Fast likelihood-free inference in the LSS **Stage-IV** era

Guillermo Franco Abellán

GRAPPA, University of Amsterdam

24th October 2024

COSMO'24 Kyoto

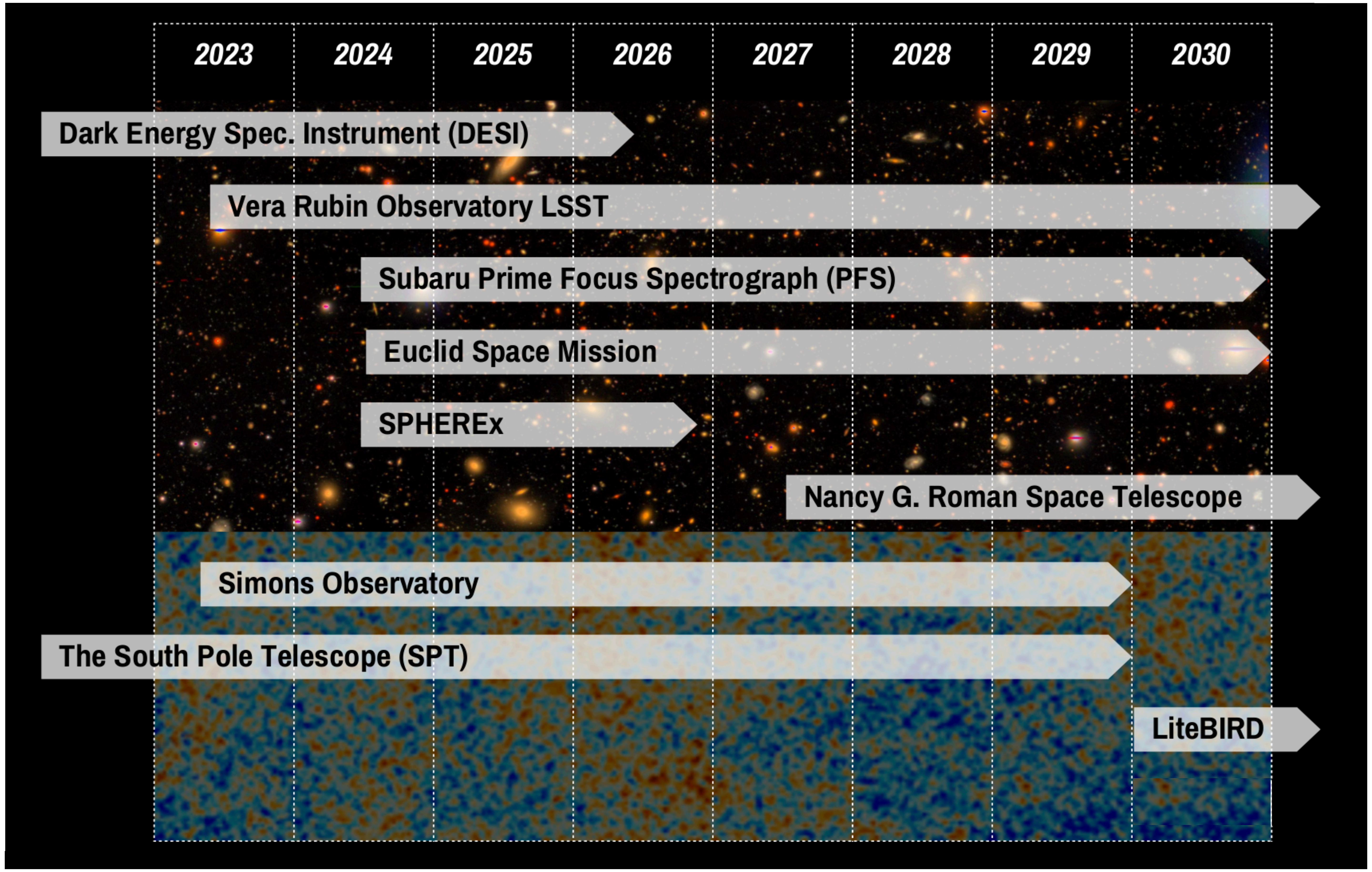
Based on arXiv:2403.14750

with Guadalupe Cañas-Herrera,
Matteo Martinelli,
Oleg Savchenko,
Davide Sciotti,
& Christoph Weniger

GRAPPA

GRavitation AstroParticle Physics Amsterdam

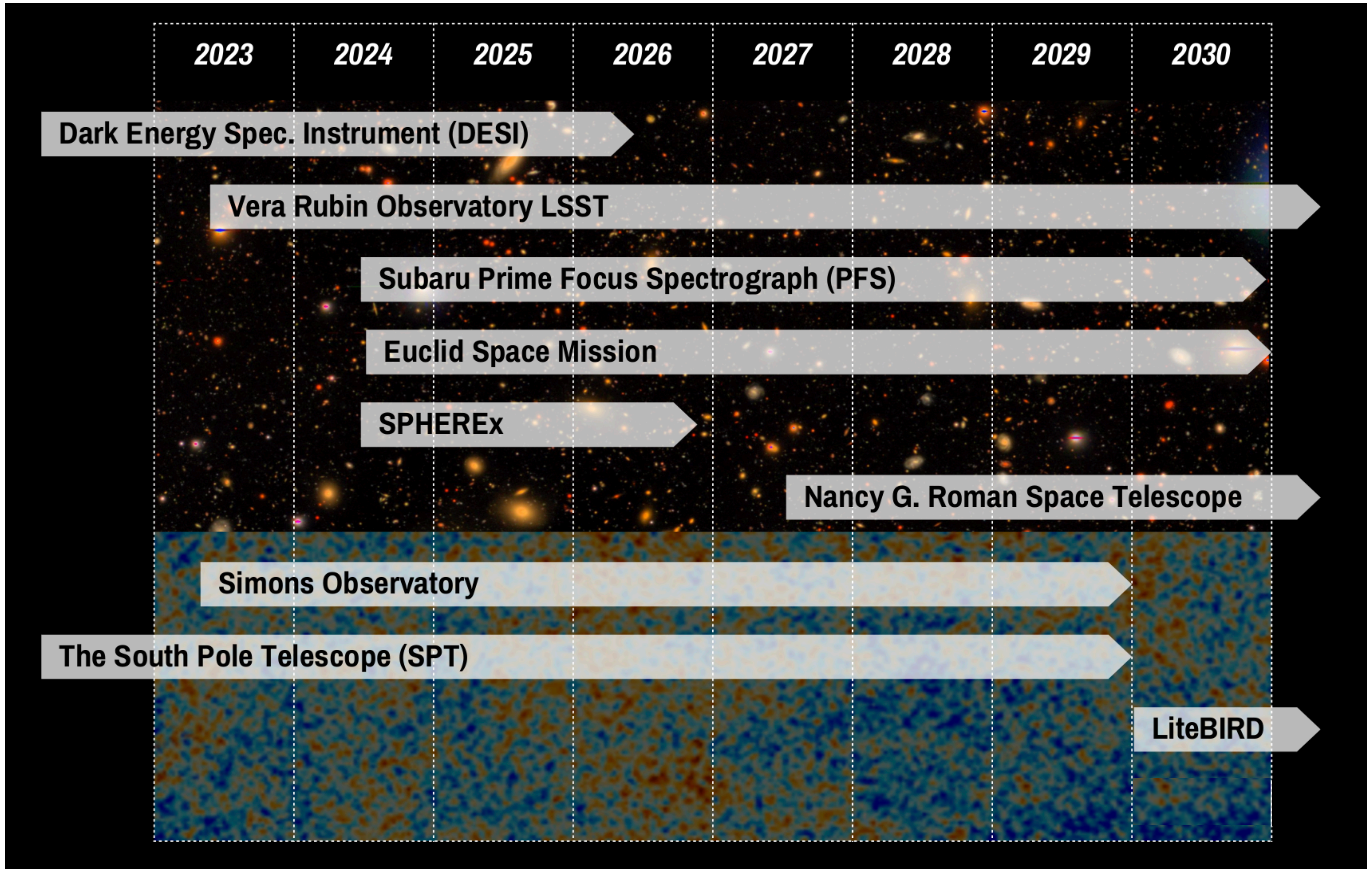
Next-generation cosmological data is becoming available…



Credit: A. Bayer

Next-generation cosmological data is becoming available...

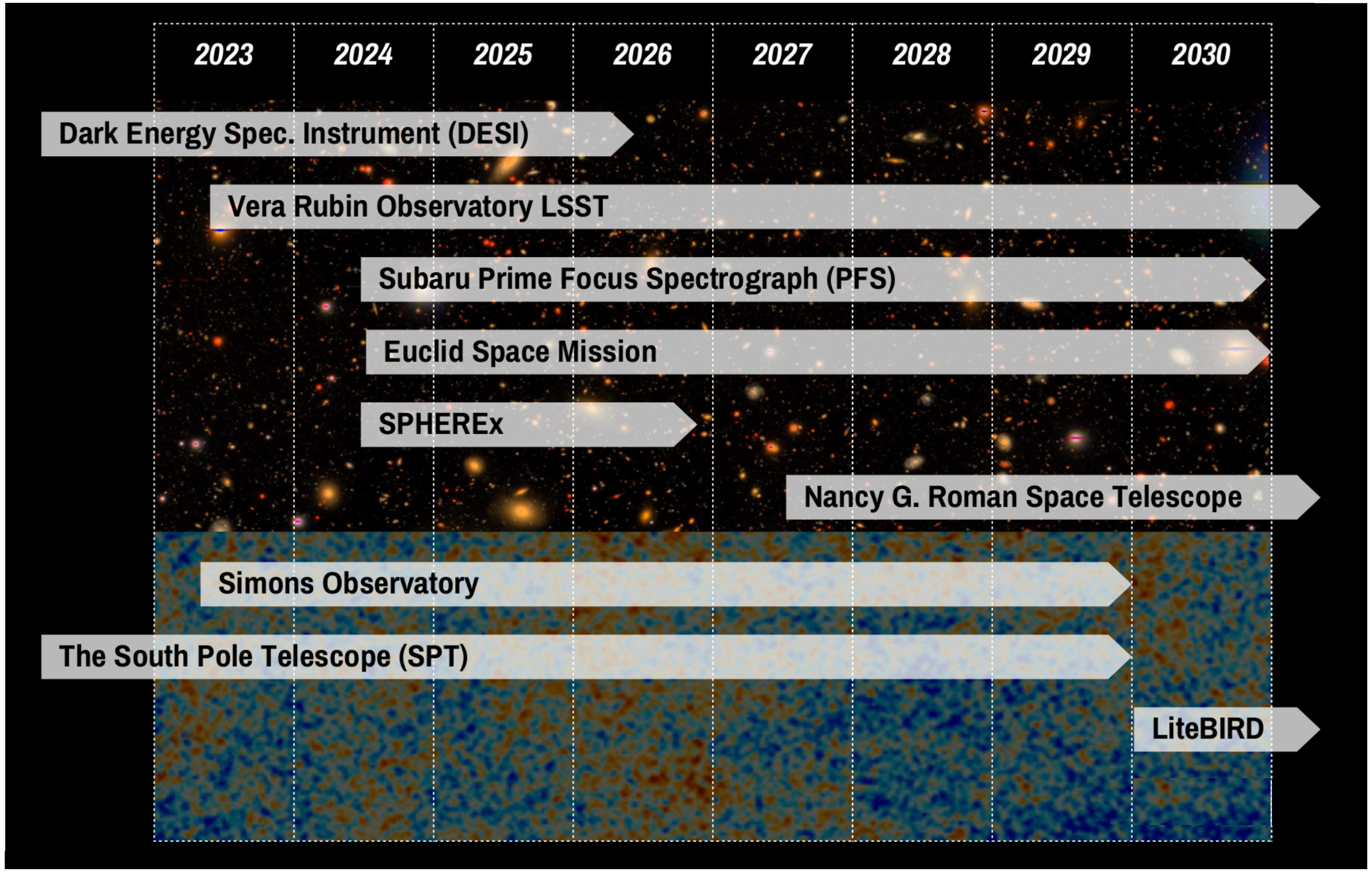
...but analysing these data will be **unfeasible** with traditional methods



Credit: A. Bayer

Next-generation cosmological data is becoming available...

...but analysing these data will be **unfeasible** with traditional methods

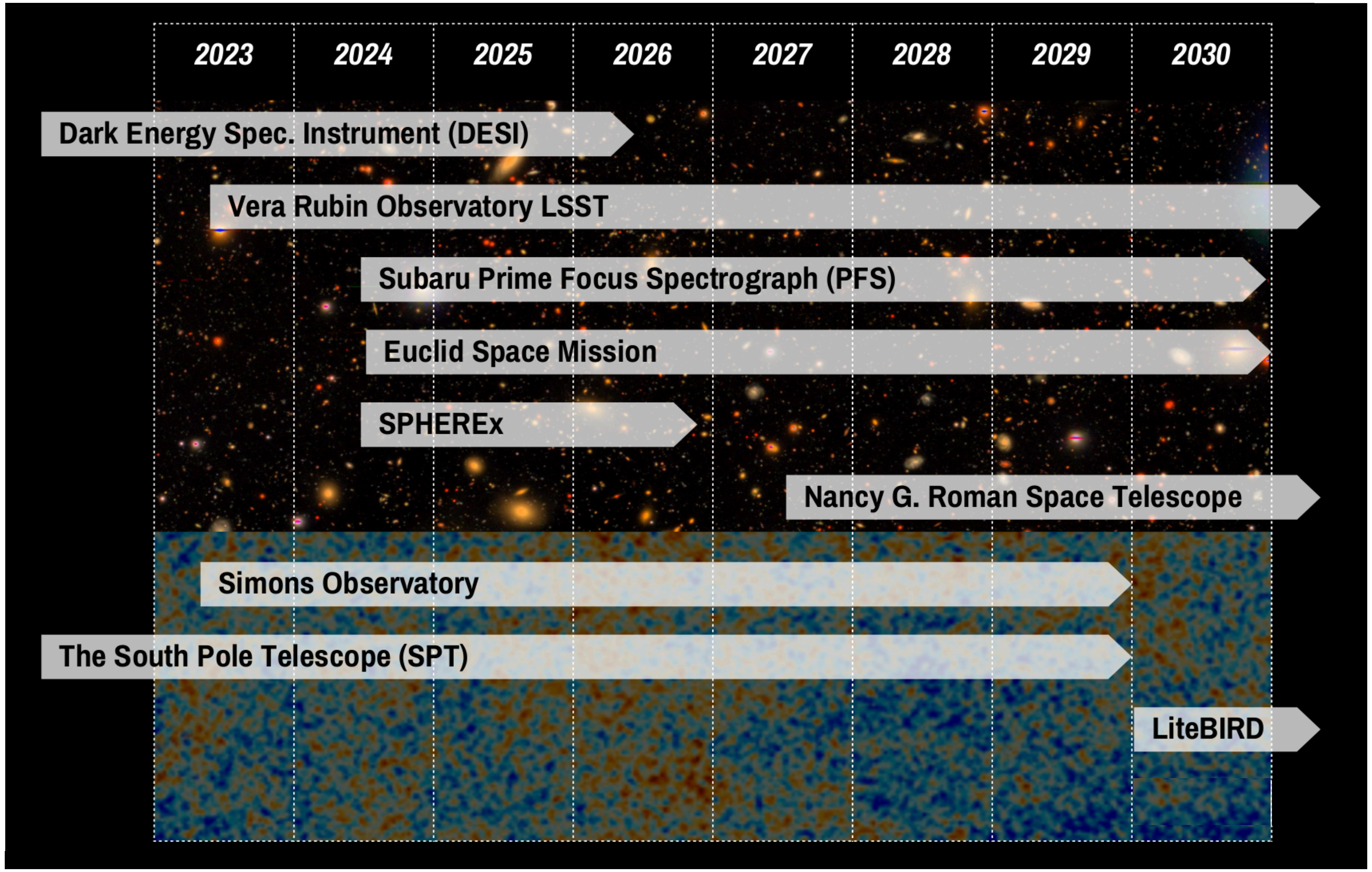


Credit: A. Bayer

Classical methods scale poorly with the dimensionality of the parameter space

Next-generation cosmological data is becoming available…

…but analysing these data will be **unfeasible** with traditional methods

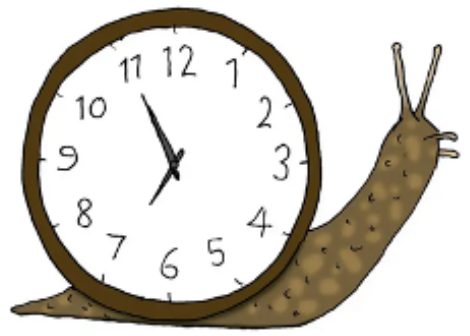


Credit: A. Bayer

Classical methods scale poorly with the dimensionality of the parameter space

Ex: recent **Euclid** forecast included **64 nuisance parameters**

Euclid. I. Overview 24



Our **GOAL**

Accelerate parameter inference
from Stage-IV photometric surveys
using a new **SBI** method called MNRE

MNRE = **M**arginal **N**eural **R**atio **E**stimation

Implemented in Swyft* [Miller+ 20]

Our **GOAL**

Accelerate parameter inference
from Stage-IV photometric surveys
using a new **SBI** method called MNRE

MNRE = **M**arginal **N**eural **R**atio **E**stimation

Implemented in Swyft* [Miller+ 20]

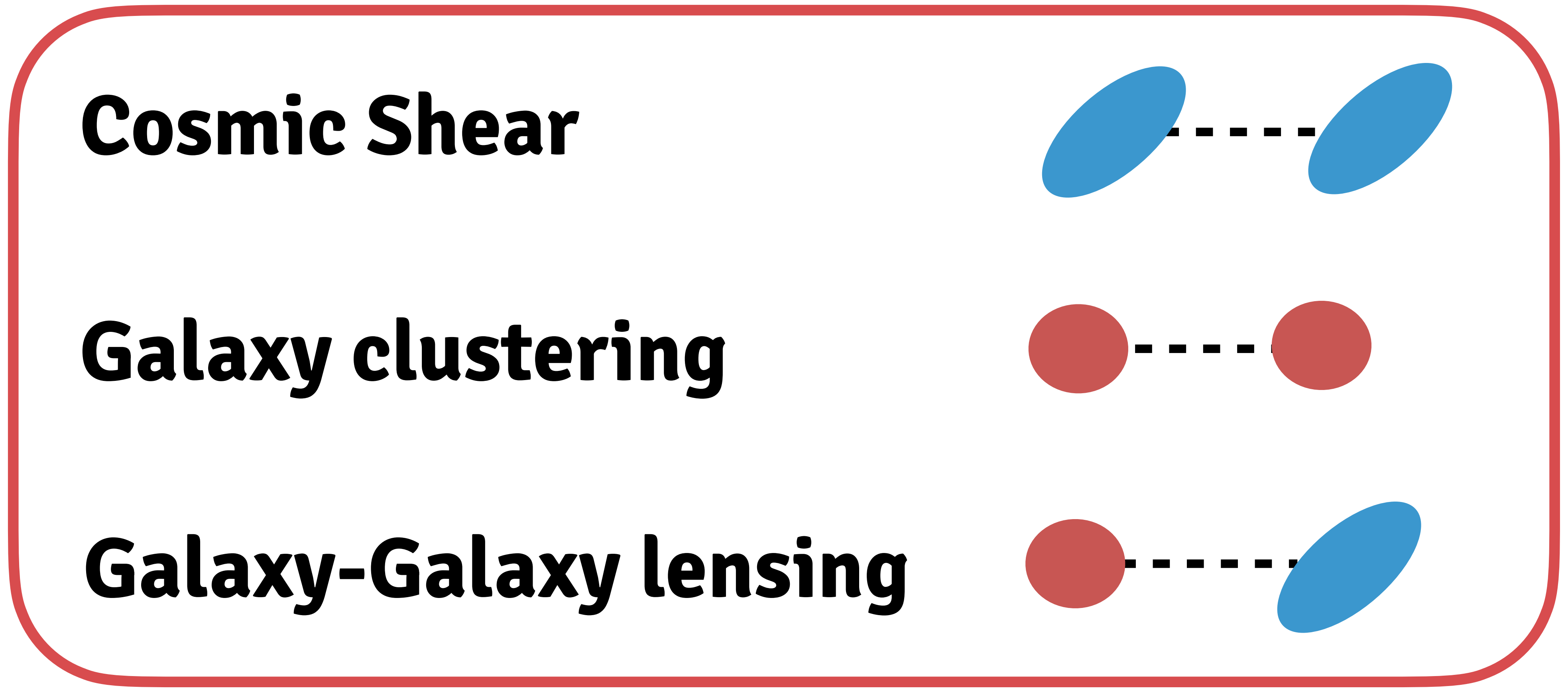
Advantages over other SBI methods:

- More flexibility in network architecture
- It can directly target marginals of interest (more efficient & flexible)

NOTE: We use a simplified simulator/likelihood (10 tomographic bins, 12 nuisance params)

Swyft 3x2pt analysis

3x2pt statistics



NOTE: We use a simplified simulator/likelihood (10 tomographic bins, 12 nuisance params)

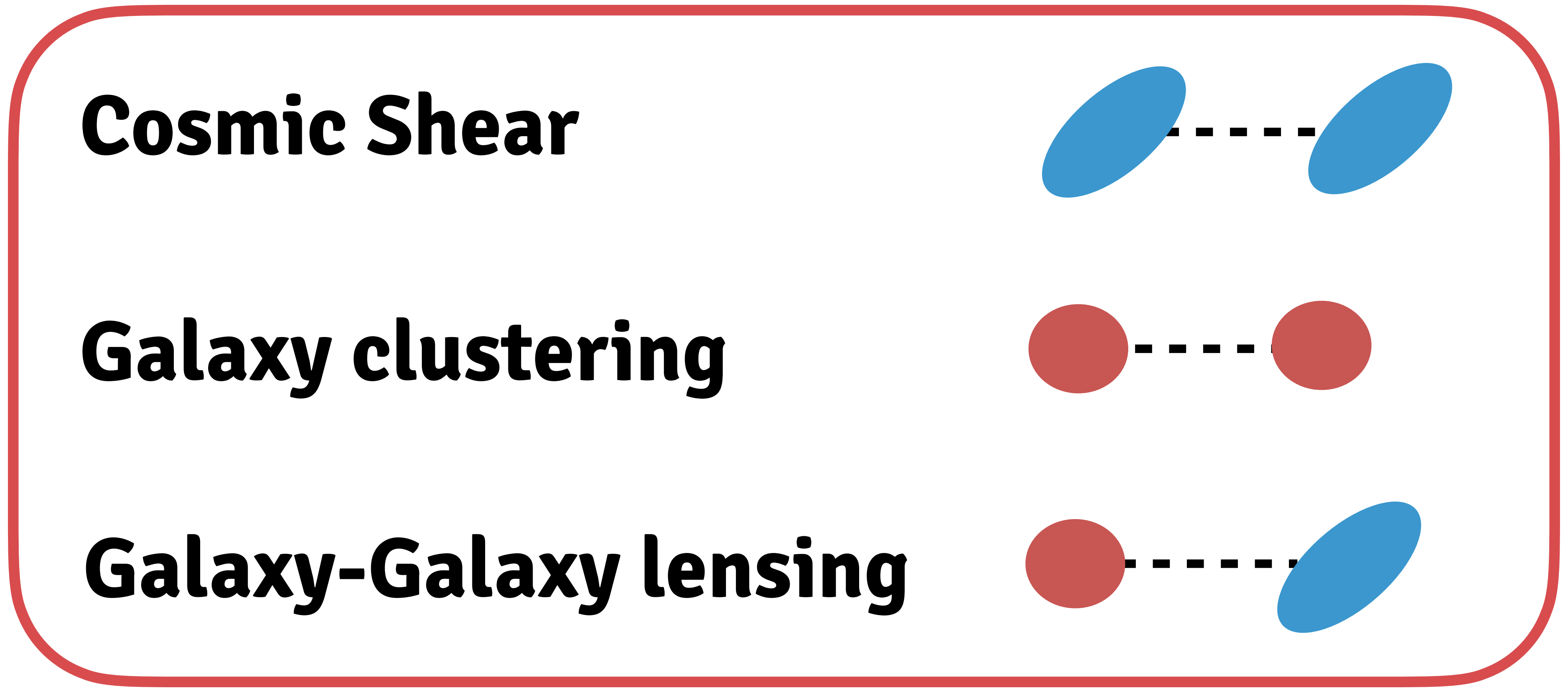
Swyft 3x2pt analysis

- We generate 50k simulations of power spectra with gaussian noise

$$\hat{C}_{ij}^{AB}(\ell) = C_{ij}^{AB}(\ell) + n_{ij}^{AB}(\ell)$$

$$\mathcal{N}(0,\mathbf{C})$$

3x2pt statistics



NOTE: We use a simplified simulator/likelihood (10 tomographic bins, 12 nuisance params)

Swyft 3x2pt analysis

- We generate 50k simulations of power spectra with gaussian noise

$$\hat{C}_{ij}^{AB}(\ell) = C_{ij}^{AB}(\ell) + \underbrace{n_{ij}^{AB}(\ell)}_{\mathcal{N}(0,\mathbf{C})}$$

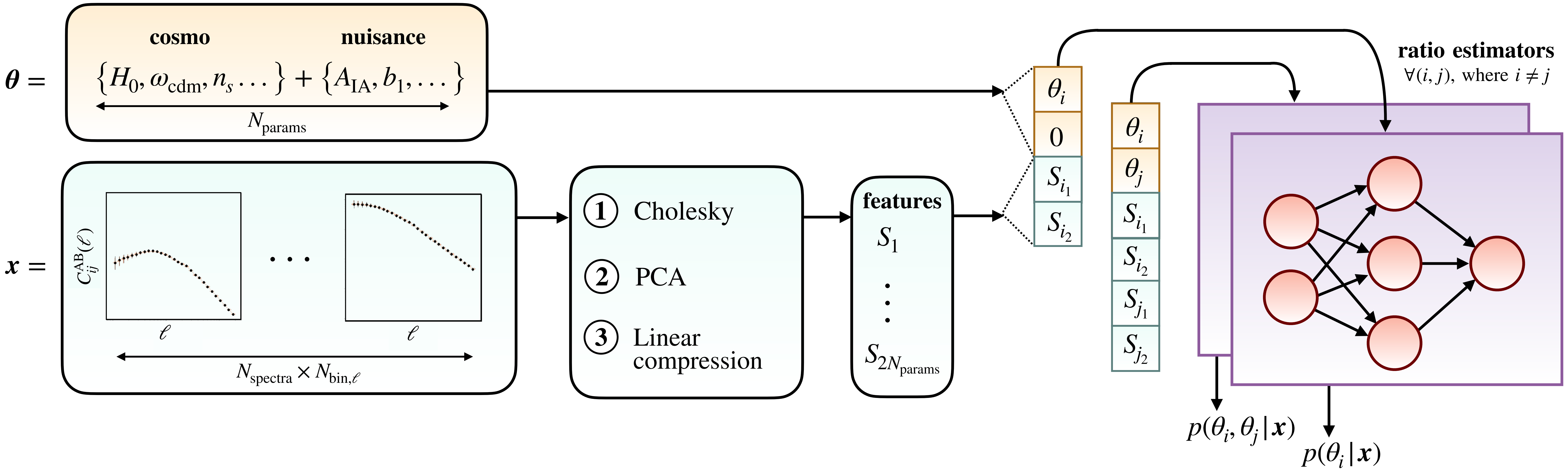
3x2pt statistics

Cosmic Shear

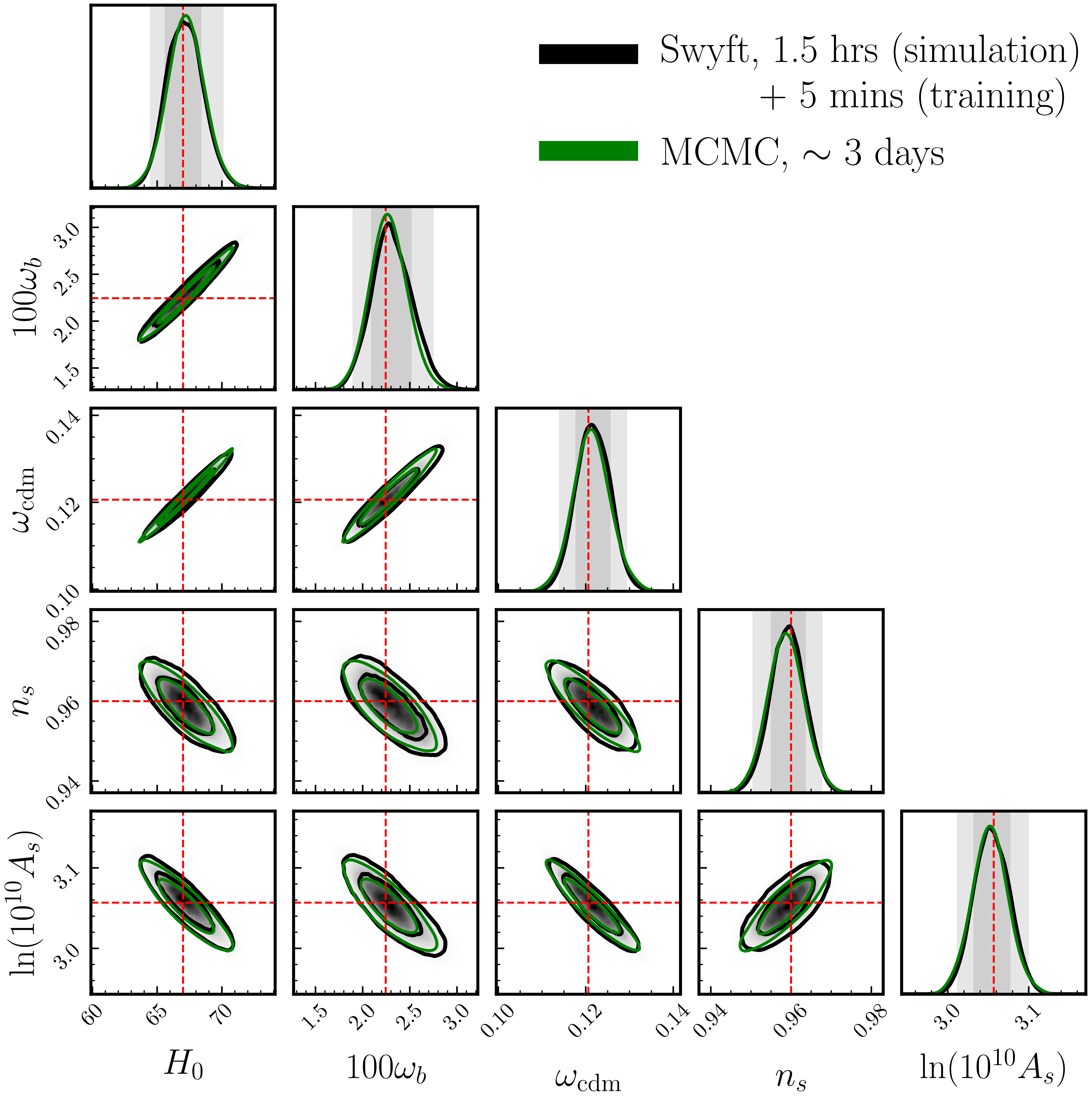
Galaxy clustering

Galaxy-Galaxy lensing

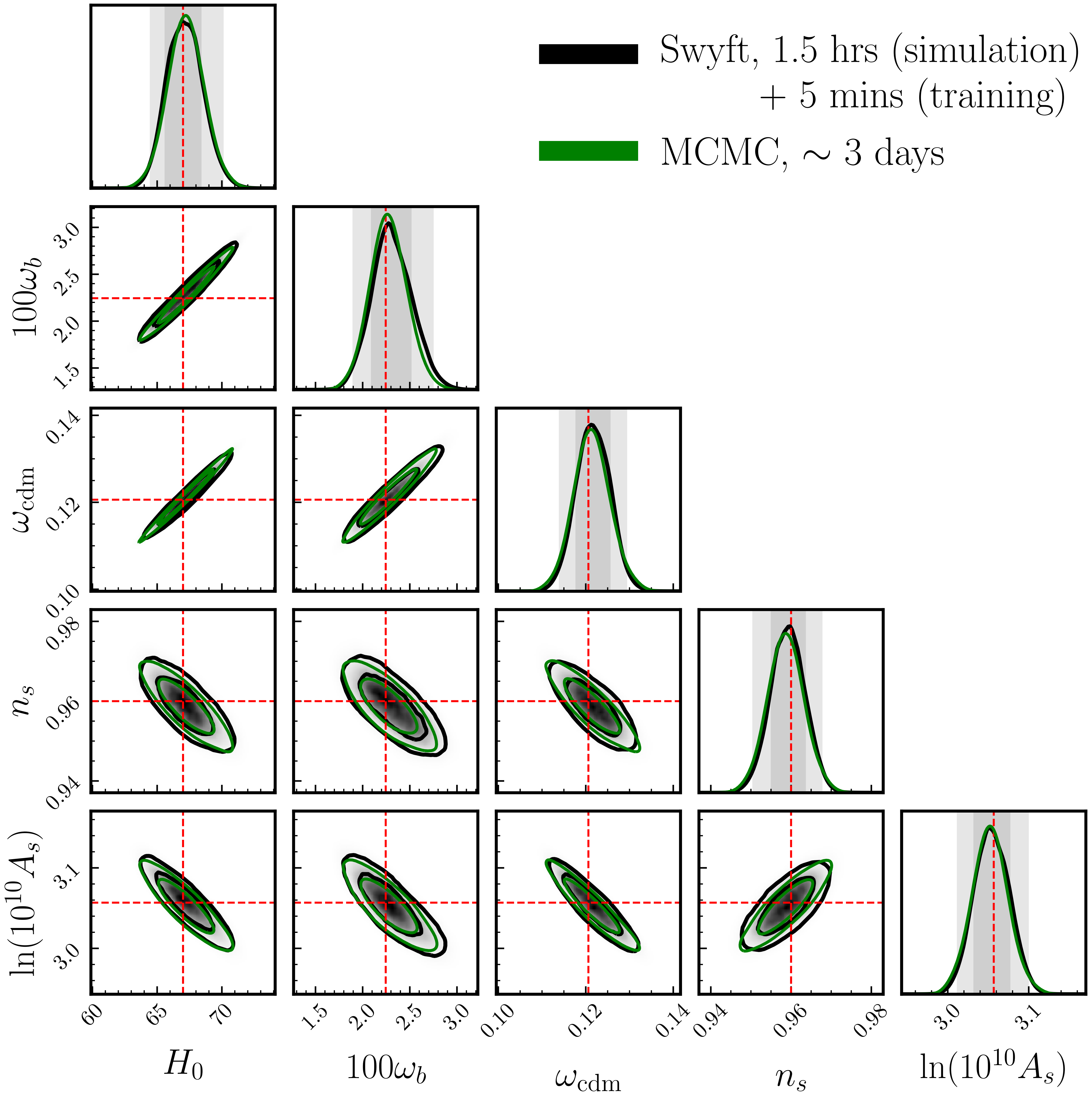
- We pre-compress spectra into features before passing this to the posterior networks



Forecast ΛCDM posteriors

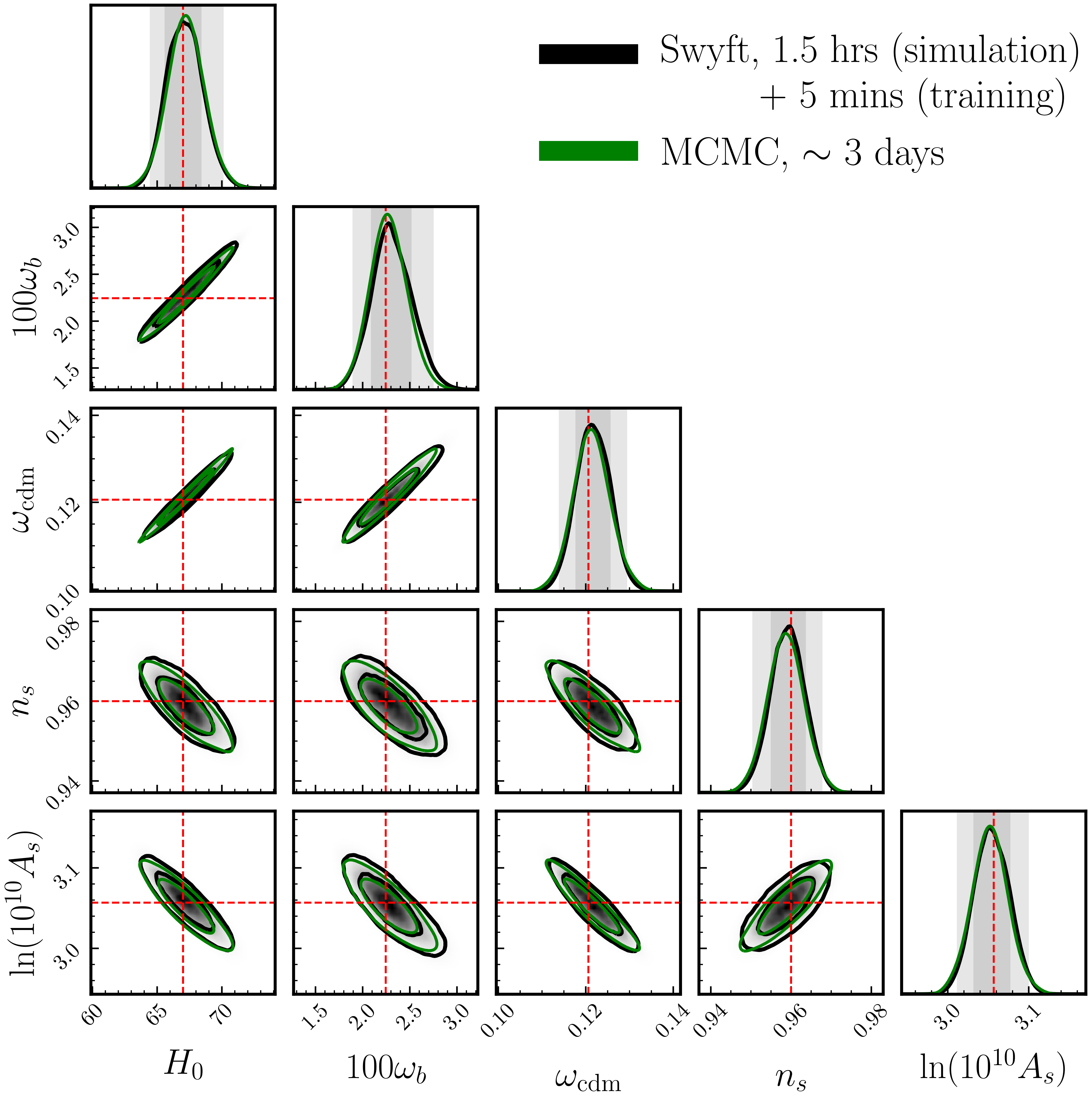


Forecast ΛCDM posteriors



MNRE & MCMC are in excellent agreement!

Forecast ΛCDM posteriors



Dramatic reduction in CPU time!

MNRE & MCMC are in excellent agreement!

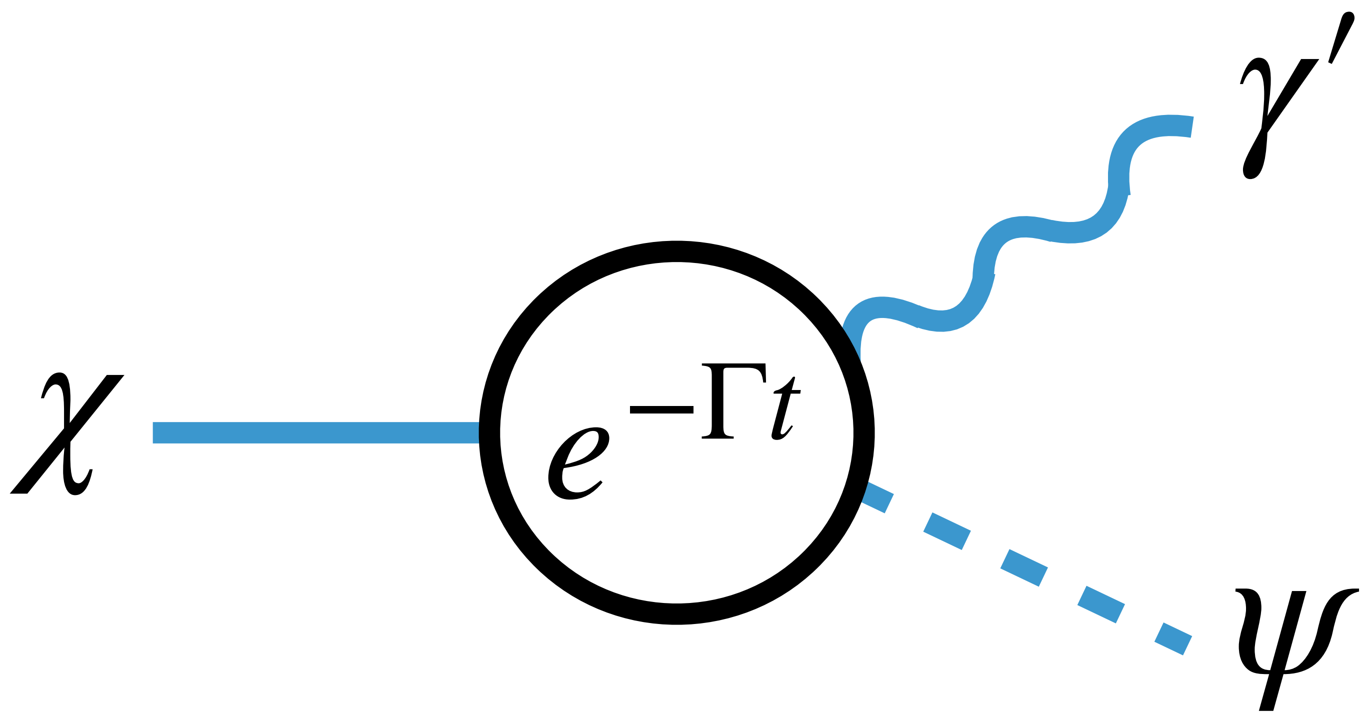
Does MNRE perform well with highly non-Gaussian posteriors?

- Does MNRE perform well with highly non-Gaussian posteriors?

- As an example, we test a model of **CDM decaying to DR + WDM** (proposed to explain the S_8 tension)

[Abellan+ 21]

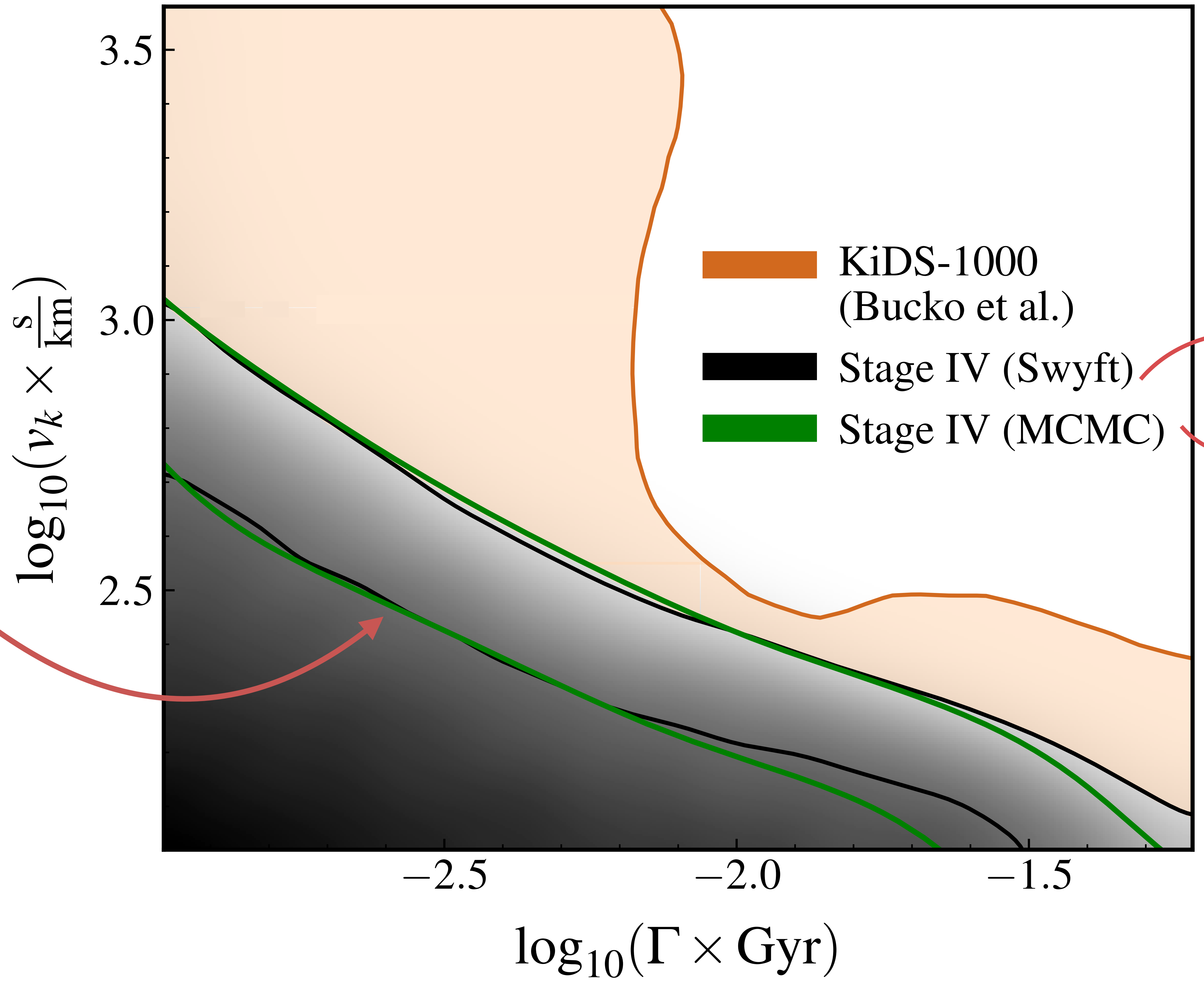
[Bucko+ 23]



Decay rate Γ
WDM velocity kick v_k

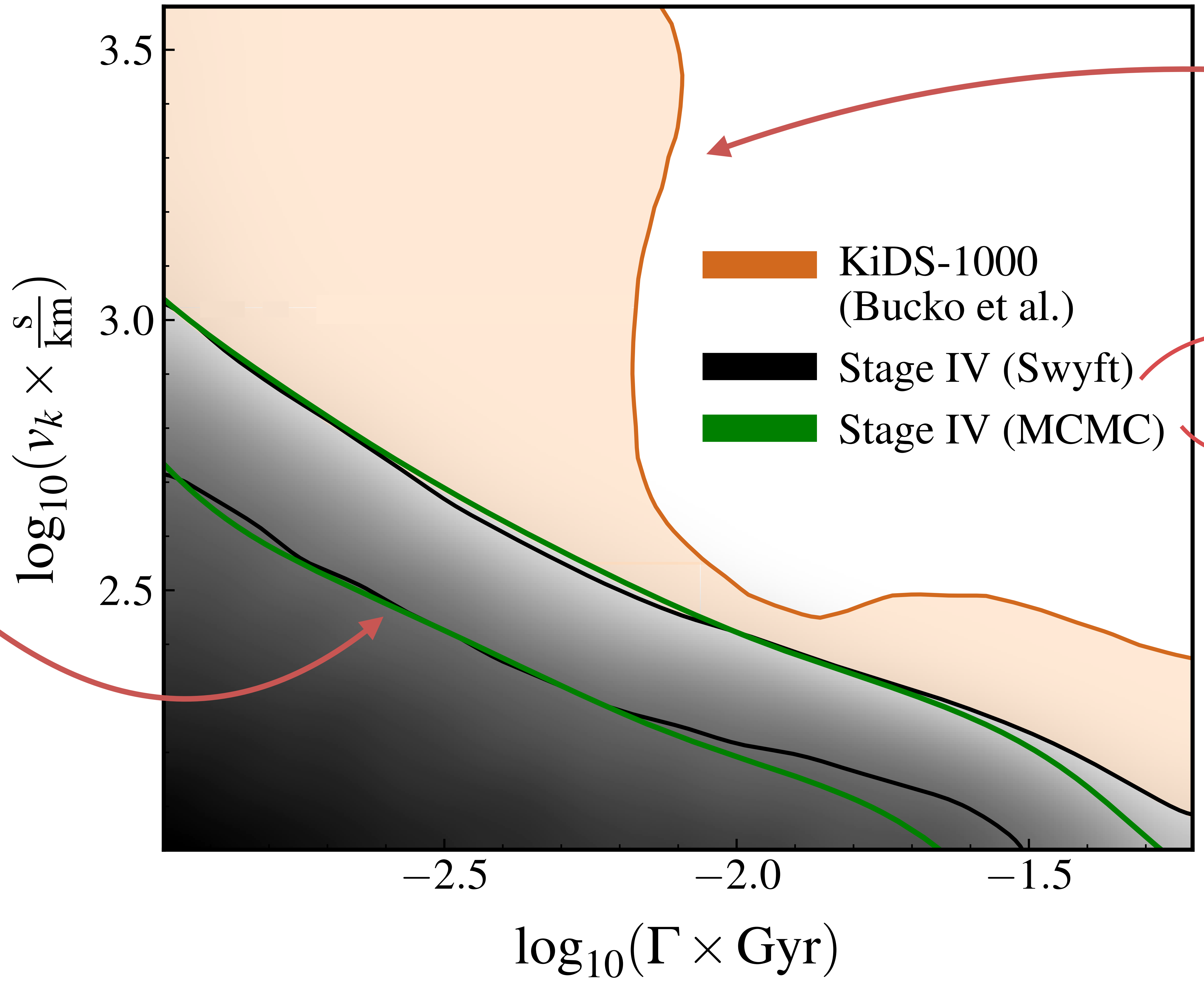
Forecast constraints on decaying DM

MNRE & MCMC in good agreement



3 hours vs. 8 days!

Forecast constraints on decaying DM



Improve current limits by ~1 order of magnitude!

3 hours vs. 8 days!

MNRE & MCMC in good agreement

Conclusions & prospects

ありがとう!

g.francoabellan@uva.nl

- Using a SBI method called MNRE, we performed parameter inference from 3x2pt Stage-IV probes with a speed-up factor of ~50 compared to MCMC

- We tested a model of decaying DM to check that MNRE performs well with very non-gaussian posteriors

- As a next step, build a Swyft simulator based on CLOE, the official Euclid likelihood (more nuisance parameters, add spectroscopic galaxy clustering)

BACK-UP

Swyft in a nutshell

Simulation-based inference

Use simulator to sample implicit likelihood

$$\mathbf{x} \sim p(\mathbf{x} \,|\, \boldsymbol{\theta})$$

Neural Ratio Estimation

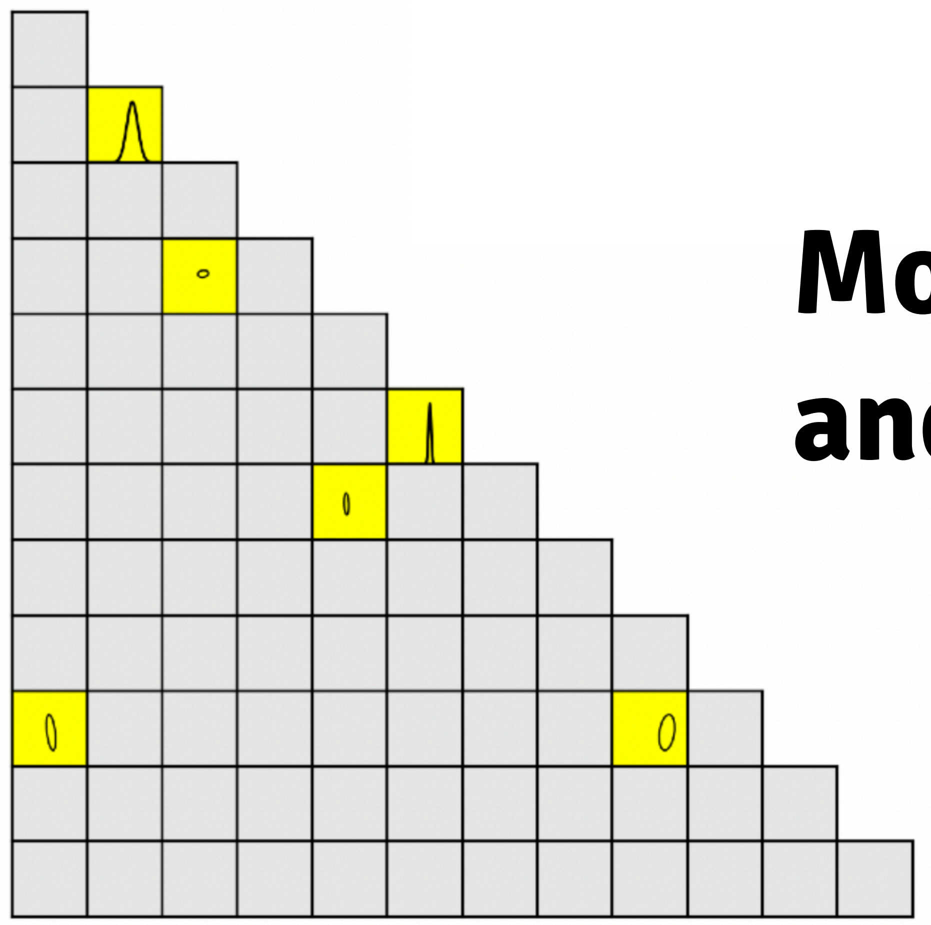
Get posteriors by training NNs to solve a binary classification problem

$$(\mathbf{x}, \boldsymbol{\theta}) \sim p(\mathbf{x}, \boldsymbol{\theta})$$

$$(\mathbf{x}, \boldsymbol{\theta}) \sim p(\mathbf{x})p(\boldsymbol{\theta})$$

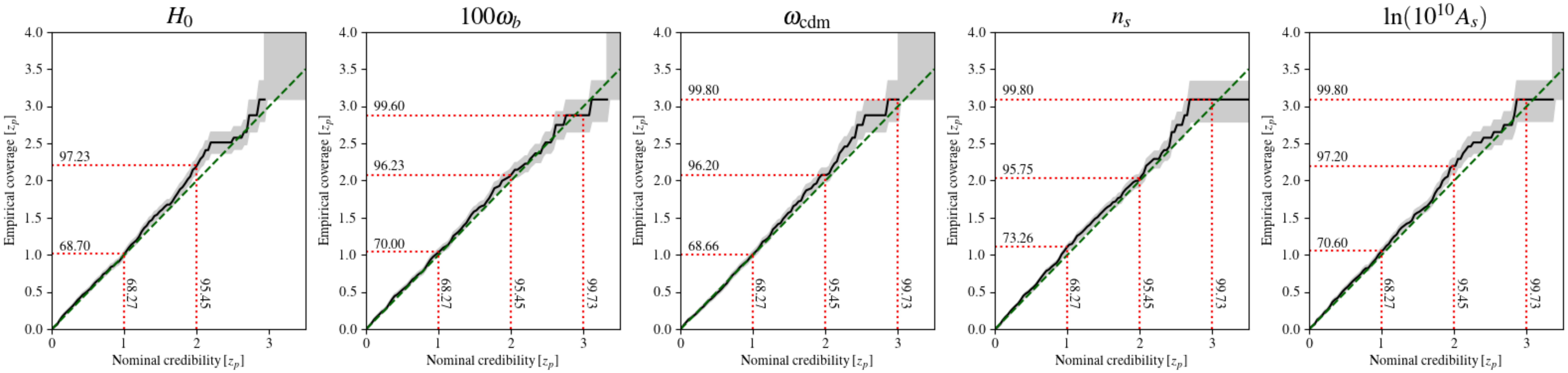
Focus on marginals

Instead of estimating all parameters, cherry-pick the ones we care about



More flexible and efficient!

Coverage test for Stage-IV 3x2pt



Empirical coverage and confidence level match to excellent precision!