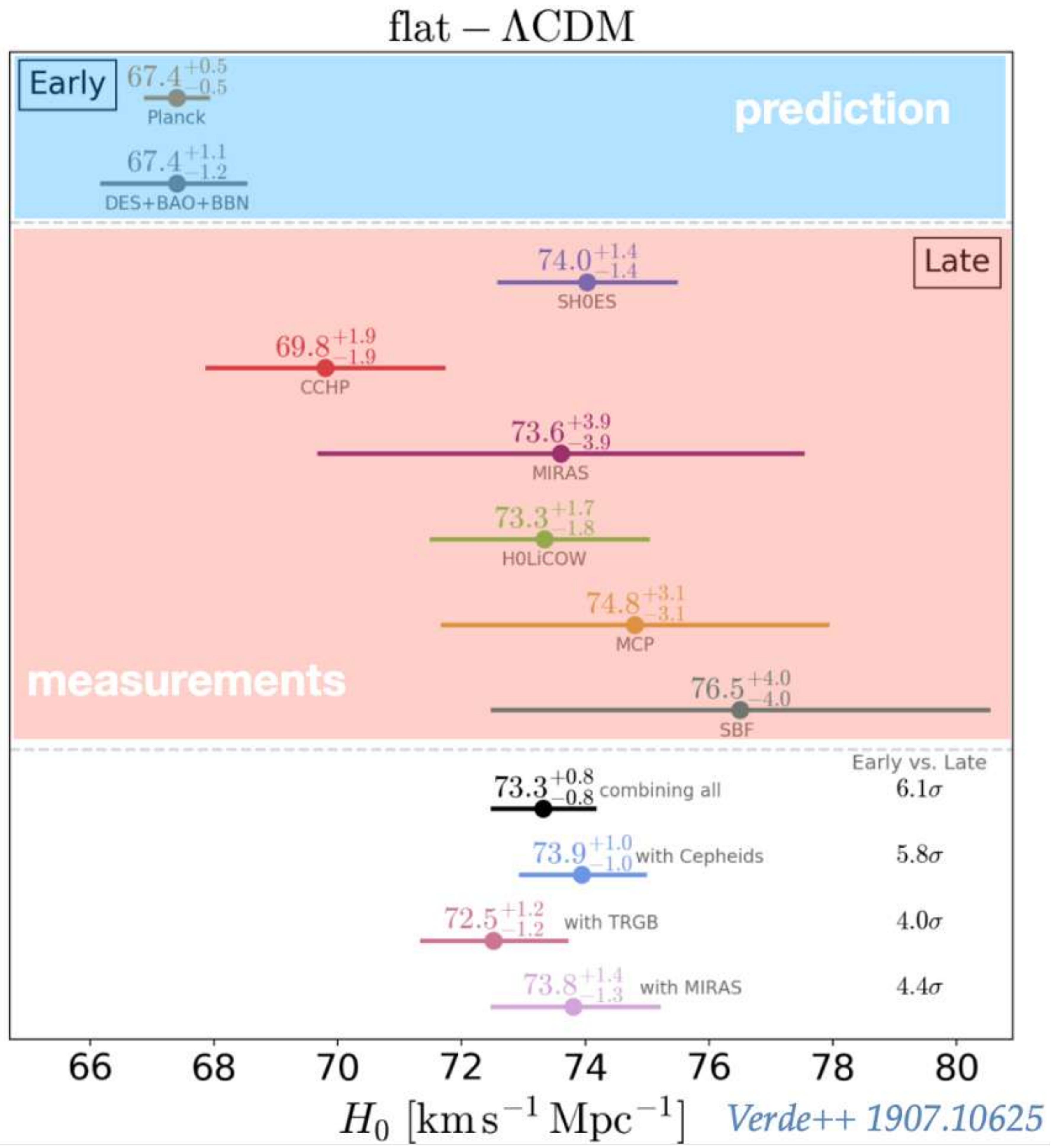
(Planck+BAO+Sn)

Systematics or new physics?

Hubble tension

Indirect

Direct



Depends on the cosmological model

New SHOES result (*Riess et al 2021*):

$$H_0 = 73.04 \pm 1.04\,\mathrm{km/s/Mpc}$$

5σ tension with Planck!

$$H_0 = 67.71 \pm 0.40\,\mathrm{km/s/Mpc}$$

(Planck+BAO+Sn)

Systematics or new physics?

The H_0 Olympics: A fair ranking of proposed models,
Schöneberg + (2021)

	ΔN_p	
Majoron	3	2
primordial B	1	3
varying m_e	1	1
varying m_e+Ω_k	2	1
EDE	3	2
NEDE	3	2
EMG	3	2

Early dark energy

Based on Laura Herold, EF 2210.1629
& Laura Herold, EF and Eiichiro Komatsu 2112.12140,
& A. Reeves, L. Herold, S. Vagnozzi , B. Sherwin, EF 2207.01501

Early dark energy

Idea: add an extra component (to increase $H(z)$) that starts acting around equality, behaves as DE and dilutes faster than matter

acting around equality → ϕ_i

behaves as DE → Initially frozen, Slow-rolls

Early dark energy

Idea: add an extra component (to increase $H(z)$) that starts acting around equality, behaves as DE and dilutes faster than matter

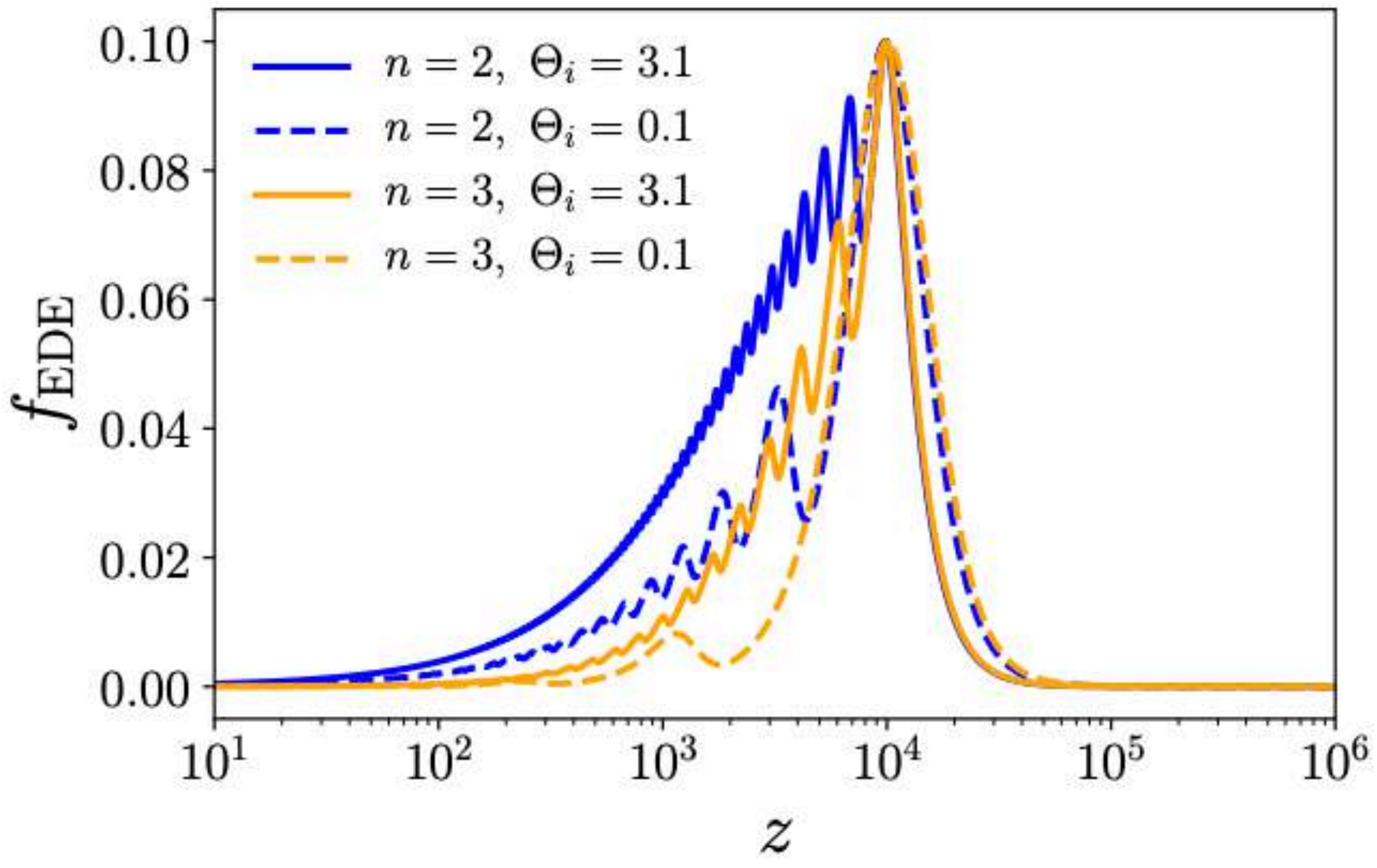
$$V(\phi) = V_0\,[1 - \cos(\phi/f)]^n$$

3 free parameters: $\{m, f, n\}$ + IC: ϕ_i

- V_0 or m $(V_0 = m^2 f^2)$: the field is ultra-light $m \sim H(z_{eq}) \sim 10^{-27}\,\text{eV}$
- f (spont. sym. breaking scale)
- n : controls the decay → needs to be hidden at late times

$$n \geq 2\ \ (w \geq 1/3)$$

Smith et al 2019



Phenomenological parameters $\{f_{\rm EDE}(z_c),\ z_c,\ n,\ \theta_i = \phi_i/f\}$

$$f_{\rm EDE}(z_c) \equiv \left.\frac{\rho_{\rm EDE}}{\rho_{\rm tot}}\right|_{z_c} = \left.\frac{\rho_{\rm EDE}}{(3M_{pl}^2 H^2)}\right|_{z_c}$$

We usually fix $n = 3$

$\Longrightarrow$ $\{f_{\rm EDE}(z_c),\ z_c,\ \theta_i\}$

Matteo Braglia et al 2020, Adrià Gómez-Valent et al 2021, …

Early dark energy

Does EDE really solves the H_0 tension?

2019

Early dark energy can resolve the Hubble tension

EDE from CMB

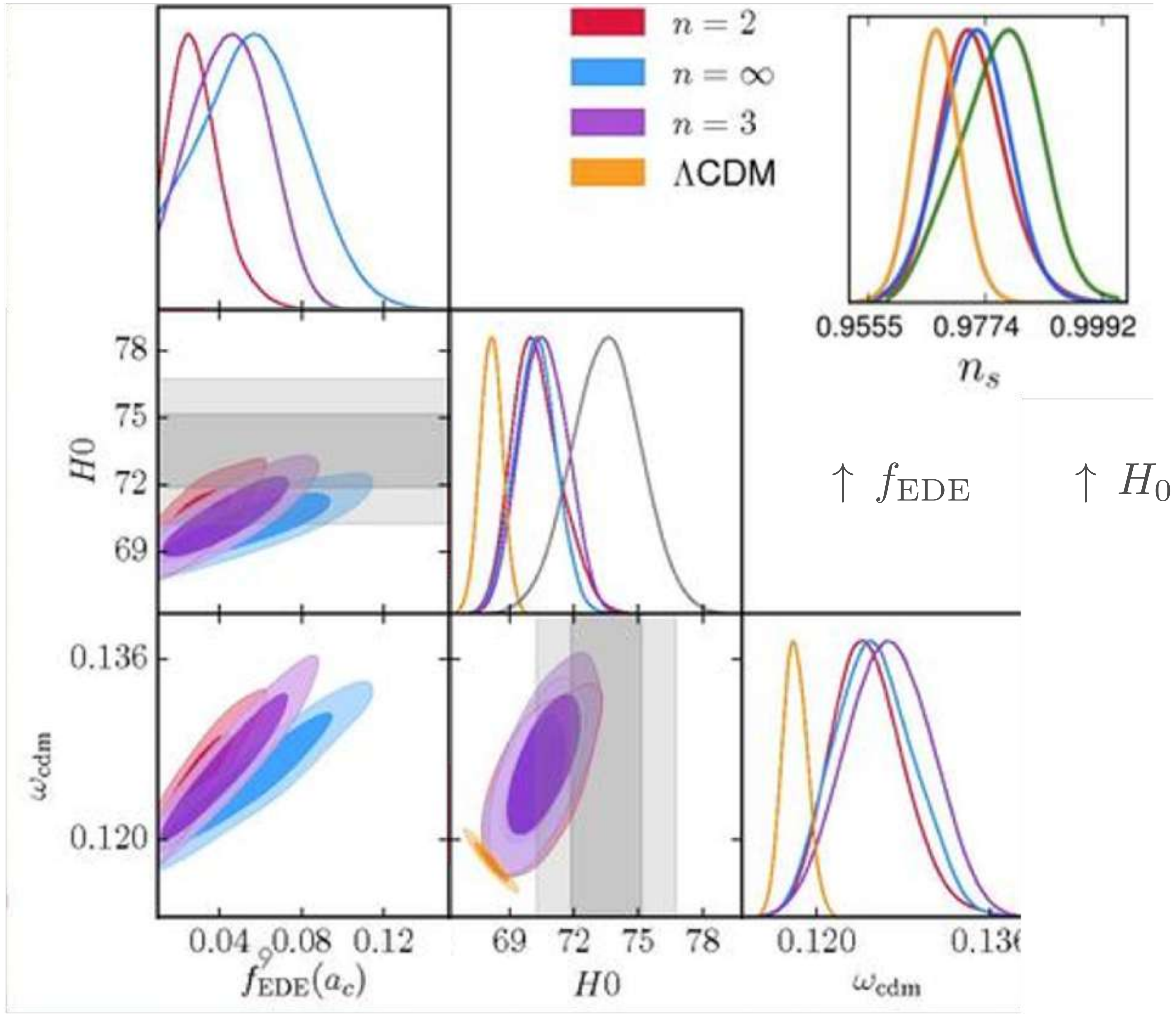
V. Poulin, T. Smith, T. Karwal, M. Kamionkowski, 2019

- For: *Planck + BOSS DR12 BAO/RSD + 6dFGS + Pantheon + SHOES 2016*

$$H_0 = 71.49 \pm 1.20\,\mathrm{km/s/Mpc}$$

Solves the tension!

$\uparrow f_{\mathrm{EDE}}$ $\uparrow H_0$



$\uparrow f_{\mathrm{EDE}}$ $\uparrow H_0$

But with:

- More DM
- Higher n_s

$$w_{cdm} = \Omega_{cdm}\, h^2$$

Early dark energy does NOT *restore cosmological concordance*

EDE from LSS

2020

Use LSS to constrain EDE

- CMB: *Planck* 2018 TT, TE, EE
- LSS
  - *Planck* lensing
  - "Compressed" likelihood
    - BAO
  - Weak lensing from KIDS+VIKING-450 + HSC
  - FULL SHAPE OF THE PS

Early Dark Energy Does Not Restore Cosmological Concordance

J. Colin Hill,[1,2] Evan McDonough,[3] Michael W. Toomey,[4] and Stephon Alexander[3]

Constraining Early Dark Energy with Large-Scale Structure

Mikhail M. Ivanov,[1,2] Evan McDonough,[3] J. Colin Hill,[4,5] Marko Simonović,[6] Michael W. Toomey,[7] Stephon Alexander,[7] and Matias Zaldarriaga[8]

Uniform prior in the phenomenological parameters

$$f_{\rm EDE}, z_c, \theta_i$$

Early dark energy does NOT *restore cosmological concordance*

EDE from LSS

- For: *Planck* + *BOSS DR12 BAO/RSD* + ~~*SHOES 2016*~~ + *full-shape of PS*

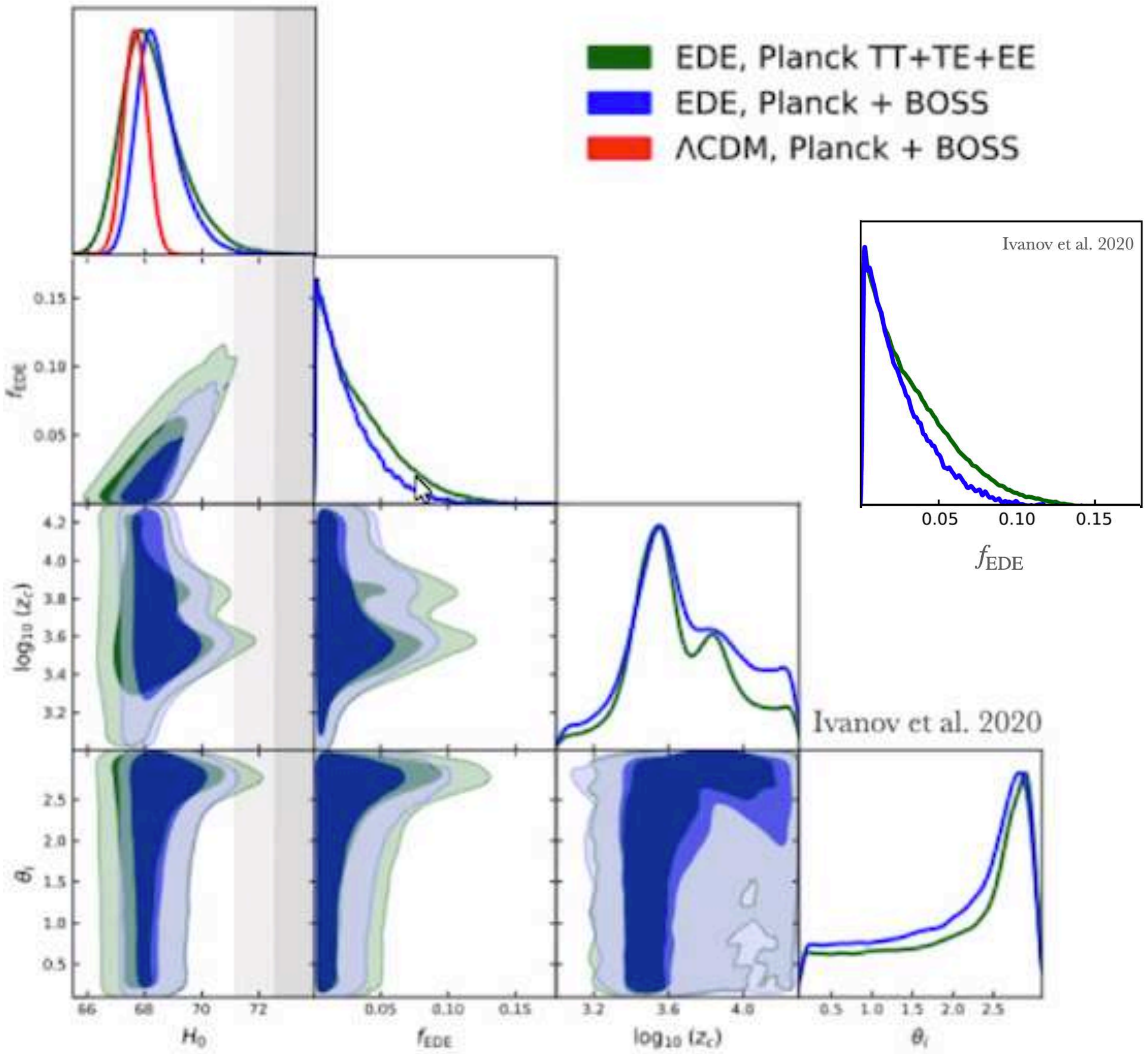
Constraints from *Planck* 2018 data + BOSS DR12		
Parameter	ΛCDM	EDE ($n = 3$)
$\boldsymbol{f_{\rm EDE}}$	–	$< 0.072\,(0.047)$
H_0 [km/s/Mpc]	$67.70\,(67.56)\ \pm 0.42$	$68.54\,(68.83)^{+0.52}_{-0.95}$
$\Omega_{\rm m}$	$0.3105\,(0.3112)^{+0.0053}_{-0.0058}$	$0.3082\,(0.3120)^{+0.0056}_{-0.0057}$

$$H_0 = 68.54^{+0.52}_{-0.95}\ \mathrm{km/s/Mpc}$$

$f_{\rm EDE} < 0.072$ (95 % CL)

3.6σ tension with SHOES !!

Adding S_8 prior $\longrightarrow$ $f_{\rm EDE} < 0.058$



Early dark energy does NOT solve Hubble tension!

Wait…

Early dark energy is ***NOT*** excluded by current LSS data

T. Smith et al (2020)

Niedermann, Sloth (2019)

Not all groups agree with this result Volume effects?

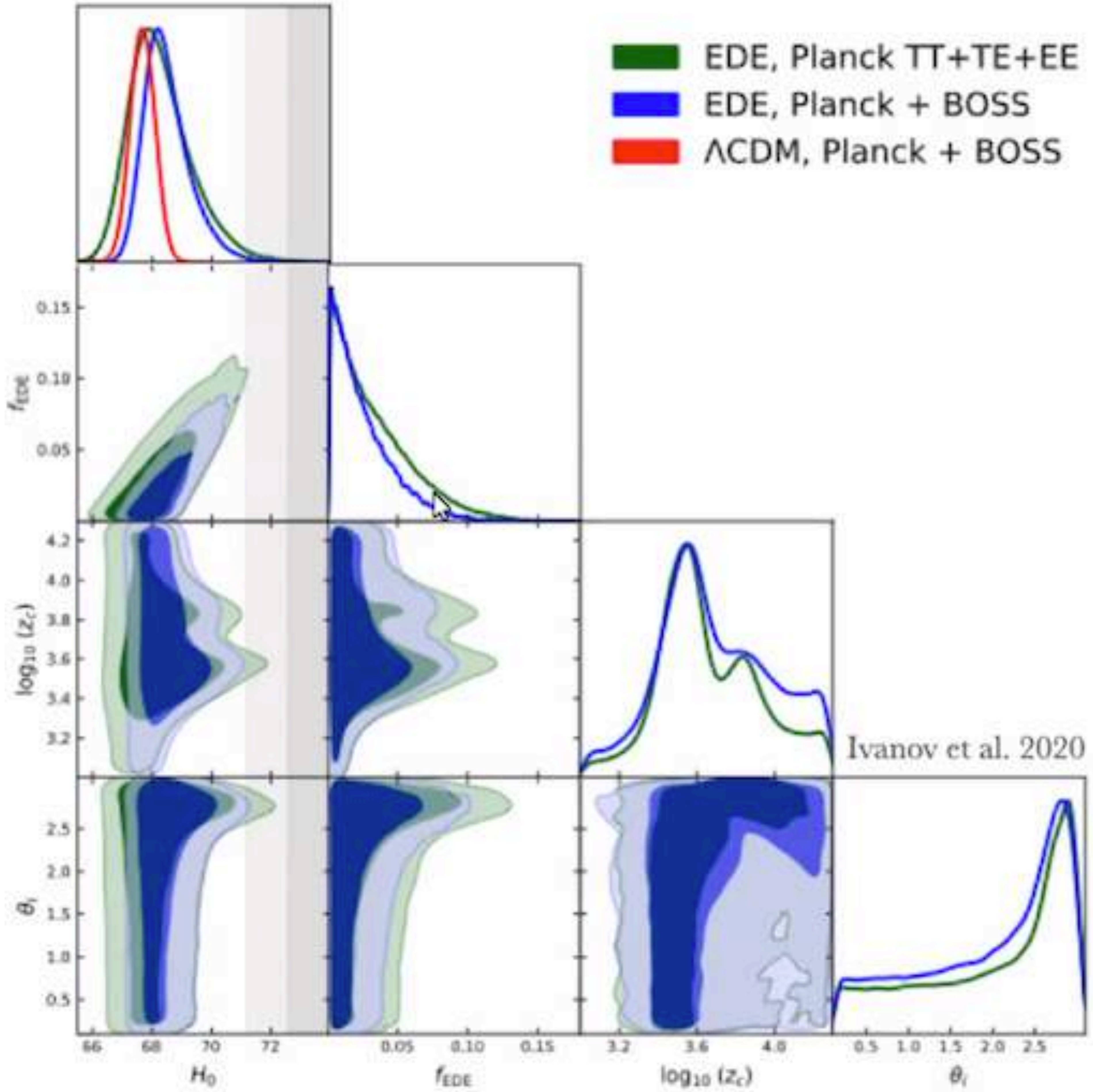
- Previous result can be a consequence of choice of priors of the EDE parameters

Volume effects: $f_{\rm EDE} \to 0$, any value of $\log(z_c)$ and θ_i, degenerate with ΛCDM

↓

Marginalization: preference for $f_{\rm EDE} \sim 0$

• $\log(z_c)$ and θ_i are not well constrained by data



Ivanov et al. 2020

→ Ivanov et al checked for volume effects in their paper, finding no evidence

Prior volume effects

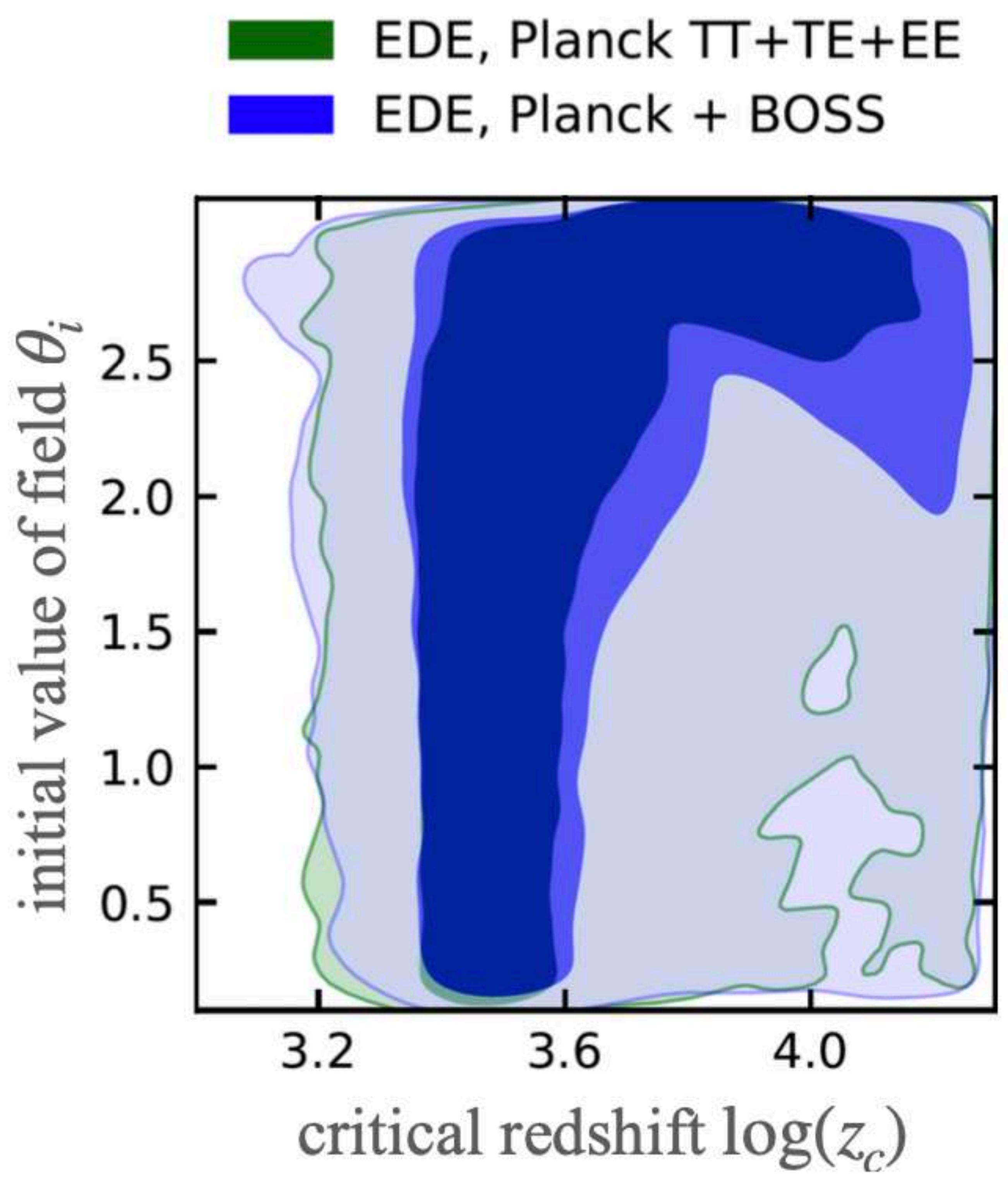
Bayesian marginalization of the full-dimensional posterior involves integrating out the nuisance dimensions

Since in addition to the value of the posterior, an integral is sensitive to the volume in these directions

→ Large parameter regions (of possibly non-maximal posterior values) are emphasized compared to smaller regions (of possibly larger posterior values).

Inescapable feature of the Bayesian method!!

(volume effect can occur even with flat priors)



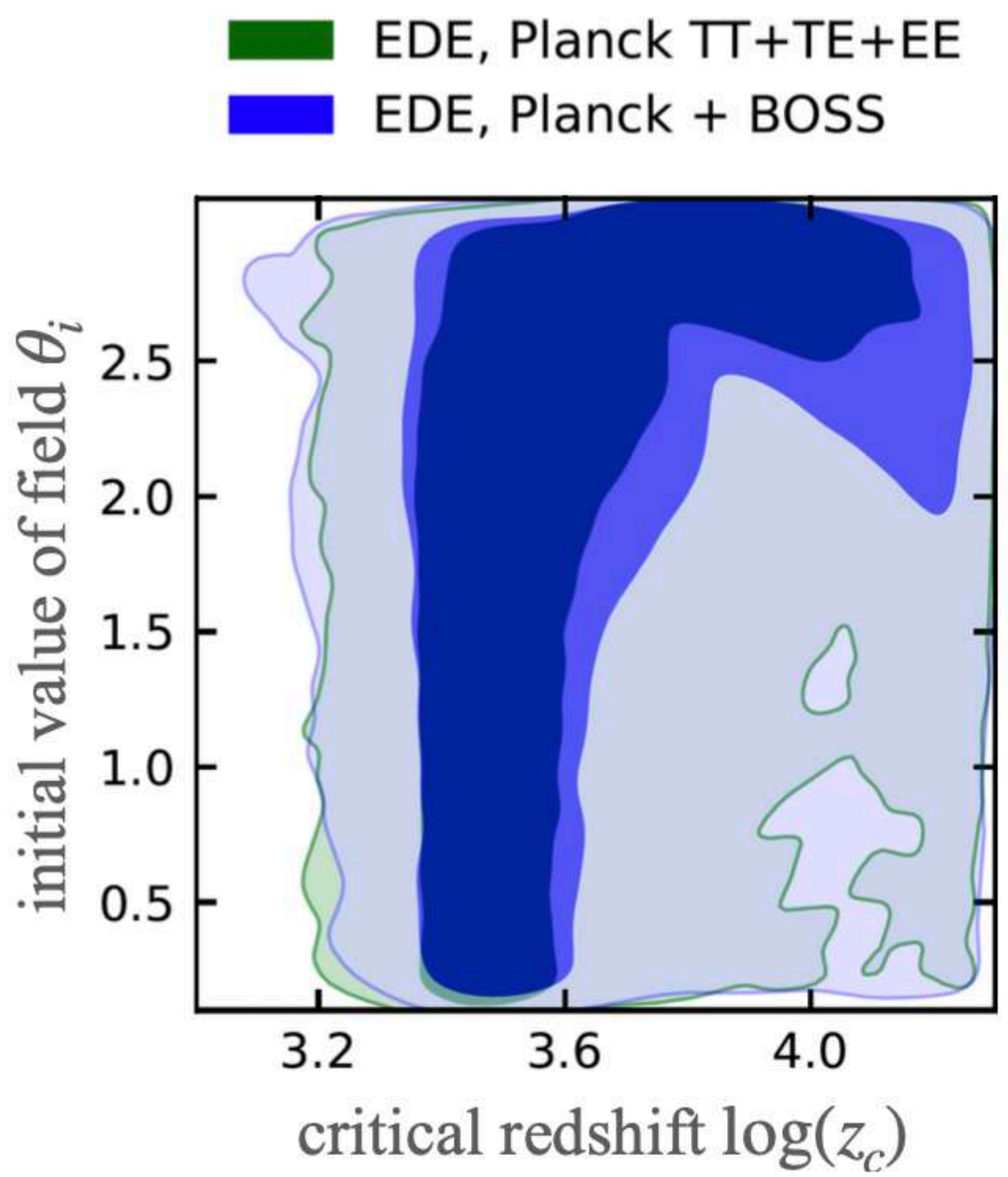
Ivanov et al. 2020

Prior volume *effects*

Prior volume effects or marginalization effects: ...appear if the posterior is dominated by the prior volume

When they appears:

- Model has too many parameters / data is not constraining.
- Posterior is very non-Gaussian.
- Parameter structure of the model generates large volume differences.



Ivanov et al. 2020

Prior volume effects

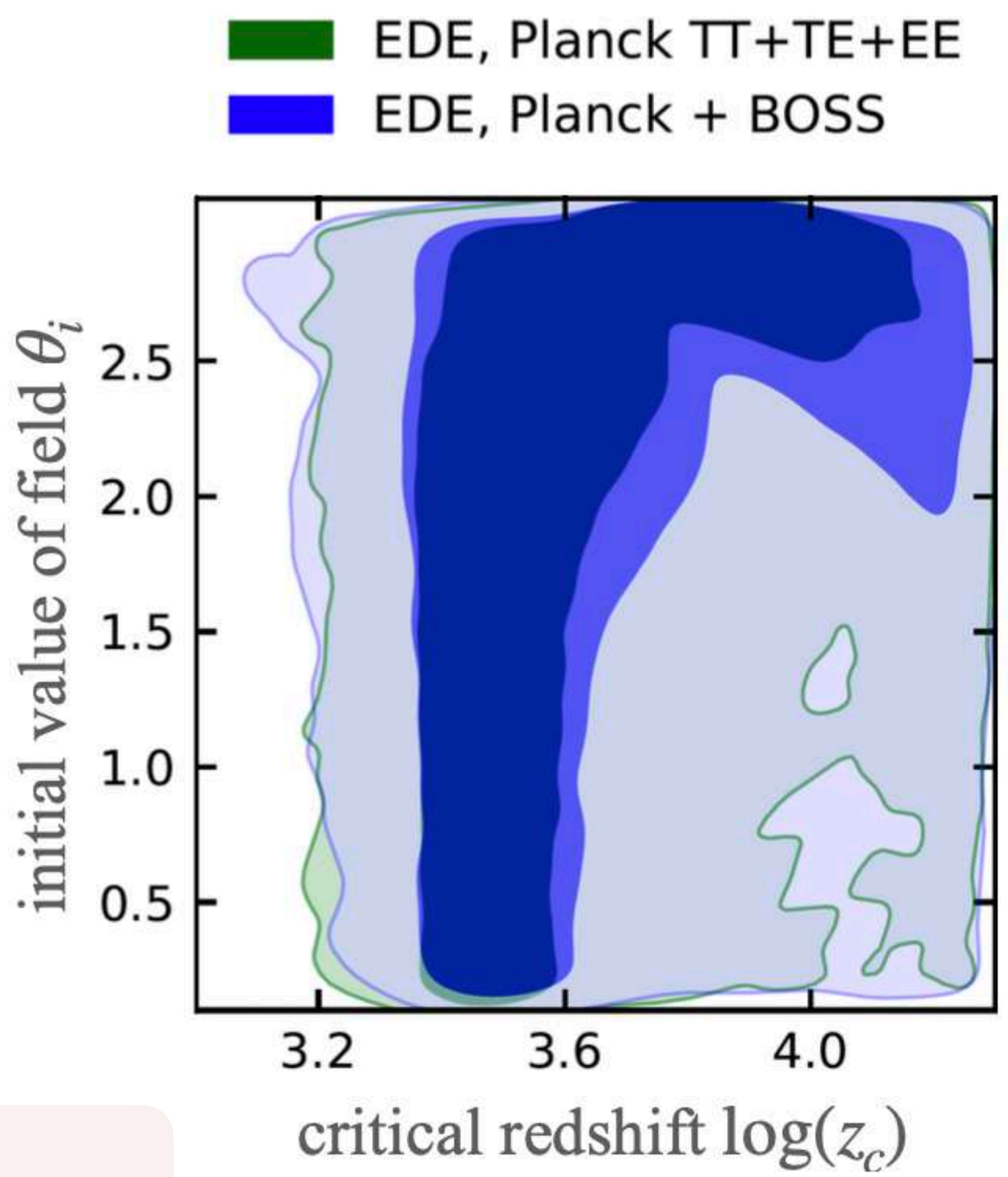
Prior volume effects or marginalization effects: ...appear if the posterior is dominated by the prior volume

When they appears:

- Model has too many parameters / data is not constraining.
- Posterior is very non-Gaussian.
- Parameter structure of the model generates large volume differences.

⟶ Bias* in the marginalized posterior

** offset induced by a projection effect*



Ivanov et al. 2020

Relevant to study the extent to which one's results are affected by volume effects!

Profile likelihood

Motivation:

Frequentist method for comparison with Bayesian to check for prior or marginalization effects

What is the profile likelihood $L(\theta)$ or $\chi^2(\theta)$?

Profile likelihood is a method in frequentist statistics, that allows to treat nuisance parameters

By splitting the full parameter space Θ into two categories:

- θ of N parameters
- ν of M (nuisance) parameters

⇒ Profile likelihood of θ is obtained by **maximization** over all parameters in the complementary set of (nuisance) parameters ν for **fixed** $\boldsymbol{\theta}$

Nuisance parameters

$$L(\boldsymbol{\theta}) = \max_{\boldsymbol{\nu}} \underbrace{L(\boldsymbol{\theta}, \boldsymbol{\nu})}_{\text{Full likelihood function}},$$

Bayesian and frequentist

Bayesian statistics

Bayesian inference: derive the posterior probability as a consequence of two antecedents: a prior probability and a "likelihood function" derived from a statistical model for the observed data.

Posterior — Likelihood — Prior

$$P(H \mid E) = \frac{P(E \mid H) \cdot P(H)}{P(E)}$$

H: hypothesis
E: evidence

- Likelihood: probability of observing E *given H*
- *Prior*: Probability of H *before* E is observed
- Posterior: probability of H *given* E, i.e., *after* E is observed. Probability of a hypothesis *given* the observed evidence.

Arman Shafieloo : "Priors and simplicity of the proposed model also matters (in model comparison)"

Frequentist approach

Frequentist statistics never uses or calculates the probability of the hypothesis, while Bayesian uses probabilities of data and probabilities of both hypothesis.

Frequentist methods do not demand construction of a prior and depend on the probabilities of observed and unobserved data.

Arman Shafieloo : "Assuming a proposed model, the probability of the observed data must not be insignificant"

Profile likelihood

Motivation:

Frequentist method for comparison with Bayesian to check for prior or marginalization effects

What is the profile likelihood $L(\theta)$ or $\chi^2(\theta)$?

Profile likelihood is a method in frequentist statistics, that allows to treat nuisance parameters

By splitting the full parameter space Θ into two categories:

- θ of N parameters
- ν of M (nuisance) parameters

⇒ Profile likelihood of θ is obtained by **maximization** over all parameters in the complementary set of (nuisance) parameters ν for **fixed θ**

Nuisance parameters

$$L(\boldsymbol{\theta}) = \max_{\boldsymbol{\nu}} \underbrace{L(\boldsymbol{\theta}, \boldsymbol{\nu})}_{\text{Full likelihood function}},$$

Profile likelihood

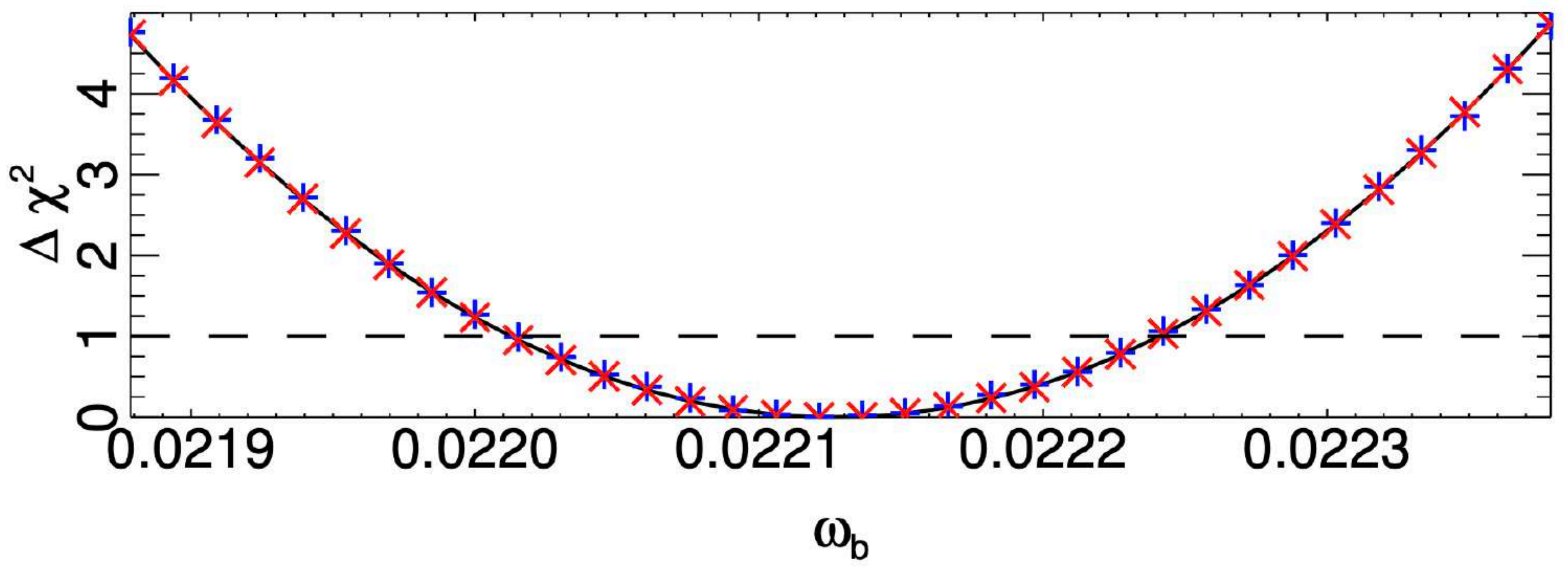
Motivation:

Frequentist method for comparison with Bayesian to check for prior or marginalization effects

What is the profile likelihood $\chi^2(\theta)$?

- Fix the parameters of interested θ to different values
- Maximize the likelihood L (minimizes $\chi^2 = -2\ln L$) wrt to the other parameters for different values of the parameter of interest
- For Gaussian distribution this gives a parabola → fit a parabola

Planck intermediate results XVI, 2014



Confidence interval:

- A confidence region, can be extracted from the likelihood ratio statistic: $\Delta\chi^2(\theta) = \chi^2(\theta) - \chi^2_{\rm min} = -2\ln(\mathscr{L}/\mathscr{L}_{\rm max})$
- For parabolic $\chi^2(\theta)$, and one dof, the c.i. is given by:

  $\Delta\chi^2 = 1, 2.7, 3.84$ for 68, 90 and 95%, respectively $\longrightarrow$ Neyman construction

$\chi^2_{\rm min}$ is obtained from global maximum likelihood estimate given the entire set of parameters

Profile likelihood

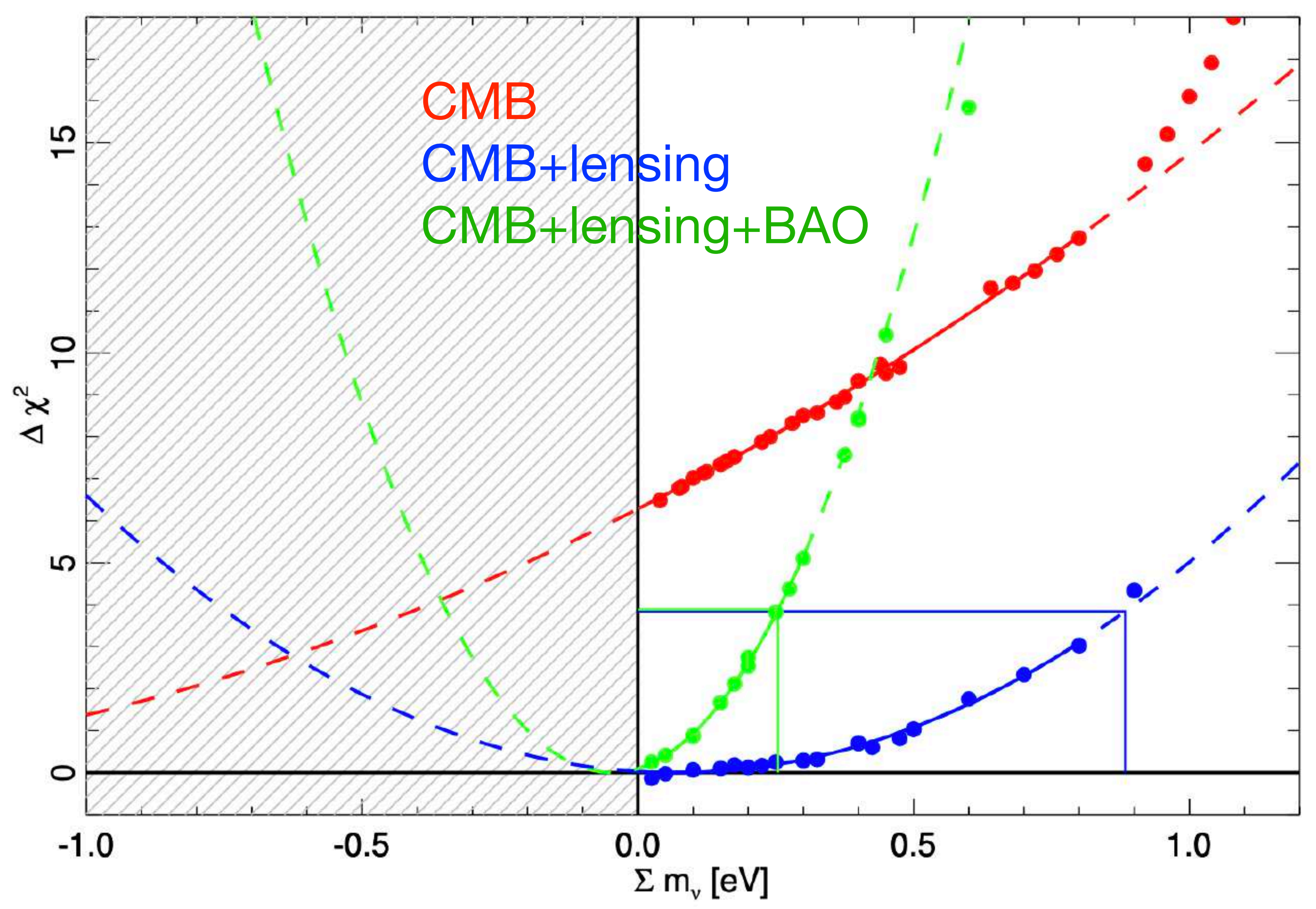
Advantages:

- Not affected by volume effects (no priors)
- Invariant under reparametrizations (since $L(\theta)$ is a MLE)
- Allows construction of confidence intervals close to boundaries

Disadvantages:

- Computationally expensive

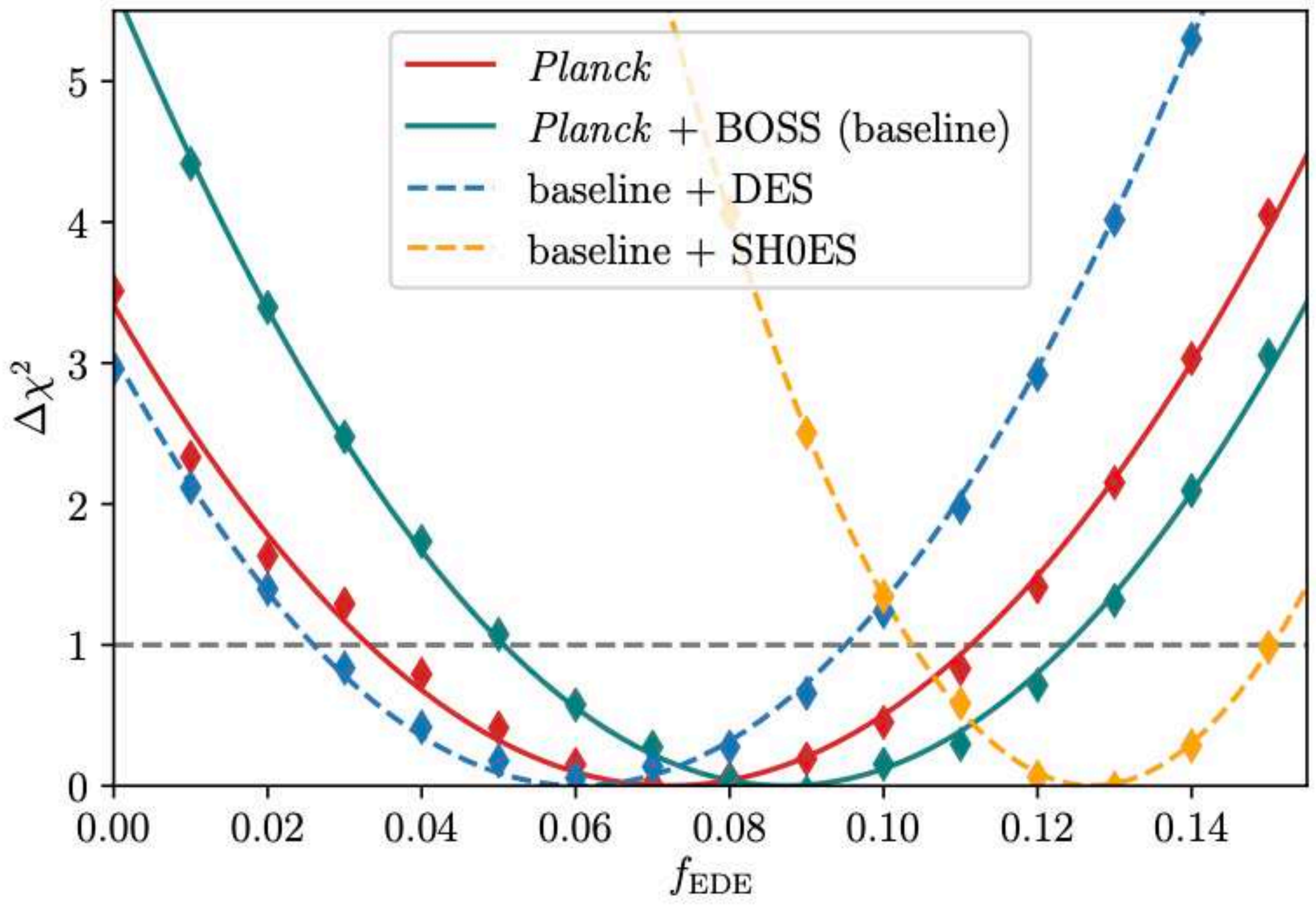
"Larger confidence intervals"



Planck intermediate results XVI, 2014

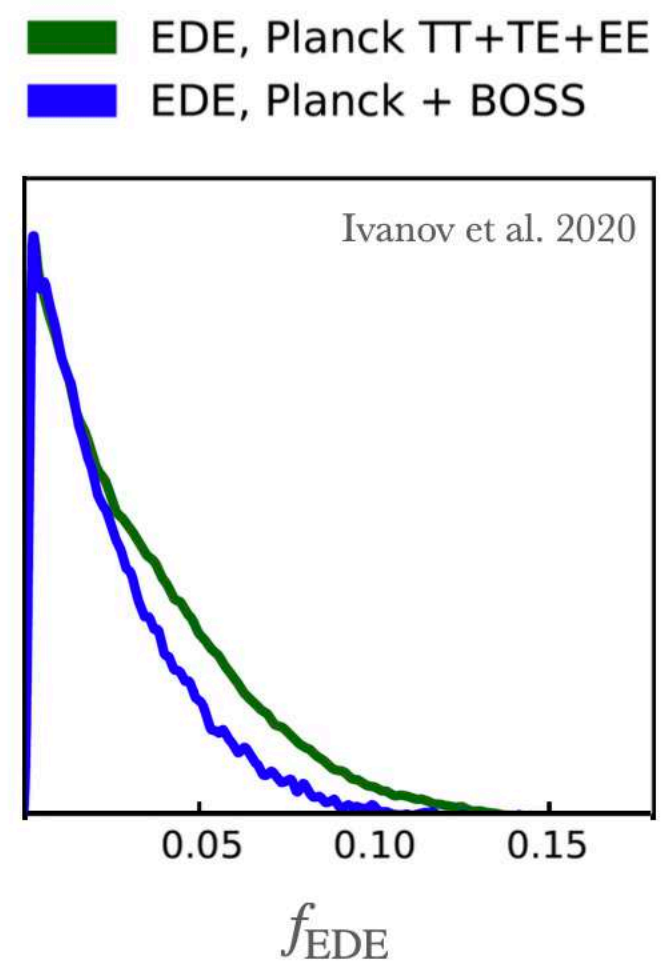
Profile likelihood for $f_{\rm EDE}$

Check for *volume effects:*
comparison to previous results



$$f_{\rm EDE}^{\rm (base)} = 0.087 \pm 0.037$$
(68%CL)

Baseline: Planck+BOSS FS

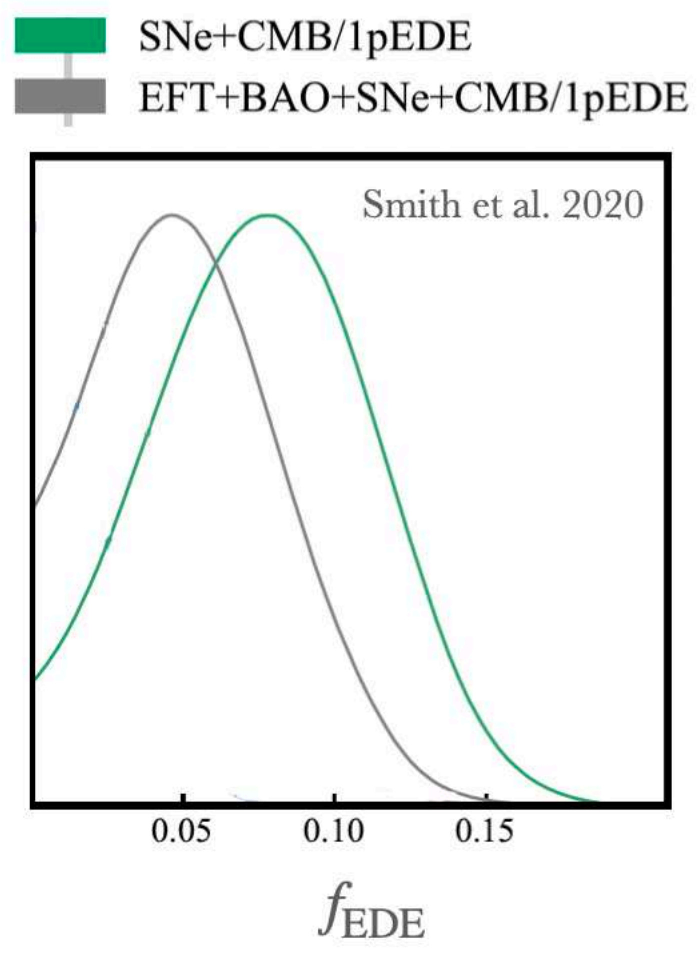


3-parameter model

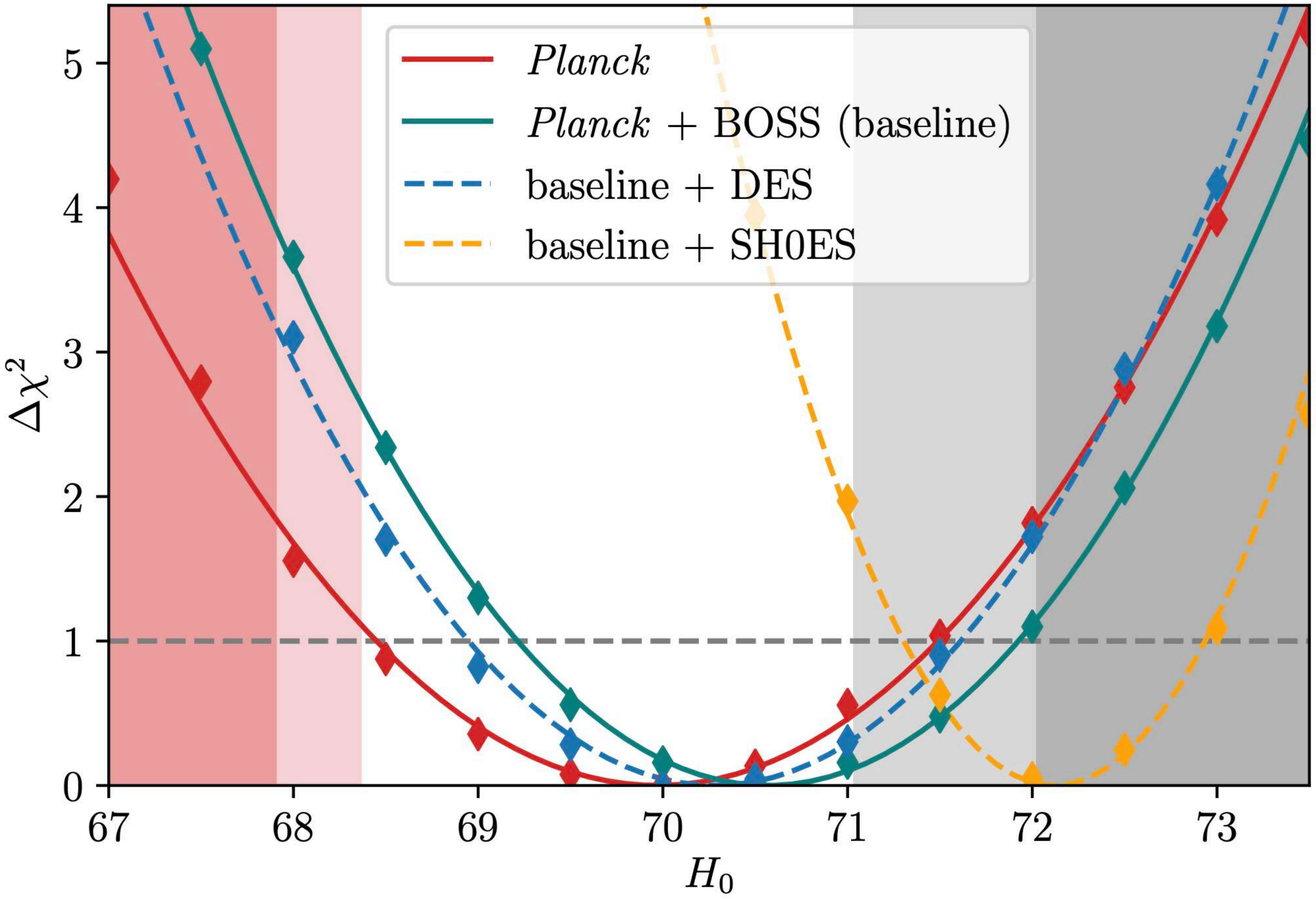
$$f_{\rm EDE} < 0.072 \quad (95\% CL)$$

1-parameter model

$$f_{\rm EDE} = 0.072 \pm 0.034$$
$(68\% CL)$



Profile likelihood for H0

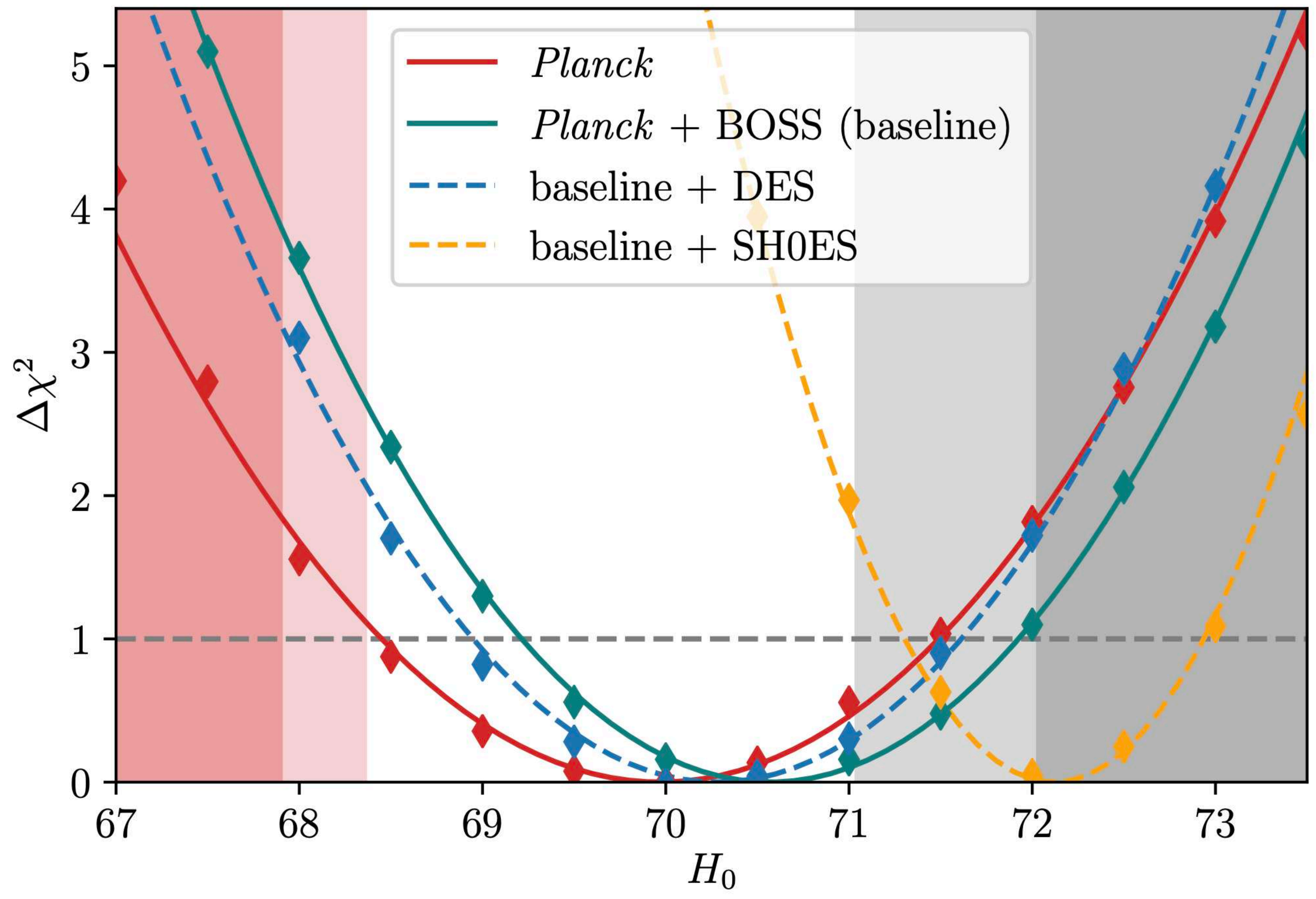


$$H_0^{\mathrm{EDE,(base)}} = 70.57 \pm 1.36$$
(68%CL)

Consistent with SHOES at 1.4σ !

Baseline: Planck+BOSS FS

Profile likelihood for H0



Data set	$\chi^2(\Lambda\text{CDM})$	$\chi^2(\text{EDE})$	$\Delta\chi^2$	f_{EDE}	H_0 (consistency w. SH0ES)
Planck	2774.24	2770.72	-3.52	0.072 ± 0.039	69.97 ± 1.52 ($1.7\,\sigma$)
Planck+BOSS (base)	3045.65	3039.98	-5.67	0.087 ± 0.037	70.57 ± 1.36 ($1.4\,\sigma$)
Baseline + DES	3052.06	3049.13	-2.93	$0.061^{+0.035}_{-0.034}$	70.28 ± 1.33 ($1.6\,\sigma$)
Baseline + SH0ES	3068.44	3042.08	-26.36	0.127 ± 0.023	72.12 ± 0.82 ($0.69\,\sigma$)

Consistent with SHOES at $< 1.4\sigma$!

No H0 tension with EDE
(For this dataset)

Baseline: Planck+BOSS FS

Profile likelihood: general

- Using profile likelihood and other complimentary statistical methods is going to be more and more important in cosmology:

    - Beyond LCDM model

      Ex.: EDE, decaying DM, nEDE, …

    - New observables or systematics

      Ex.: full-shape (nuisance parameters), clusters, …

→ **extra dofs**

Important new analysis to be incorporated in observational collaborations!

Ex.: Euclid
PFS and LiteBIRD?

Refs:

- Herold + Ferreira, 2022; + Komatsu, 2021
- Campetti, Komatsu, 2022
- Cosmology with 6 parameters in the Stage-IV era: efficient marginalisation over nuisance parameters, Hadzhiyska et al, 2023
- Other works to come…

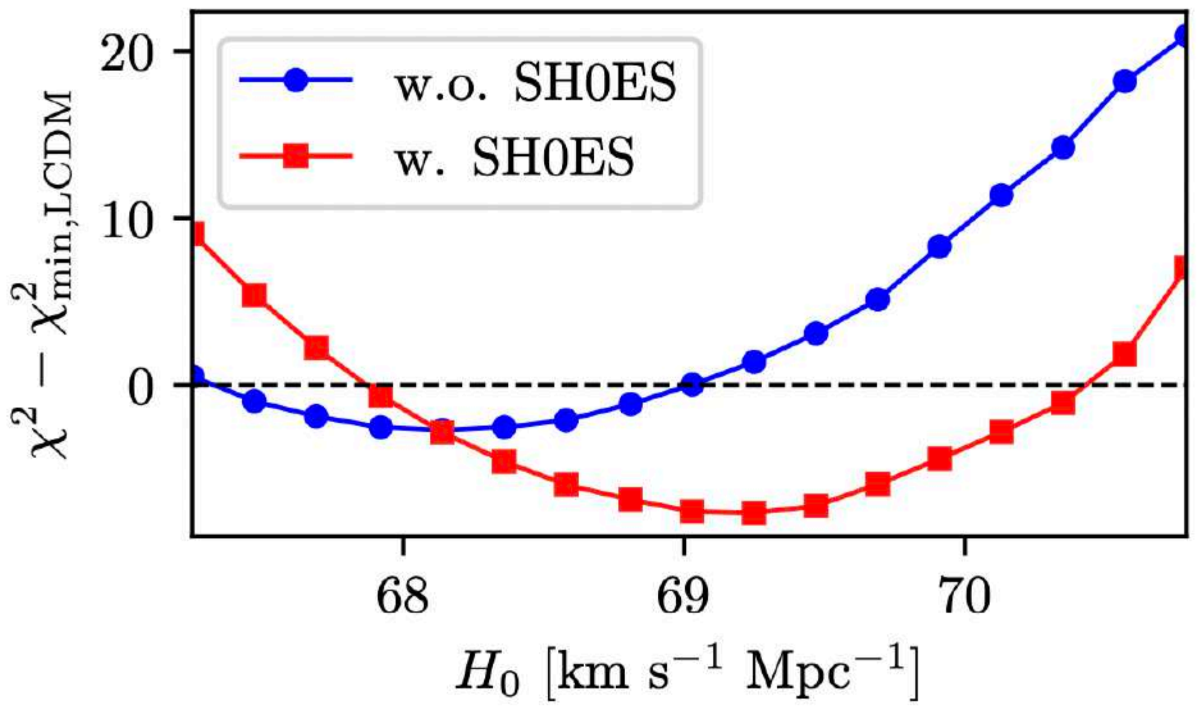
Other beyond LCDM models…

- Decaying dark matter

Holm et al. 2022

(See also (Holm et al. 2022 - DWDM)

- Previous MCMC analysis find a strong preference for either very long-lived or very short-lived dark matter.



This work:

- This preference is due to volume effects - drives the model towards the standard ΛCDM limit

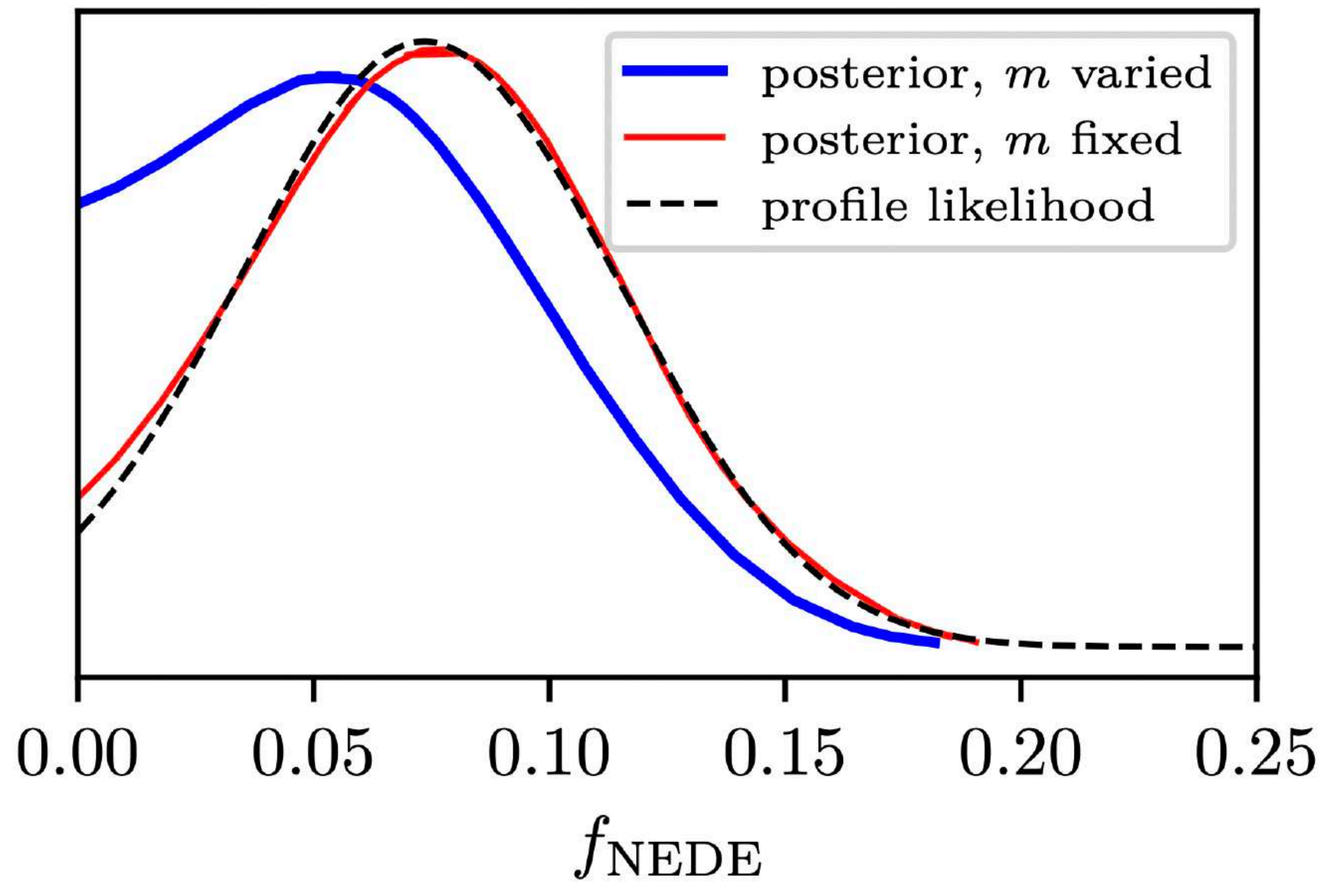
- Using profile likelihoods, they instead find that best-fitting parameters are in an intermediate regime - ~3% of cold dark matter decays just prior to recombination.

- New EDE

Cruz et al. 2023

$$\{f_{\mathrm{NEDE}}, z_{\mathrm{decay}}, w_{\mathrm{NEDE}}\}$$

- Stronger evidence for NEDE with PL, than MCMC



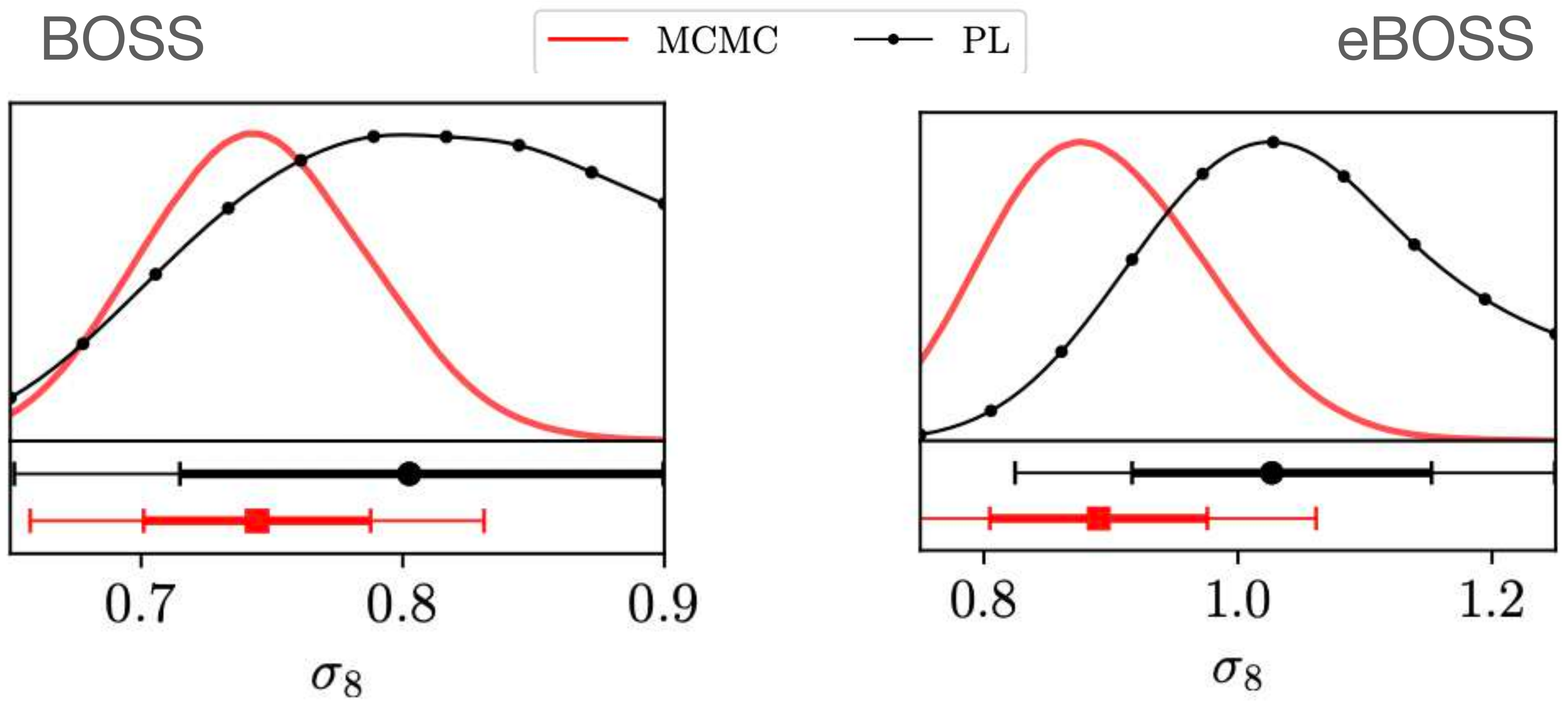
- Shows that fixing the trigger mass an s an appropriate method of avoiding volume effects

Prior effects in EFTofLSS analyses of full-shape BOSS and eBOSS data

E.Holm, L. Herold, T. Simon, EF, S. Hannestad, V. Poulin, T. Tram 2023

- Previous MCMC analysis have shown that the constraints from BOSS full-shape data using EFTofLSS depend on the choice of prior on the EFT nuisance parameters

- We explore this prior dependence using profile likelihood for BOSS, eBOSS and Planck



Shift of 1σ (BOSS) 1.6σ (eBOSS)

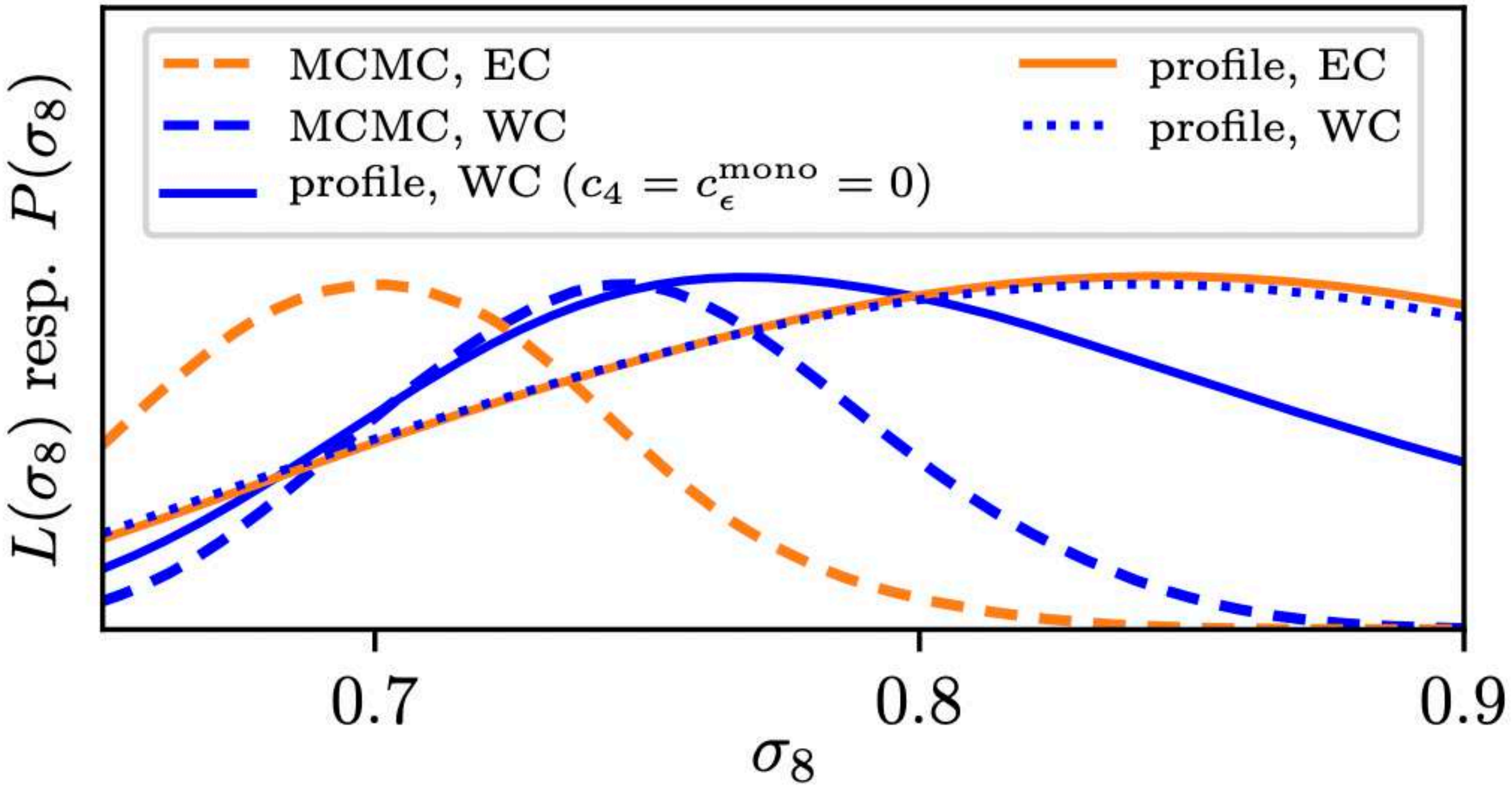
Widening of intervals by factors of 1.9 (BOSS) 1.6 (eBOSS)

- We find that the priors on the EFT parameters in the Bayesian inference are informative and that prior volume effects are important.

Prior effects in *EFTofLSS analyses of full-shape BOSS and eBOSS data*

E.Holm, L. Herold, T. Simon, EF, S. Hannestad, V. Poulin, T. Tram 2023

- EC vs. WC parametrizations and comparison to MCMC

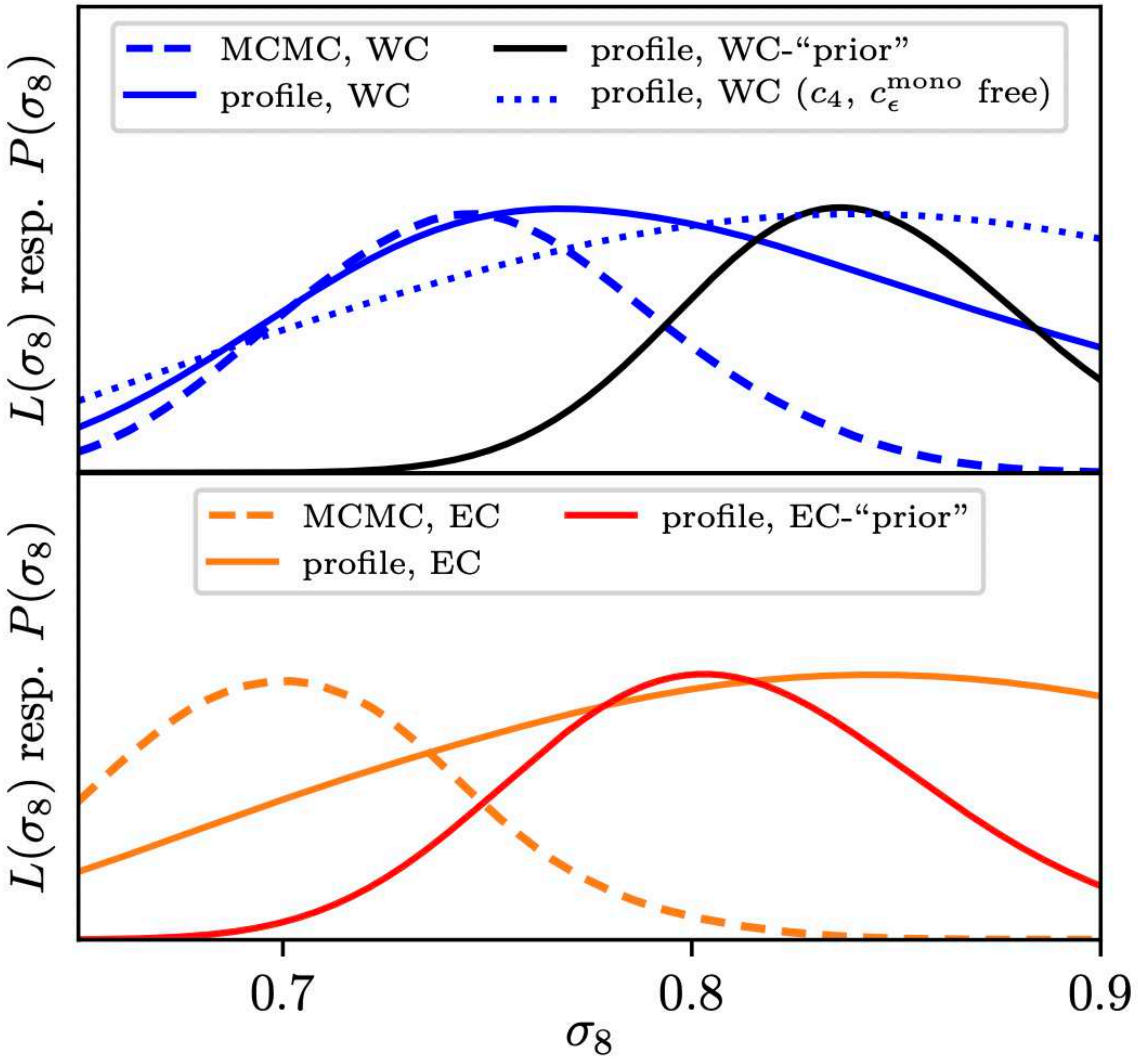


Marginalized MCMC posteriors (dashed) and profile likelihoods (solid) of σ8 within the WC (blue) and EC parametrizations (orange), for BOSS+BAO data

Prior effects in EFTofLSS analyses of full-shape BOSS and eBOSS data

E.Holm, L. Herold, T. Simon, EF, S. Hannestad, V. Poulin, T. Tram 2023

- EC vs. WC parametrizations and comparison to MCMC



Marginalized MCMC posteriors (dashed) and profile likelihoods (solid) of σ8 within the WC (blue) and EC parametrizations (orange), for BOSS+BAO data, *including profile likelihoods with Gaussian data likelihoods on the EFT parameters*

Other σ_8 estimates

- DES x KIDS analysis

DES Y3 + KIDS-1000: CONSISTENT COSMOLOGY COMBINING COSMIC SHEAR SURVEYS

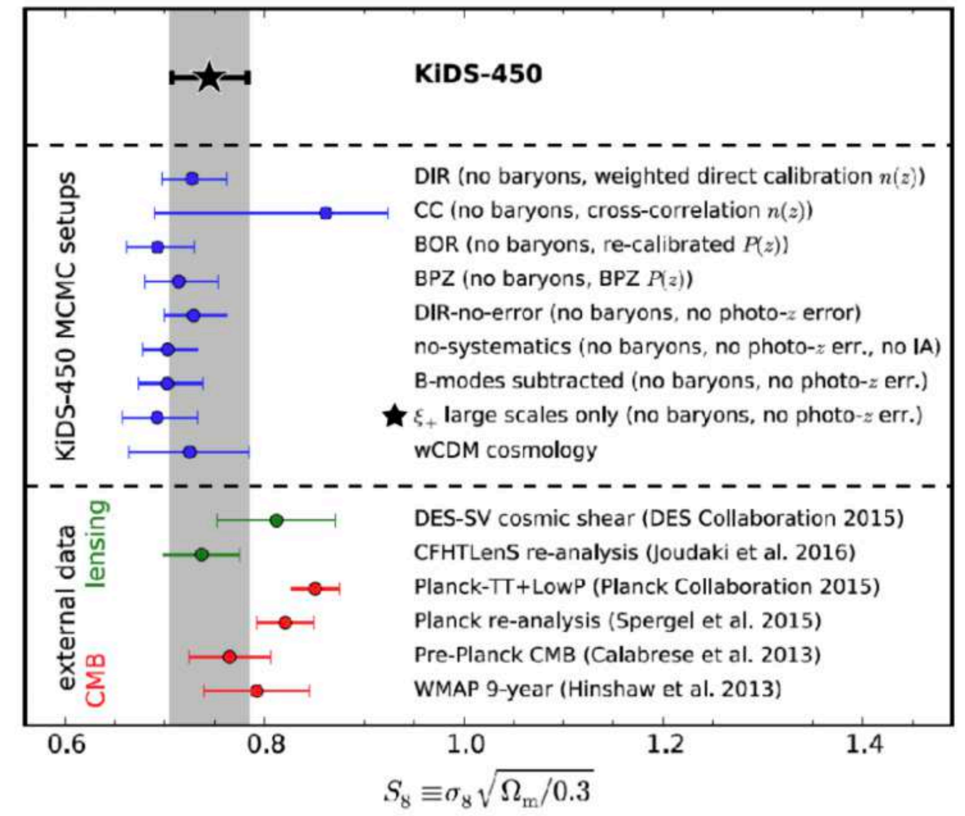
DARK ENERGY SURVEY AND KILO-DEGREE SURVEY COLLABORATION:
T. M. C. ABBOTT,[1] M. AGUENA,[2] A. ALARCON,[3] O. ALVES,[4] A. AMON,[5,6] F. ANDRADE-OLIVEIRA,[4] M. ASGARI,[7] S. AVILA,[8] D. BACON,[9] K. BECHTOL,[10] M. R. BECKER,[3] G. M. BERNSTEIN,[11] E. BERTIN,[12,13] M. BILICKI,[14] J. BLAZEK,[15] S. BOCQUET,[16] D. BROOKS,[17] P. BURGER,[18] D. L. BURKE,[19,20] H. CAMACHO,[21,2] A. CAMPOS,[22] A. CARNERO ROSELL,[23,2,24] M. CARRASCO KIND,[25,26] J. CARRETERO,[8] F. J. CASTANDER,[27,28] R. CAWTHON,[29] C. CHANG,[30,31] R. CHEN,[32] A. CHOI,[33] C. CONSELICE,[34,35] J. CORDERO,[34] L. N. DA COSTA,[2] M. E. S. PEREIRA,[36] R. DALAL,[37] C. DAVIS,[19] J. T. A. DE JONG,[38,39] J. DEROSE,[40] S. DESAI,[41] H. T. DIEHL,[42] S. DODELSON,[22,43] P. DOEL,[17] C. DOUX,[11,44] A. DRLICA-WAGNER,[30,42,31] A. DVORNIK,[45] K. ECKERT,[11] T. F. EIFLER,[46,47] J. ELVIN-POOLE,[48] S. EVERETT,[47] X. FANG,[49,46] I. FERRERO,[50] A. FERTÉ,[20] B. FLAUGHER,[42] O. FRIEDRICH,[6] J. FRIEMAN,[42,31] J. GARCÍA-BELLIDO,[51] M. GATTI,[11] G. GIANNINI,[8] B. GIBLIN,[52,53] D. GRUEN,[16] R. A. GRUENDL,[25,26] G. GUTIERREZ,[42] I. HARRISON,[54] W. G. HARTLEY,[55] K. HERNER,[42] C. HEYMANS,[53,45] H. HILDEBRANDT,[45] S. R. HINTON,[56] H. HOEKSTRA,[38] D. L. HOLLOWOOD,[57] K. HONSCHEID,[58,59] H. HUANG,[46,60] E. M. HUFF,[47] D. HUTERER,[4] D. J. JAMES,[61] M. JARVIS,[11] N. JEFFREY,[17] T. JELTEMA,[57] B. JOACHIMI,[62] S. JOUDAKI,[63] A. KANNAWADI,[64] E. KRAUSE,[46] K. KUEHN,[65,66] K. KUIJKEN,[38] N. KUROPATKIN,[42] P.-F. LEGET,[19] P. LEMOS,[67,68,69,70] S. LI,[38] X. LI,[71,72] A. R. LIDDLE,[73] M. LIMA,[74,2] C. LIN,[45,75] H. LIN,[42] N. MACCRANN,[76] C. MAHONY,[45] J. L. MARSHALL,[77] J. MCCULLOUGH,[19] J. MENA-FERNÁNDEZ,[78] F. MENANTEAU,[25,26] R. MIQUEL,[79,8] J. J. MOHR,[80,16] J. MUIR,[81] J. MYLES,[82,19,20] N. NAPOLITANO,[83] A. NAVARRO-ALSINA,[84] R. L. C. OGANDO,[85] A. PALMESE,[22] S. PANDEY,[11] Y. PARK,[86] M. PATERNO,[42] J. A. PEACOCK,[53] D. PETRAVICK,[25] A. PIERES,[2,85] A. A. PLAZAS MALAGÓN,[87] A. PORREDON,[58,59,53] J. PRAT,[30,31] M. RADOVICH,[88] M. RAVERI,[89] R. REISCHKE,[45] R. P. ROLLINS,[34] A. K. ROMER,[90] A. ROODMAN,[19,20] E. S. RYKOFF,[19,20] S. SAMUROFF,[15] C. SÁNCHEZ,[11] E. SANCHEZ,[78] J. SANCHEZ,[91] P. SCHNEIDER,[18] L. F. SECCO,[31] I. SEVILLA-NOARBE,[78] H. SHAN,[92,93,94] E. SHELDON,[95] T. SHIN,[96] C. SIFÓN,[97] M. SMITH,[98] M. SOARES-SANTOS,[4] B. STÖLZNER,[45] E. SUCHYTA,[99] M. E. C. SWANSON, G. TARLE,[4] D. THOMAS,[9] C. TO,[58] M. A. TROXEL,[32] T. TRÖSTER,[100]

May 2023

Identified and quantified the marginalization (prior volume or projection) effects given some choices in the analysis.
Using MAP (maximum a posteriori)

Prior volume effects might be biasing σ_8 from cosmic shear

Analysis in preparation!



Projection effects

With the ***increase statistical power of new observations*** that have the goal to lead to a more precise determination of the cosmological parameters, **new systematic effects** are present leading to a larger number of nuisance parameter that also need to be fitted in the data analysis - systematics more prominent!

This **inflation of the number of parameters** can lead to difficulties in the statistical analysis and ***bias*** the inference of cosmological parameters.

For the standard MCMC analysis these are:

Ex.: Prior volume effects, weight volume effects, … ⟶ Inescapable feature of the Bayesian method!

Relevant to study the extent to which one's results are affected by volume effects!

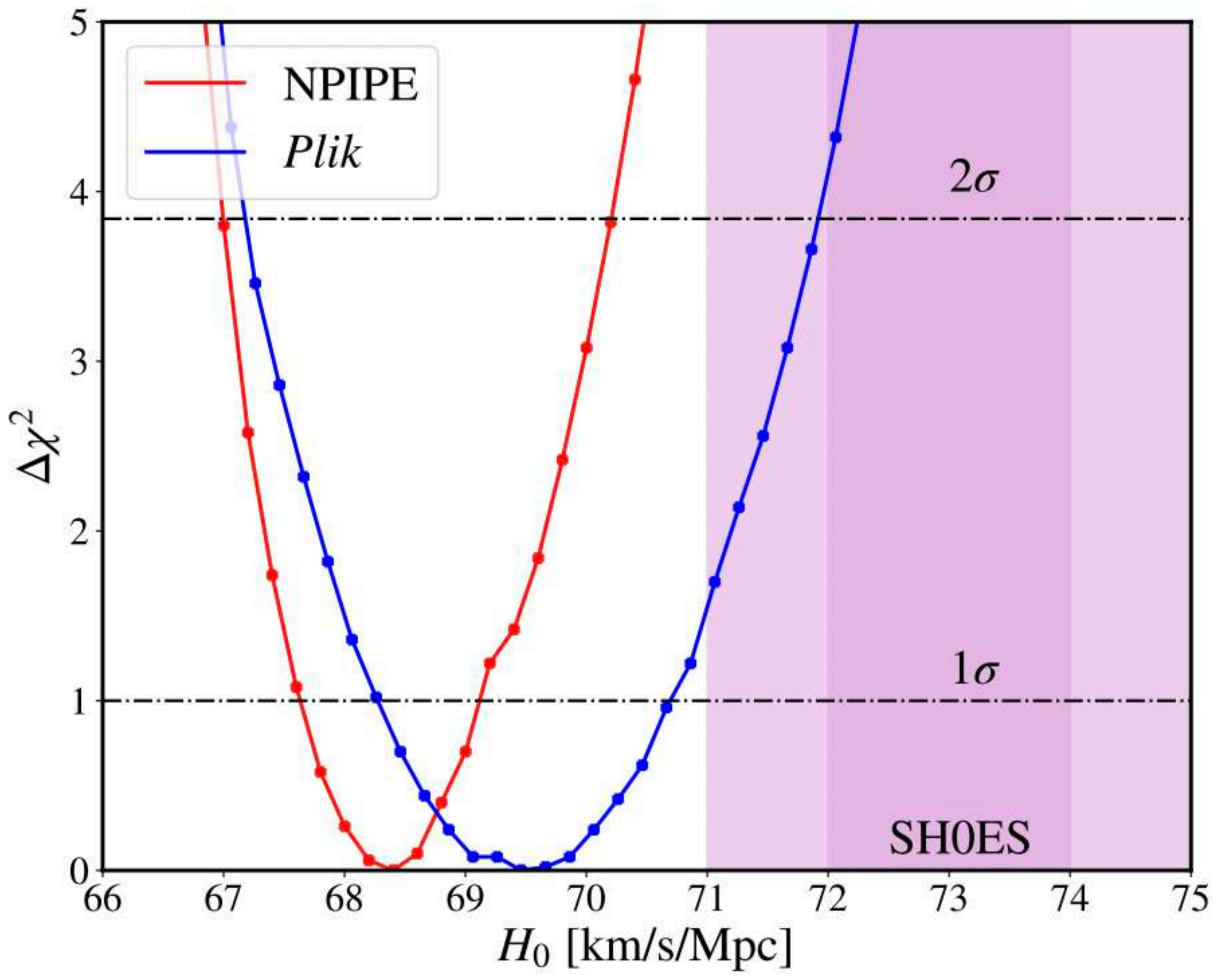
Need complementary statistical methods to deal with that, like for example, the profile likelihood

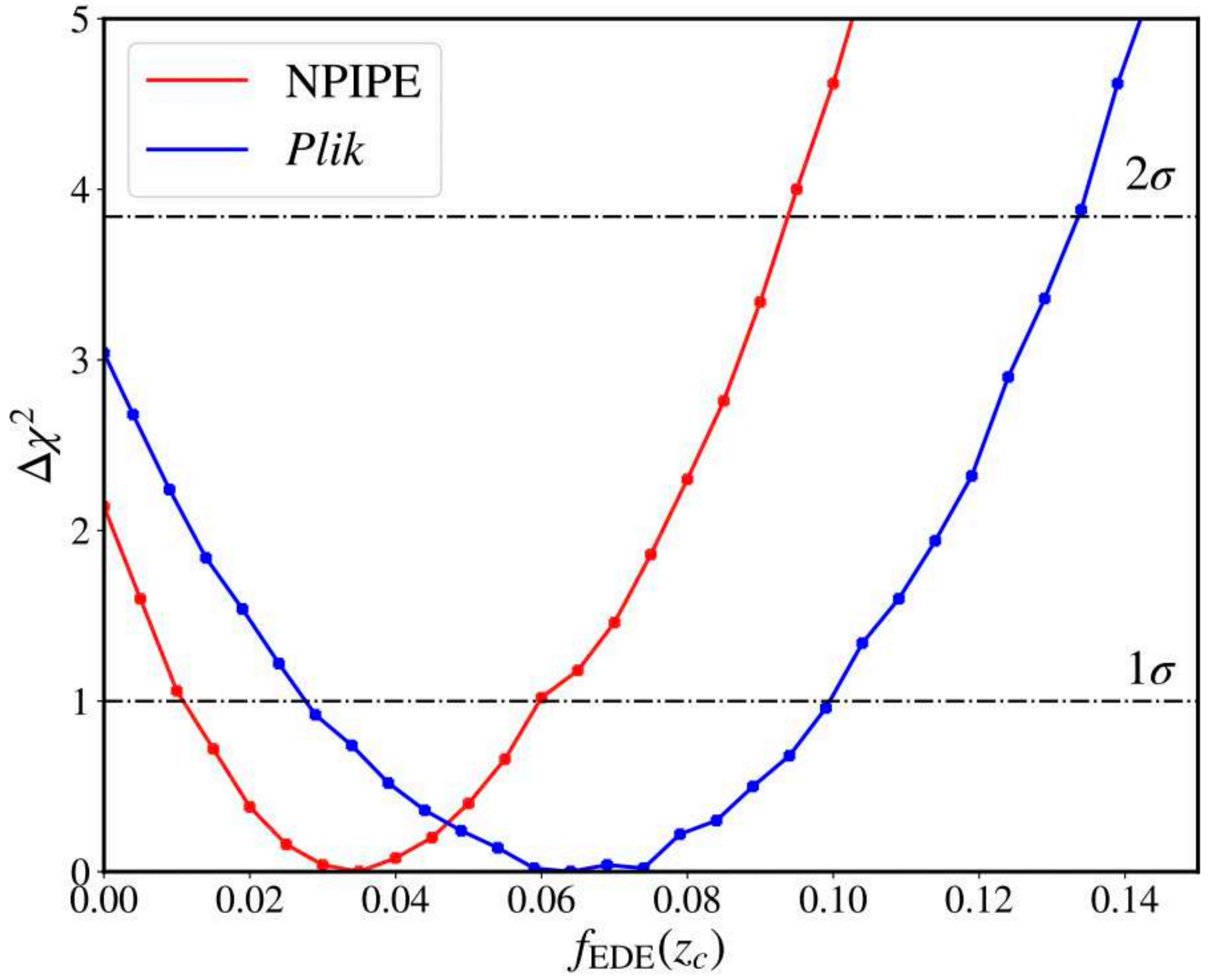
Early dark energy - reanalysis

George Efstathiou, Erik Rosenberg,, and Vivian Poulin 2023, *Improved Planck constraints on axion-like early dark energy as a resolution of the Hubble tension*

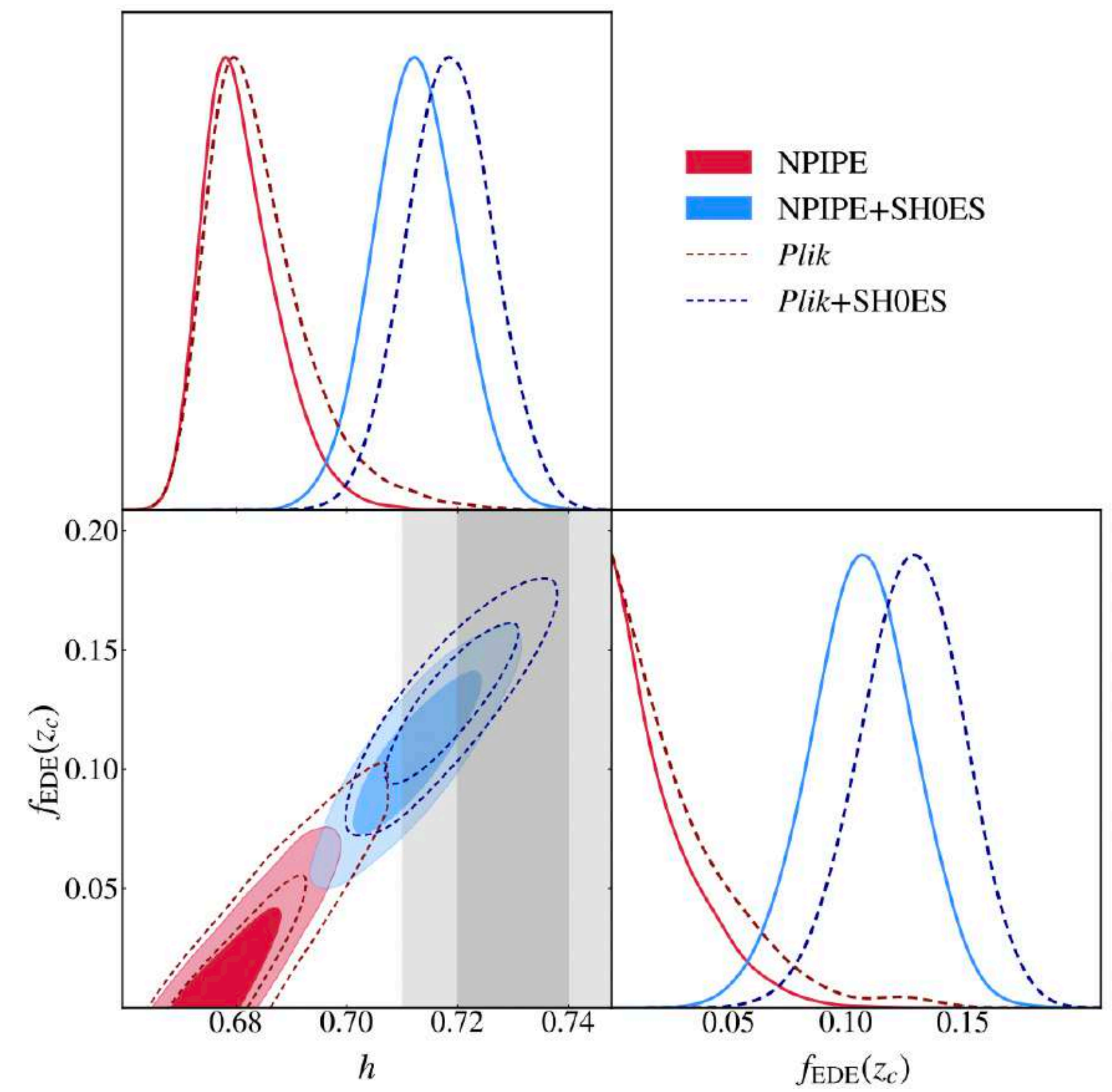
Reanalysis using new *Planck* likelihood (NPIPE)

Profile likelihood:





MCMC:



PL (MCMC) analysis: $H_0^{\mathrm{EDE,(NPIPE)}} = 68.37 \pm 0.0075\,(68.11^{+0.0047}_{-0.0082})$

Distance from SHOES at 3.7σ

Volume effects present in the old likelihood.

NO volume effects with new likelihood!

Prior/marginalization effects

Inflation in the number of nuisance parameters or beyond LCDM parameters that enter the statistical analysis leading to possible marginalization or prior volume effects in standard MCMC analysis

Inherent from Bayesian analysis. When strong, can influence inferred parameters

Early Dark Energy

Volume effects are important: full MCMC result differs from the profile likelihood

PL analysis: $H_0^{\mathrm{EDE,(Plik)}} = 70.57 \pm 1.36$
Distance from SHOES at 1.4σ

PL (MCMC) analysis: $H_0^{\mathrm{EDE,(NPIPE)}} = 68.37 \pm 0.0075\,(68.11^{+0.0047}_{-0.0082})$
Distance from SHOES at 3.7σ

Summary

Complementary statistical methods necessary for current and future parameter inference analysis - profile likelihood

~~Bayesian vs Frequentist~~ Bayesian & Frequentist & ML & …

STATISTICAL THINKING

Statistics: Are you Bayesian or Frequentist?

The fastest way to diagnose your statistical alignment

Statistical methods are available to us to use. Each statistical method should be used in the appropriate situation.
No right or wrong, no better or worse, no preference.

Not use blindly

Arman Shafieloo: "Many statistical tools not used properly in cosmology"

→ or interpreted

Systematics dominated era

Higher quality data → effects of systematic more important
Particularly serious for Stage IV experiments or cosmological tensions!

Careful to avoid misinterpretations or "fake" new discoveries!


Thank you!