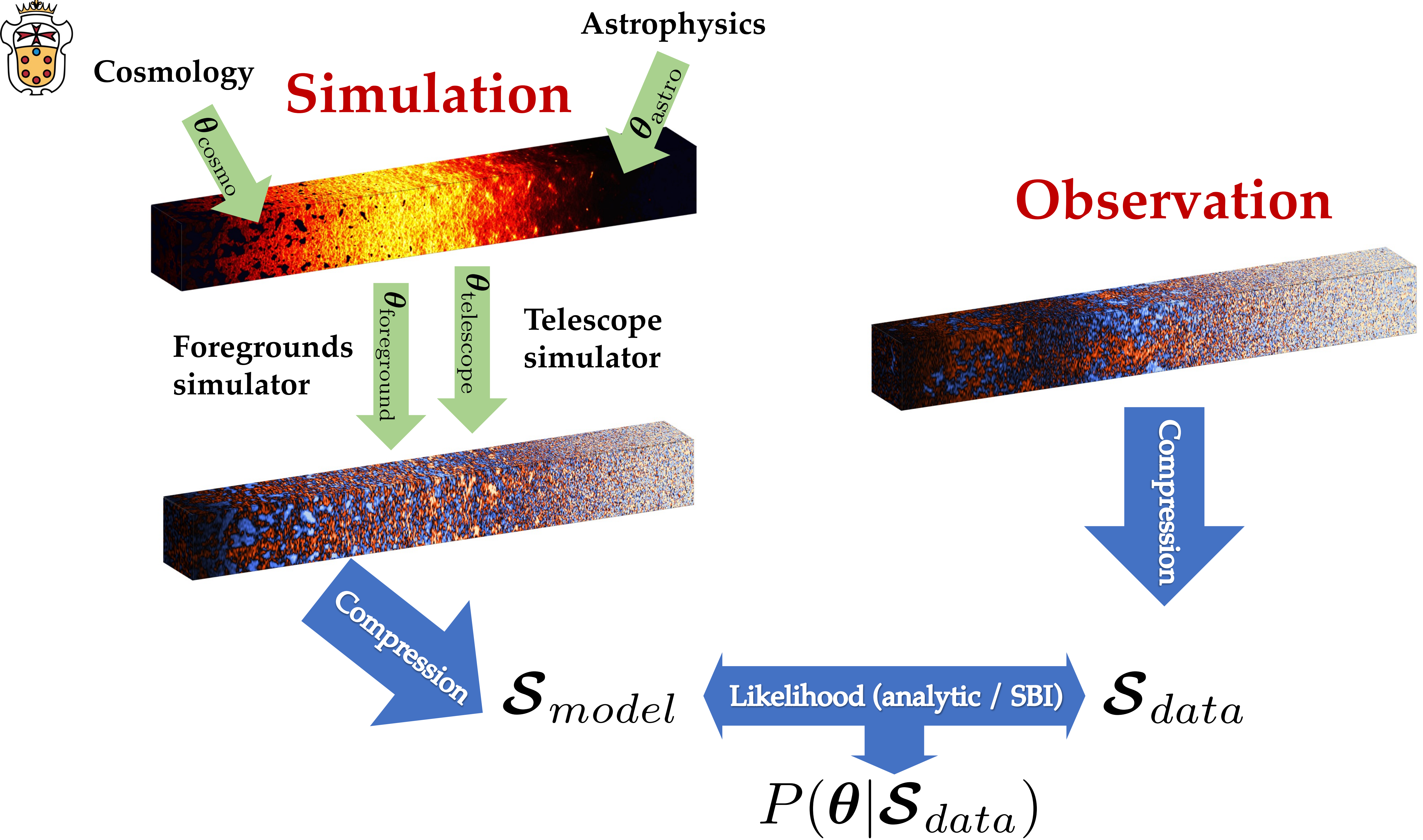
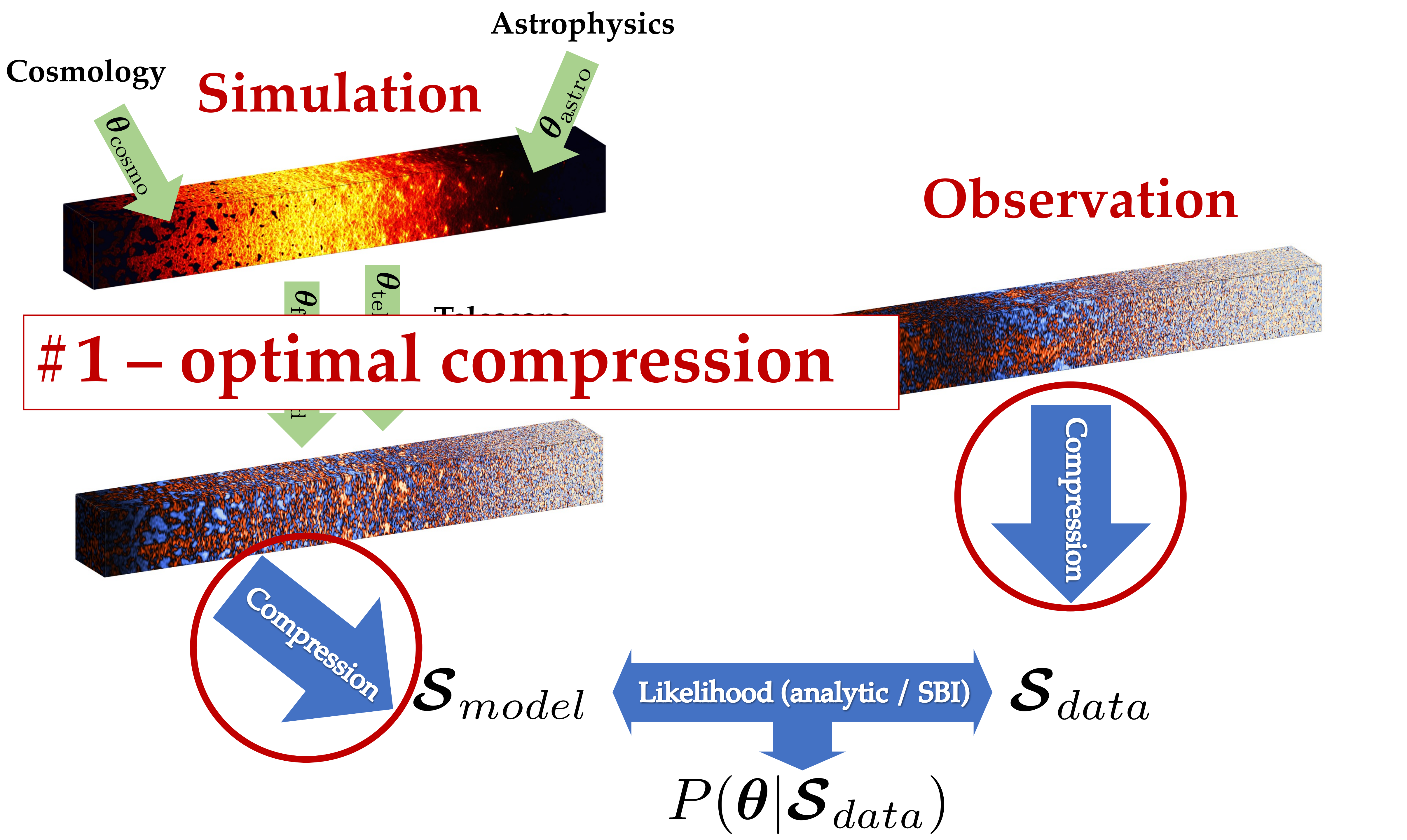
SCUOLA
NORMALE
SUPERIORE

Simulation-Based Inference of the cosmic 21-cm signal

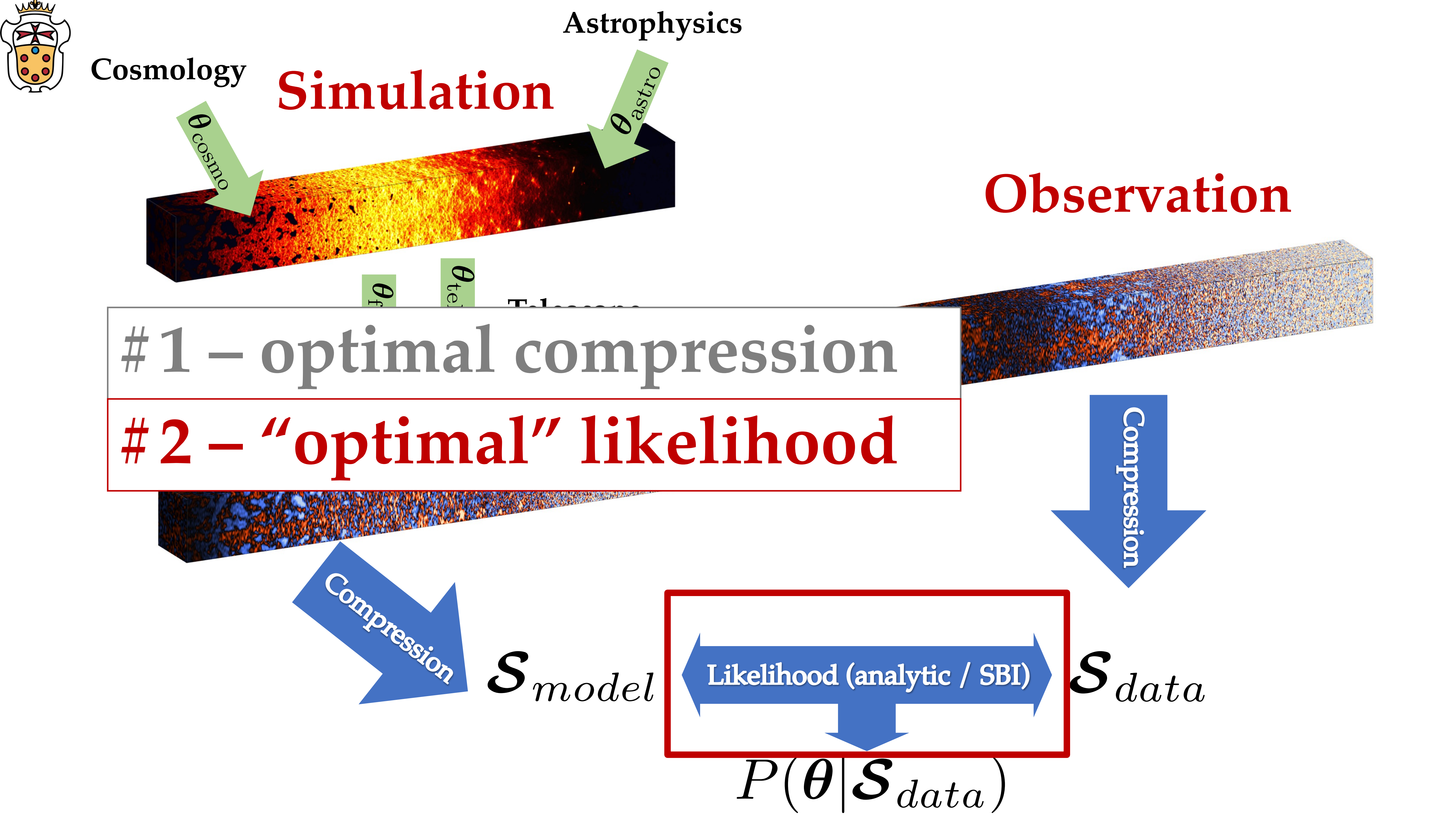
David Prelogović
PhD @ SNS Pisa
Supervisor: prof. Andrei Mesinger



Astrophysics
Cosmology
Simulation
θ_{cosmo}
θ_{astro}
Observation
Foregrounds simulator
$\theta_{foreground}$
$\theta_{telescope}$
Telescope simulator
Compression
Compression
$\mathcal{S}_{model}$
Likelihood (analytic / SBI)
$\mathcal{S}_{data}$
$P(\boldsymbol{\theta}|\mathcal{S}_{data})$



Astrophysics
Cosmology
Simulation
θ_{cosmo}
θ_{astro}
Observation
#1 – optimal compression
Compression
Compression
$\boldsymbol{s}_{model}$
Likelihood (analytic / SBI)
$\boldsymbol{s}_{data}$
$P(\boldsymbol{\theta}|\boldsymbol{s}_{data})$

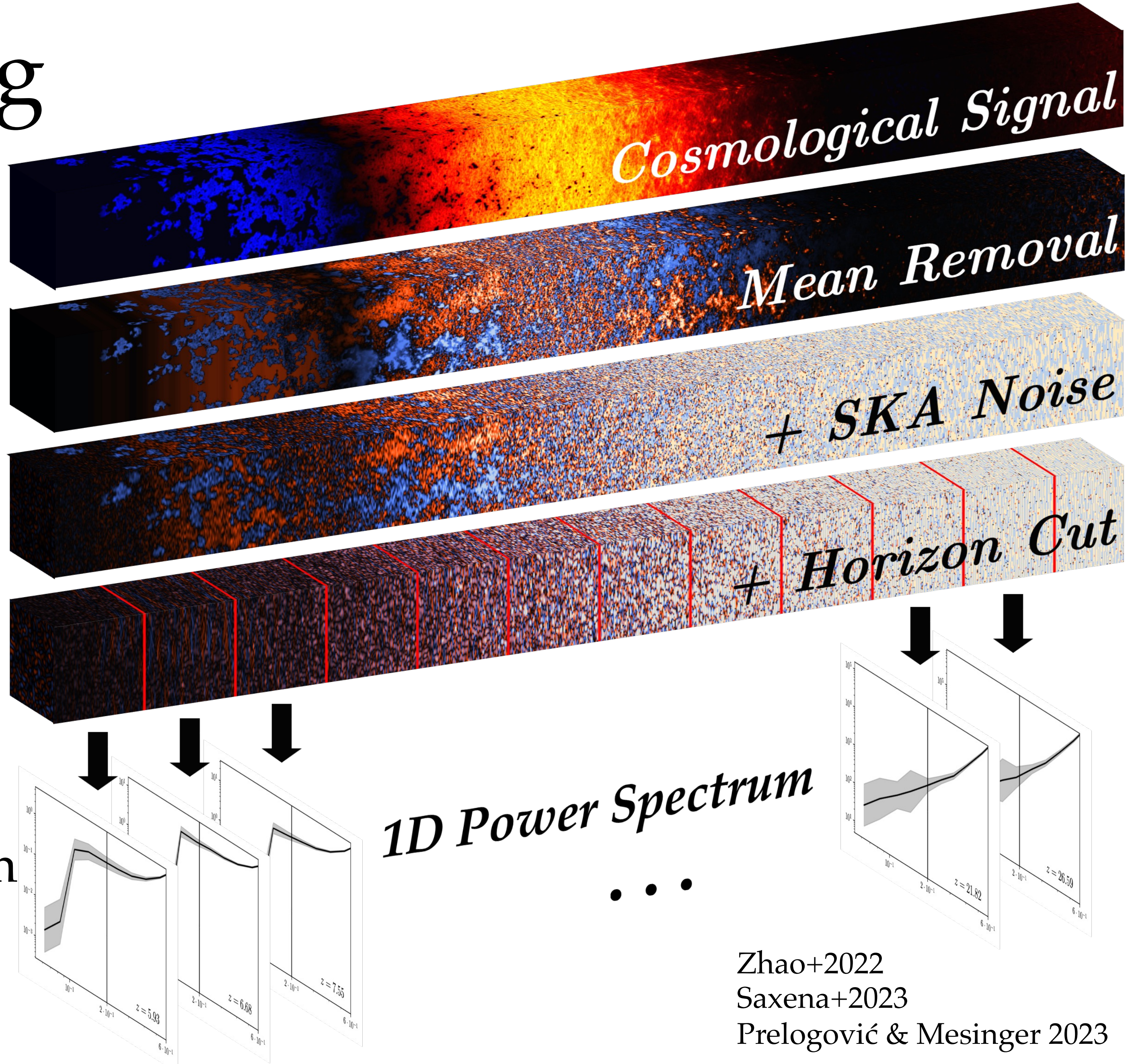


Astrophysics
Cosmology
Simulation
θ_{cosmo}
θ_{astro}
Observation
#1 – optimal compression
#2 – “optimal” likelihood
Compression
Compression
s_{model}
Likelihood (analytic / SBI)
s_{data}
$P(\theta|s_{data})$

SCUOLA NORMALE SUPERIORE

Forward-modelling

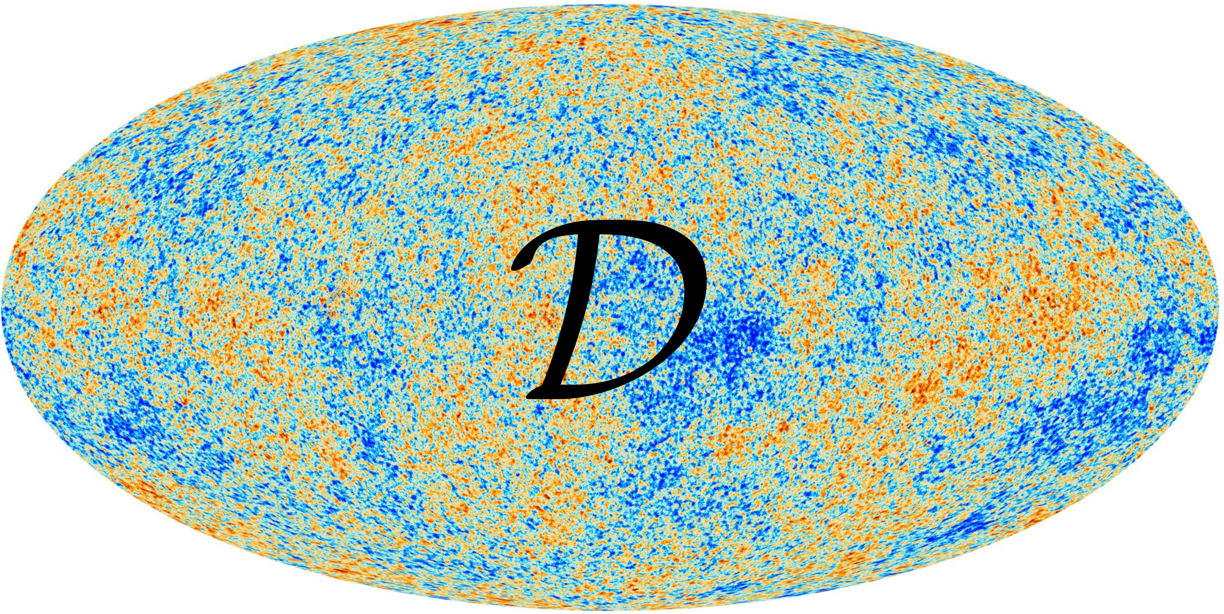
- Signal simulation: 21cmFAST
- SKA not sensitive to the mean
- UV coverage of the SKA and its noise
- Removing modes contaminated with foregrounds
- Compressing it to the power spectrum

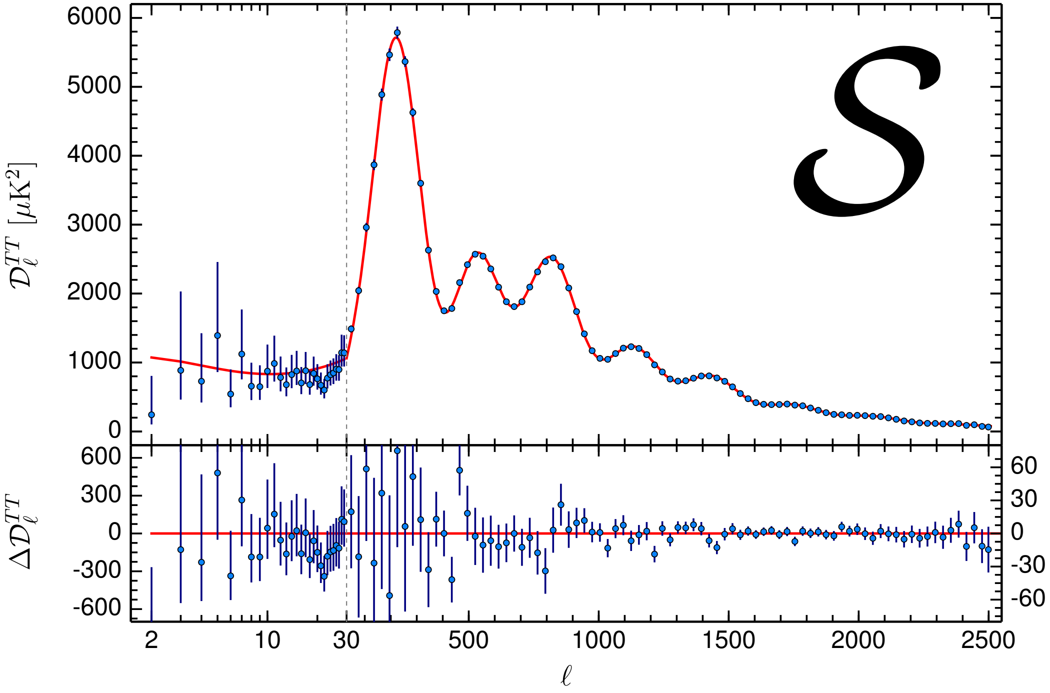


Zhao+2022
Saxena+2023
Prelogović & Mesinger 2023

SCUOLA
NORMALE
SUPERIORE

Example: CMB

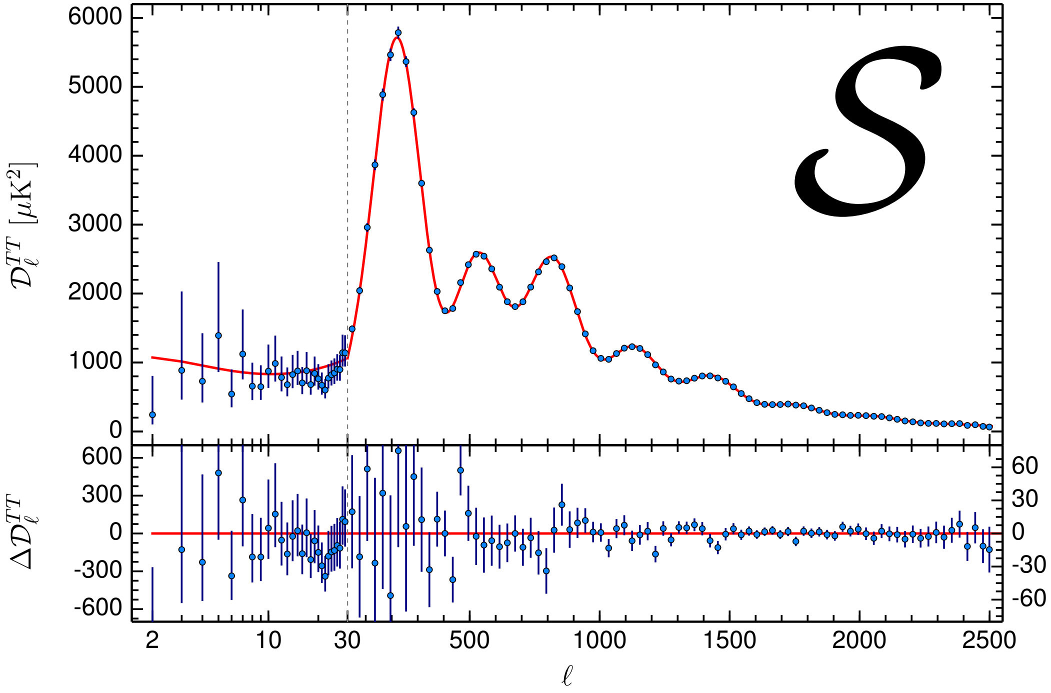




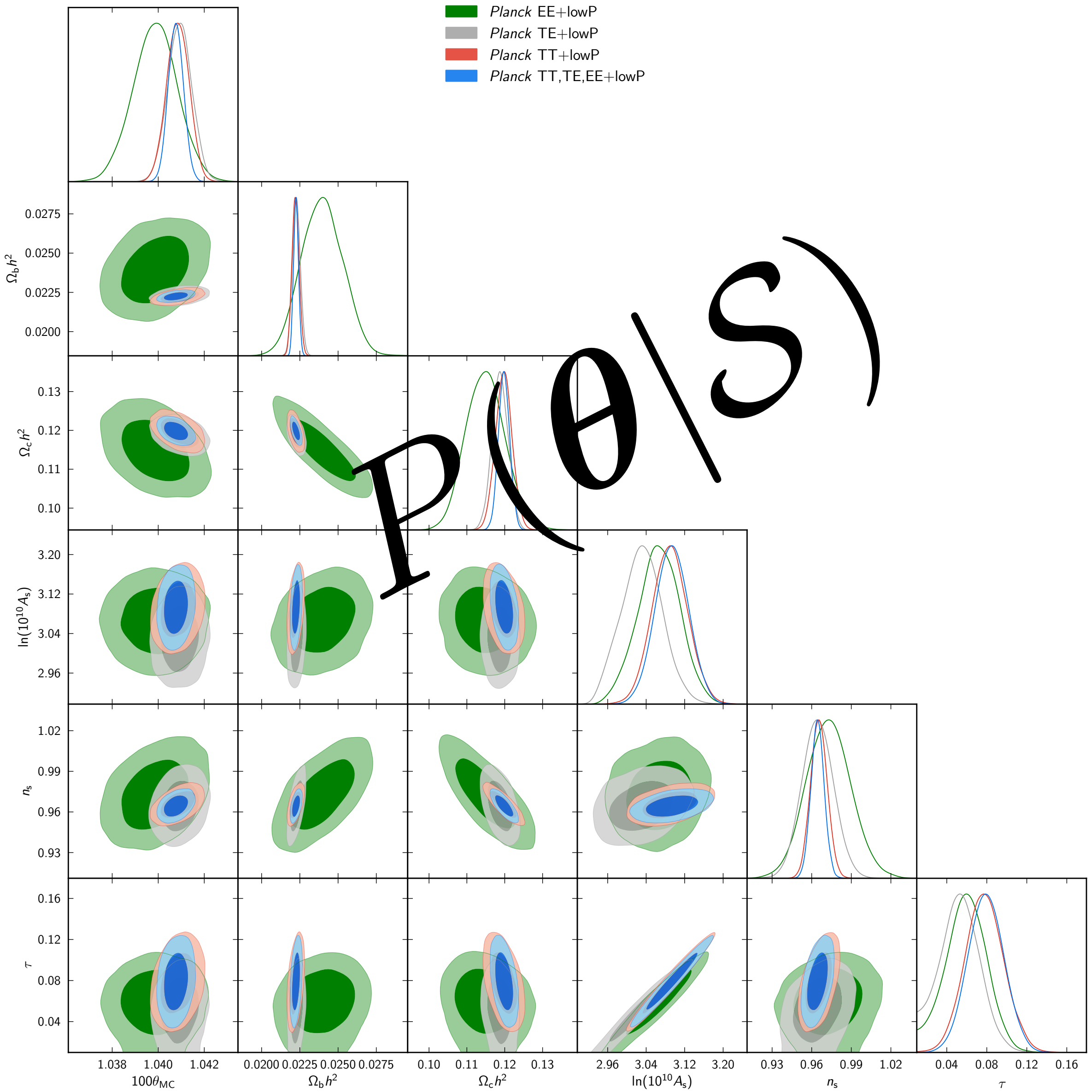
- Full sky map compressed to PS

Planck 2015

Example: CMB



→



- Full sky map compressed to PS
- From it we infer the cosmology
  - Known compression, known likelihood

Planck 2015

Likelihood of the 21-cm

- Analytically not available
- Numerically non-tractable

Likelihood of the 21-cm

- Analytically not available
- Numerically non-tractable
- Eg. – power spectrum

$$\mathcal{L}(\boldsymbol{\sigma}, \boldsymbol{\mu}; \mathbf{x}) = \frac{1}{(2\pi)^{n/2}\sqrt{\Pi_i \sigma_i^2}} e^{-\frac{1}{2}\sum_i (x_i - \mu_i)^2/\sigma_i^2}$$

$$\mathcal{L}(\boldsymbol{\Sigma}, \boldsymbol{\mu}; \mathbf{x}) = \frac{1}{(2\pi)^{n/2}\sqrt{|\Sigma|}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{x}-\boldsymbol{\mu})}$$

SCUOLA NORMALE SUPERIORE

Likelihood of the 21-cm

- Analytically not available
- Numerically non-tractable
- Eg. – power spectrum

$$\mathcal{L}(\boldsymbol{\sigma}, \boldsymbol{\mu}; \mathbf{x}) = \frac{1}{(2\pi)^{n/2}\sqrt{\Pi_i \sigma_i^2}} e^{-\frac{1}{2}\sum_i (x_i - \mu_i)^2/\sigma_i^2}$$

$$\mathcal{L}(\boldsymbol{\Sigma}, \boldsymbol{\mu}; \mathbf{x}) = \frac{1}{(2\pi)^{n/2}\sqrt{|\Sigma|}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{x}-\boldsymbol{\mu})}$$

Classical inference simplifications

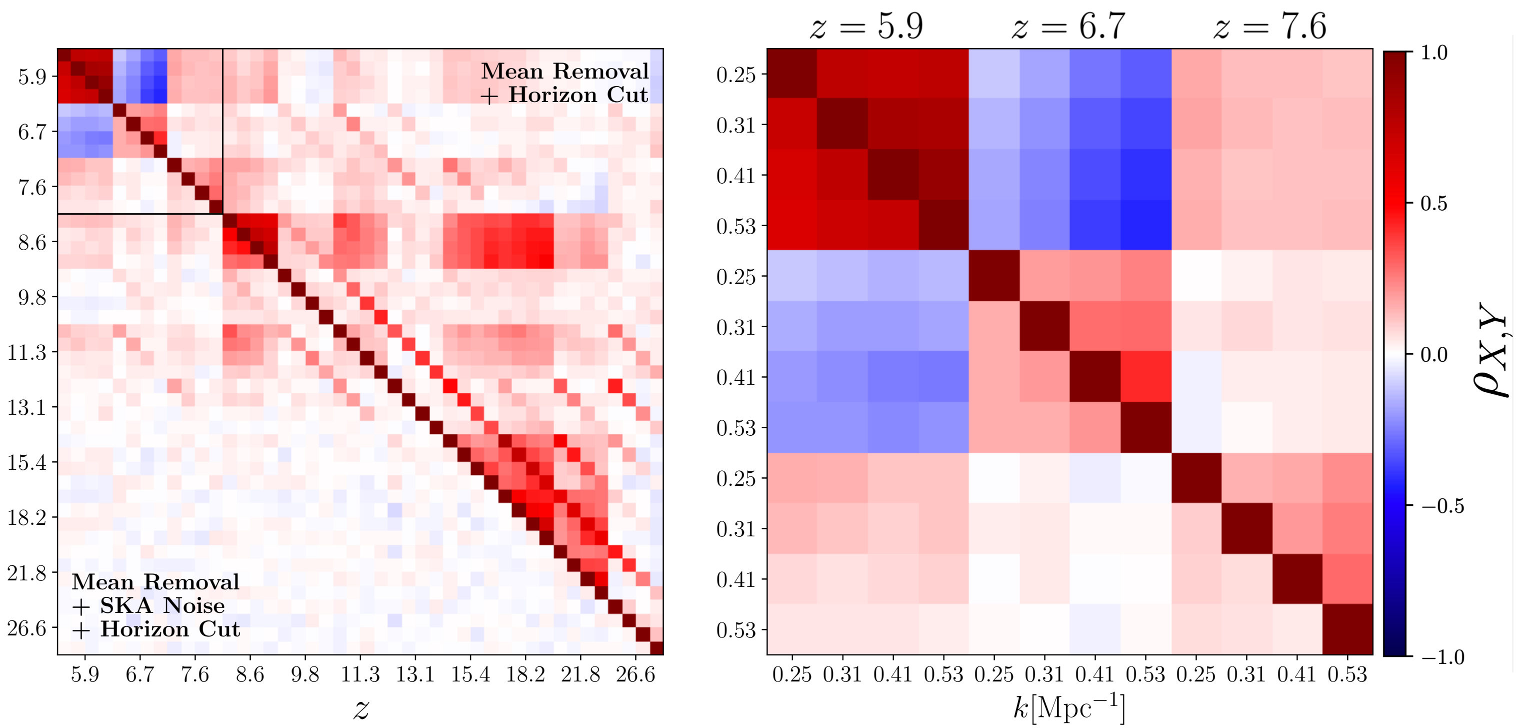
1) μ estimated from one sim.
2) $\Sigma = \Sigma_{\theta \mathrm{fid}}$
3) No higher order terms

–> Simulation-Based Inference

SCUOLA
NORMALE
SUPERIORE

Likelihood of the 21-cm PS

- Is inclusion of the covariance worthwhile?



- Large correlations across z and k
- Only diagonal taken into account in usual analyses

Prelogović & Mesinger 2023

SCUOLA NORMALE SUPERIORE

Simulation-Based Inference (SBI)

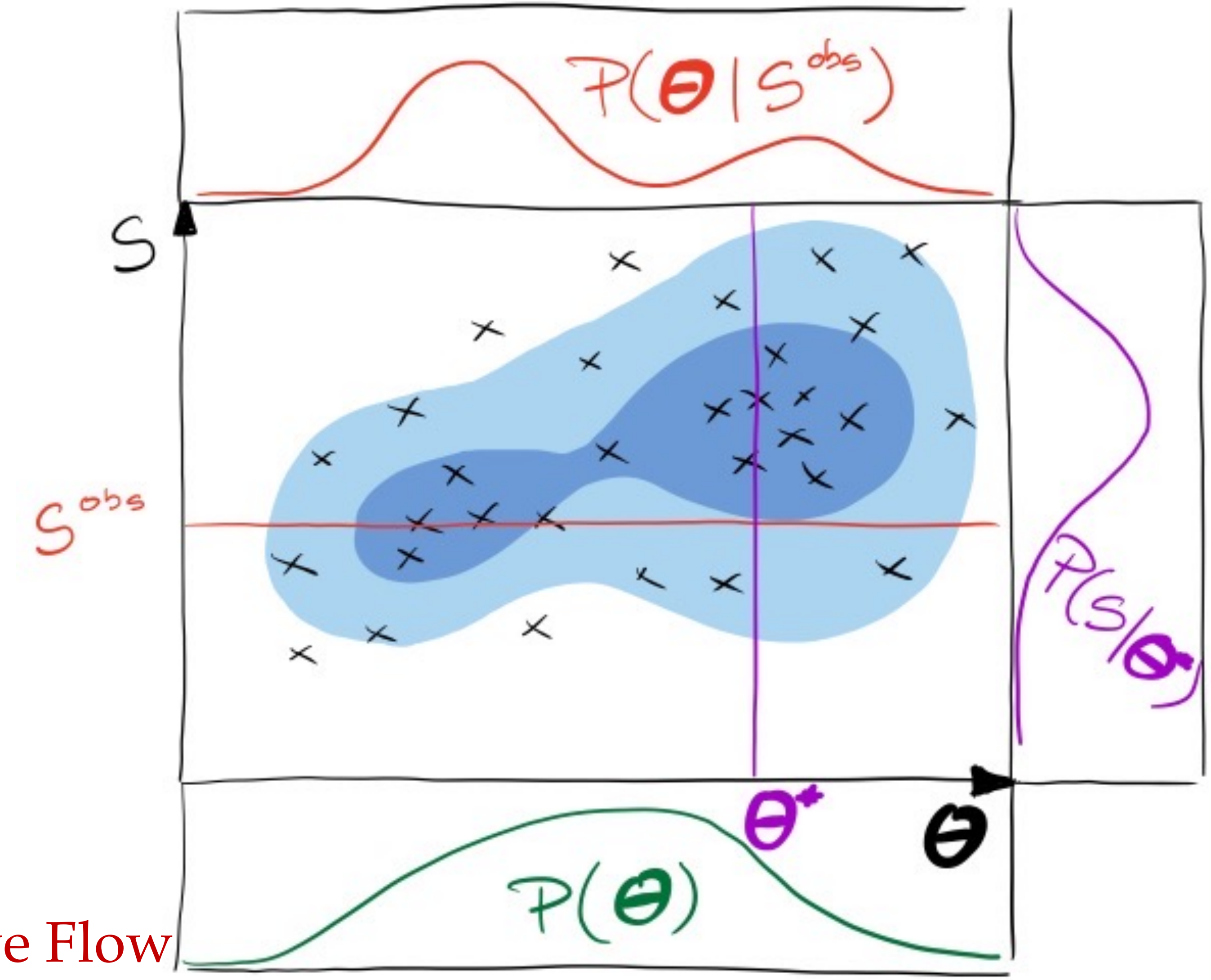
- Pull from the prior $\boldsymbol{\theta}^* \sim P(\boldsymbol{\theta})$
- Pull from the likelihood

$$\mathcal{S}^* \sim P(\mathcal{S}|\boldsymbol{\theta}^*)$$

i.e. simulate and compress

$$\mathcal{S}^* = \text{compress}(\text{simulate}(\boldsymbol{\theta}^*))$$

- Repeat many times
- Fit the distribution with **NDEs**
  - Gaussian
  - Gaussian mixture
  - Conditional Masked Autoregressive Flow



Credit: Tom Charnock

SCUOLA
NORMALE
SUPERIORE

Inference for the 21-cm PS

- Two classic likelihoods
  - CLASSIC fixed var
  - CLASSIC fixed cov
- Neural density estimators
  - NDE fixed var
  - NDE fixed cov
  - NDE varying var
  - NDE varying cov
  - NDE CMAF
  - NDE Gauss mixture

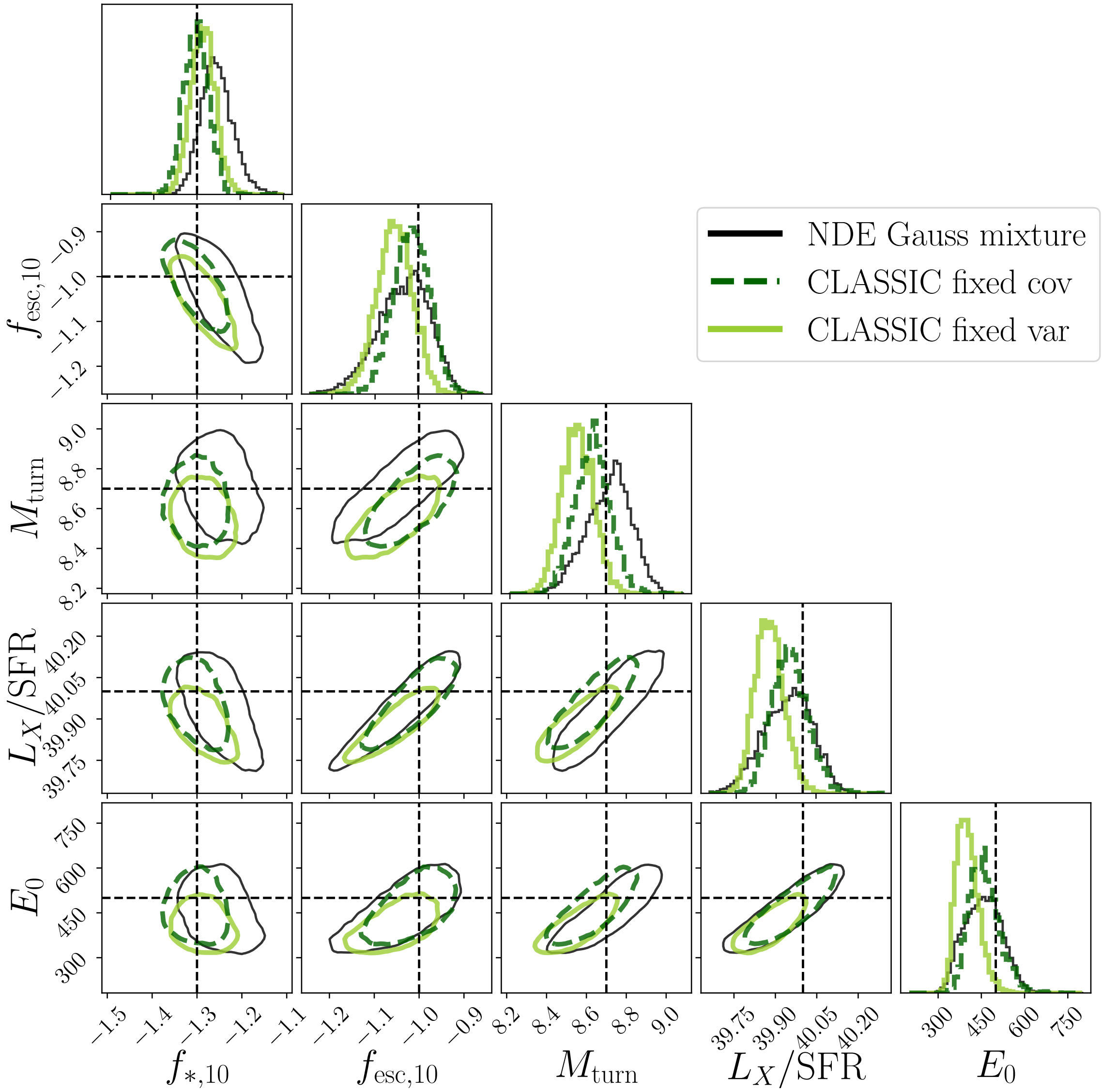
- Fixed covariance (variance)
- Mean estimated from one simulation

1) Fitting the mean

2) Fitting the mean and covariance

3) Adding all higher-order moments and multimodalities

Inference for the 21-cm PS



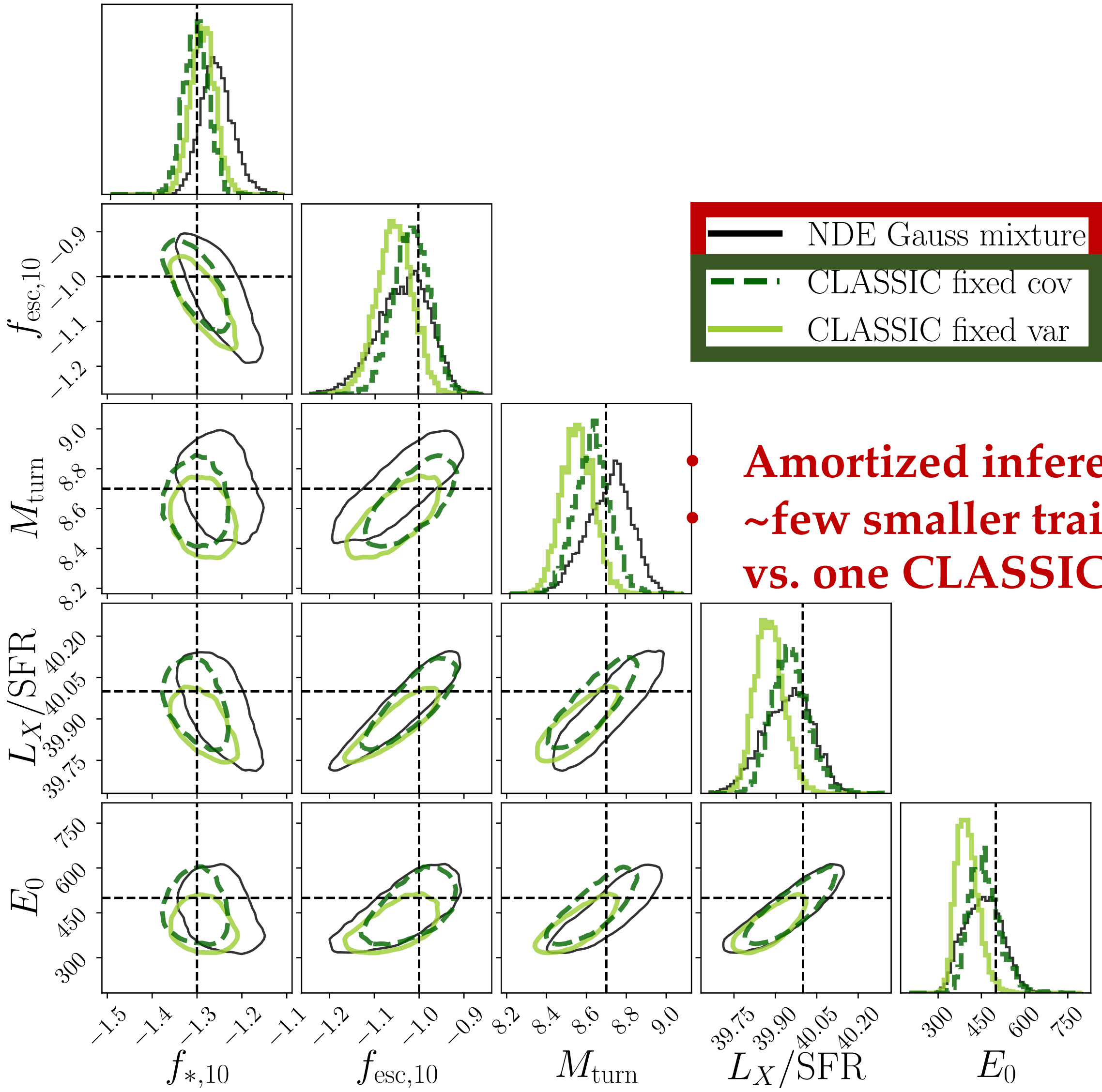
Including more realistic likelihood ≠ more constraining posterior

SCUOLA NORMALE SUPERIORE

Prelogović & Mesinger 2023

Inference for the 21-cm PS

Including more realistic likelihood ≠ more constraining posterior



0.1-1 CPU h

100 000 CPU h

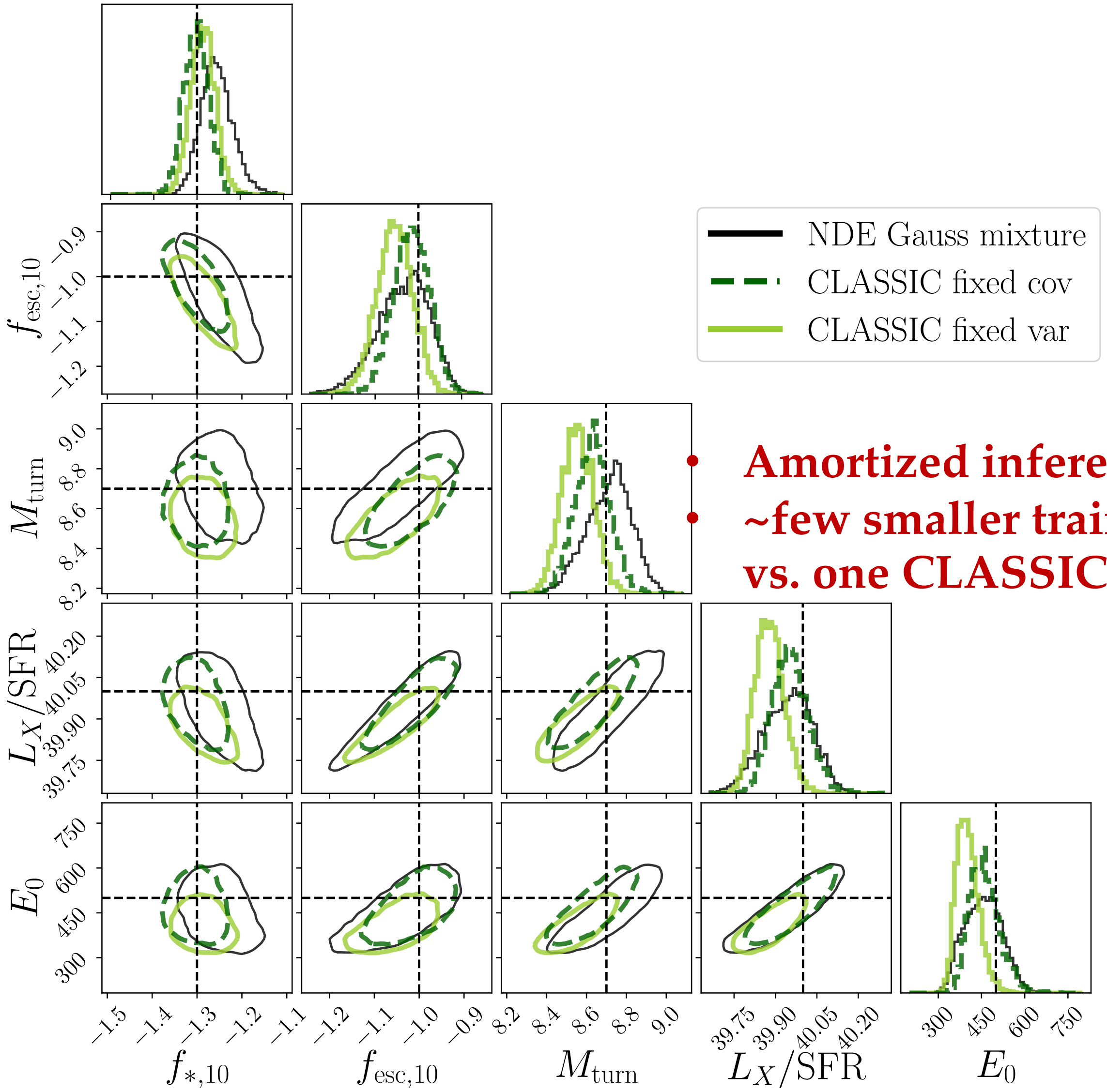
- **Amortized inference**
- **~few smaller training DB vs. one CLASSIC run**

SCUOLA
NORMALE
SUPERIORE

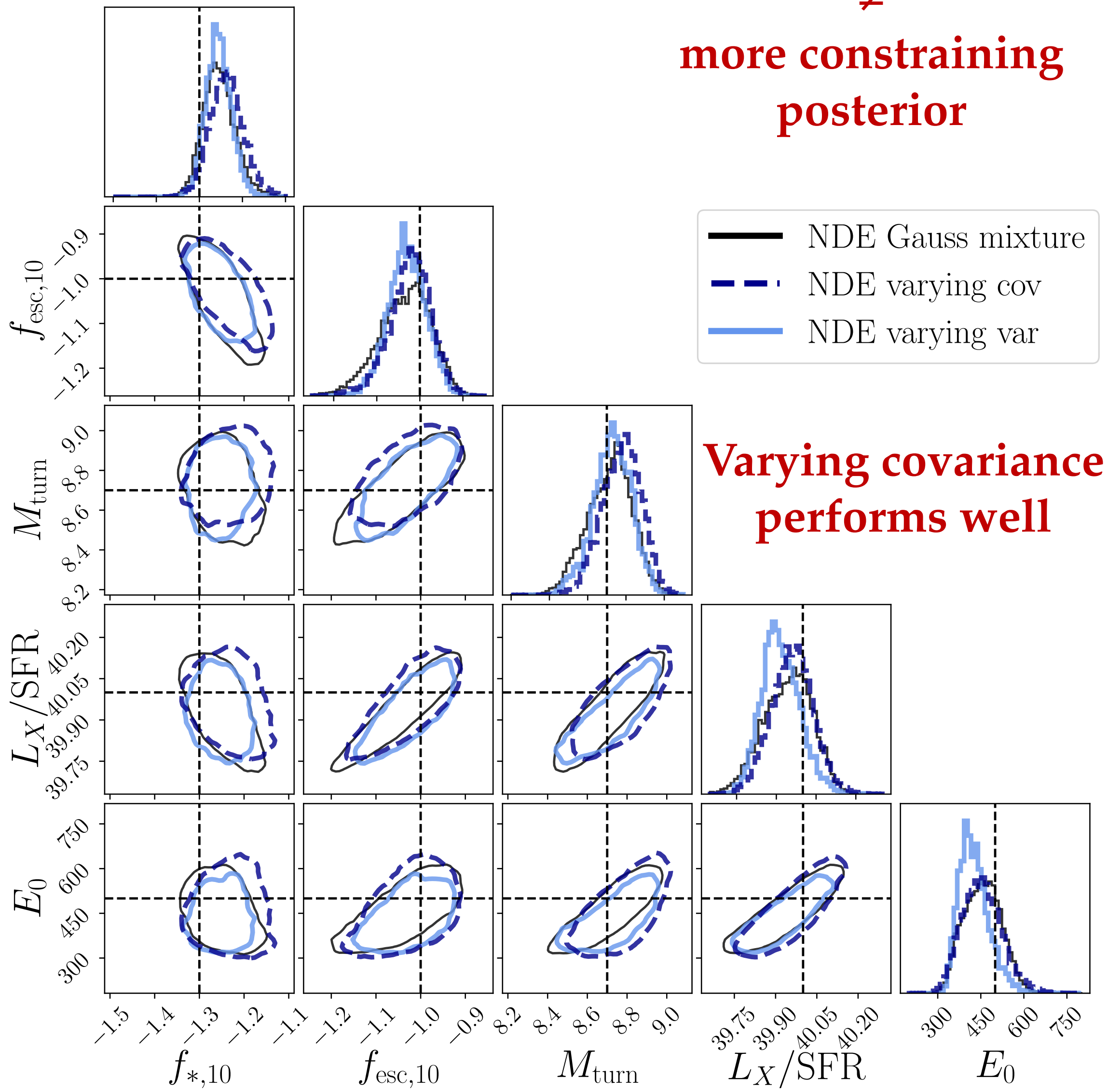
Prelogović & Mesinger 2023

Inference for the 21-cm PS

Including more realistic likelihood ≠ more constraining posterior



- **Amortized inference**
- **~few smaller training DB vs. one CLASSIC run**



Varying covariance performs well

Inference for the 21-cm PS

BUT:

This is only qualitative description, and only for the mock observation

- How does it perform for other points in the parameter space?
- Did the training converge?
- Can we quantify the best model?

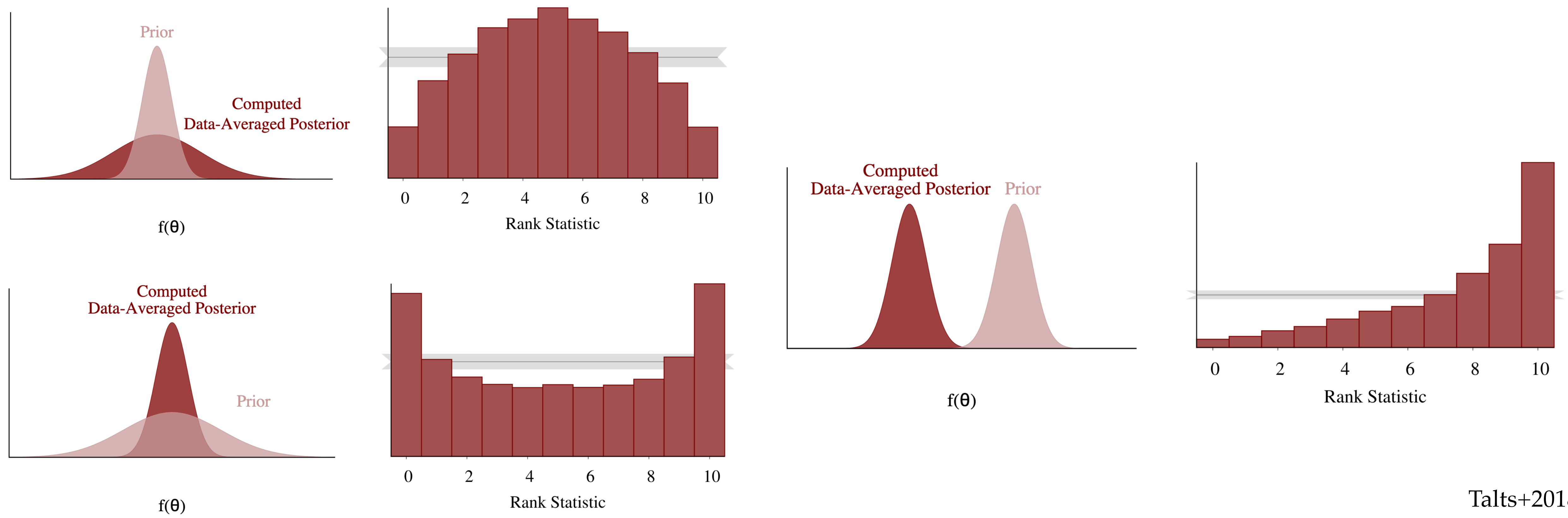
–> Simulation-Based Calibration

Simulation Based Calibration (SBC)

- ***"prior" = "data averaged posterior"*** $P(\boldsymbol{\theta}) = \int P(\boldsymbol{\theta}|\tilde{\boldsymbol{y}})\, P(\tilde{\boldsymbol{y}}|\tilde{\boldsymbol{\theta}}) P(\tilde{\boldsymbol{\theta}})\, \mathrm{d}\tilde{\boldsymbol{y}}\, \mathrm{d}\tilde{\boldsymbol{\theta}}$

Simulation Based Calibration (SBC)

- ***"prior" = "data averaged posterior"*** $P(\boldsymbol{\theta}) = \int P(\boldsymbol{\theta}|\tilde{\boldsymbol{y}})\, P(\tilde{\boldsymbol{y}}|\tilde{\boldsymbol{\theta}}) P(\tilde{\boldsymbol{\theta}})\, d\tilde{\boldsymbol{y}}\, d\tilde{\boldsymbol{\theta}}$
- SBC – casting integral into 1D rank statistics distribution



Talts+2018

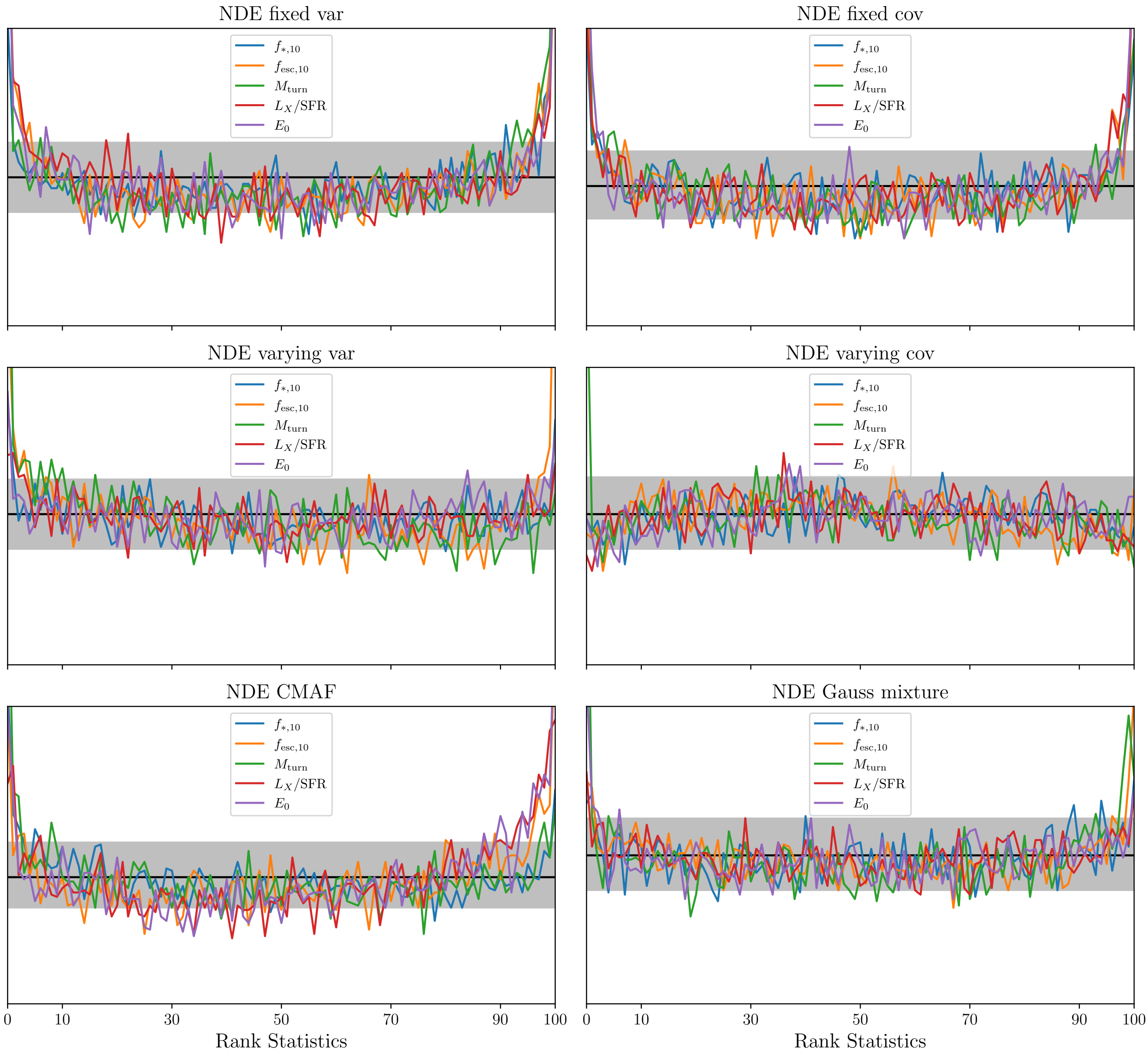
SCUOLA NORMALE SUPERIORE

SBC for 21-cm PS

- *Expensive to compute!*
  - *10 000 posteriors*

- However, once likelihood is trained, no new simulations are needed
  - "amortized inference"
- Procedure would be useful for classic inference too, but is impossible to compute

- NDE fixed var & cov – overconfident
- NDE varying var & cov – good
- NDE CMAF – overconfident
- NDE Gauss mixture – the best



Prelogović & Mesinger 2023

Conclusions

- 21cm – likelihood unavailable
- Simulation-Based Inference
  - Offers a way out
  - With crucial caveats to be taken care of
    - Consistency checks (SBC)
- Future (current) prospects
  - PS for non-Gaussian data is lossy
  - Other summaries paired with SBI might perform better

arXiv: 2305.03074
github: dprelogo/21cmLikelihoods