Less is enough:
extending ΛCDM with representation learning

Davide Piras (and many others)

UNIVERSITÉ DE GENÈVE

UCL

But first… let me apologise

But first… let me apologise

- Title was not convincing

Less is enough:
extending ΛCDM with representation learning

But first… let me apologise

- Title was not convincing

Less is enough:
extending ΛCDM with representation learning

arXiv > astro-ph > arXiv:2303.17059

Astrophysics > Instrumentation and Methods for Astrophysics

[Submitted on 29 Mar 2023]

As a matter of colon: I am NOT digging cheeky titles (no, but actually yes :>)

Joanne Tan, Tie Sien Suk

But first… let me apologise

- Title was not convincing

parameters

Less is enough:

extending ΛCDM with representation learning

But first… let me apologise

- Title was not convincing

parameters

~~Less~~ is enough:

extending ΛCDM with representation learning

- So I did what any AI researcher would do…

But first… let me apologise

- Title was not convincing

Less parameters is enough:
extending ΛCDM with representation learning

- So I did what any AI researcher would do…

- … I asked ChatGPT

Enriching ΛCDM:
extending cosmological models with representation learning

But first… let me apologise

- Yes, I also tried the newer versions

But first… let me apologise

- Yes, I also tried the newer versions
- Gave me the same answer…

But first… let me apologise

- Yes, I also tried the newer versions
- Gave me the same answer…
- … just faster

Less is enough:

extending ΛCDM with representation learning

Less is enough:

extending ΛCDM with representation learning

ΛCDM extensions

- ΛCDM is good

ΛCDM extensions

- ΛCDM is good

Λ : cosmological constant
CDM: cold dark matter

[insert standard cosmological image here]

ΛCDM extensions

- ΛCDM is good… but not the entire story

ΛCDM extensions

- ΛCDM is good… but not the entire story
    - Tensions (H_0, S_8)

ΛCDM extensions

- ΛCDM is good… but not the entire story
  - Tensions (H_0, S_8)
  - What is dark matter?

ΛCDM extensions

- ΛCDM is good… but not the entire story

  - Tensions (H_0, S_8)
  - What is dark matter?
  - And dark energy?

ΛCDM extensions

- ΛCDM is good… but not the entire story
    - Tensions (H_0, S_8)
    - What is dark matter?
    - And dark energy?
    - …

ΛCDM extensions

- ΛCDM is good… but not the entire story
- Beyond-ΛCDM models add extra parameters

ΛCDM extensions

- ΛCDM is good… but not the entire story

- Beyond-ΛCDM models add extra parameters

CDM

$$\boxed{\Omega_b \ \Omega_m \ h \ n_s \ A_s}$$

ΛCDM extensions

- ΛCDM is good... but not the entire story

- Beyond-ΛCDM models add extra parameters

CDM

$\boxed{\Omega_b\ \Omega_m\ h\ n_s\ A_s}$

$w_0 w_a$CDM

$\boxed{\Omega_b\ \Omega_m\ h\ n_s\ A_s}\boxed{w_0\ w_a}$

ΛCDM extensions

- ΛCDM is good… but not the entire story

- Beyond-ΛCDM models add extra parameters

CDM

$\boxed{\Omega_b\ \Omega_m\ h\ n_s\ A_s}$

f(R)

$\boxed{\Omega_b\ \Omega_m\ h\ n_s\ A_s}\boxed{f_{R0}}$

ΛCDM extensions

- ΛCDM is good… but not the entire story

- Beyond-ΛCDM models add extra parameters

CDM

$\boxed{\Omega_b \; \Omega_m \; h \; n_s \; A_s}$

Dvali-Gabadadze-Porrati

$\boxed{\Omega_b \; \Omega_m \; h \; n_s \; A_s}\boxed{\Omega_{rc}}$

ΛCDM extensions

- ΛCDM is good… but not the entire story

- Beyond-ΛCDM models add extra parameters

CDM

Ω_b Ω_m h n_s A_s

…

Ω_b Ω_m h n_s A_s …

ΛCDM extensions

- ΛCDM is good… but not the entire story
- Beyond-ΛCDM models add extra parameters
- Find common parameterisation of all these models?

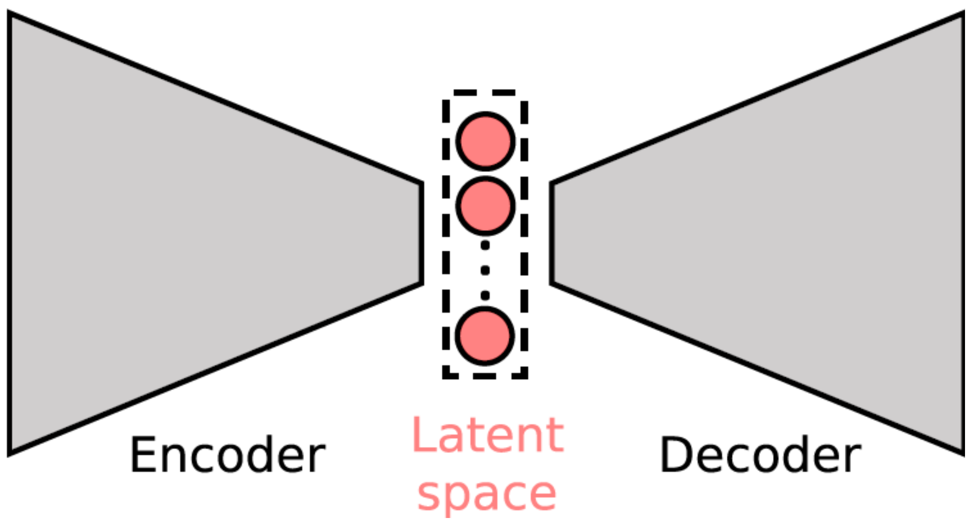
Less is enough:

extending ΛCDM with representation learning

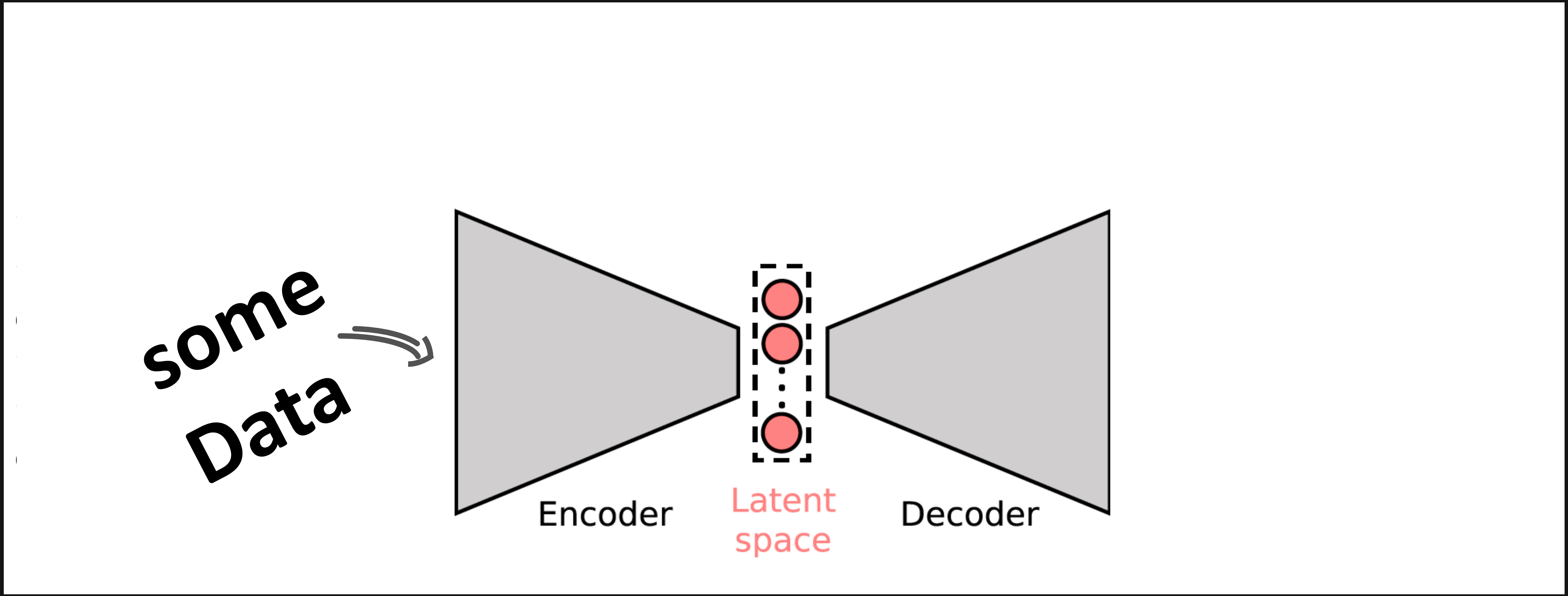
Less is enough:

extending ΛCDM **with representation learning**

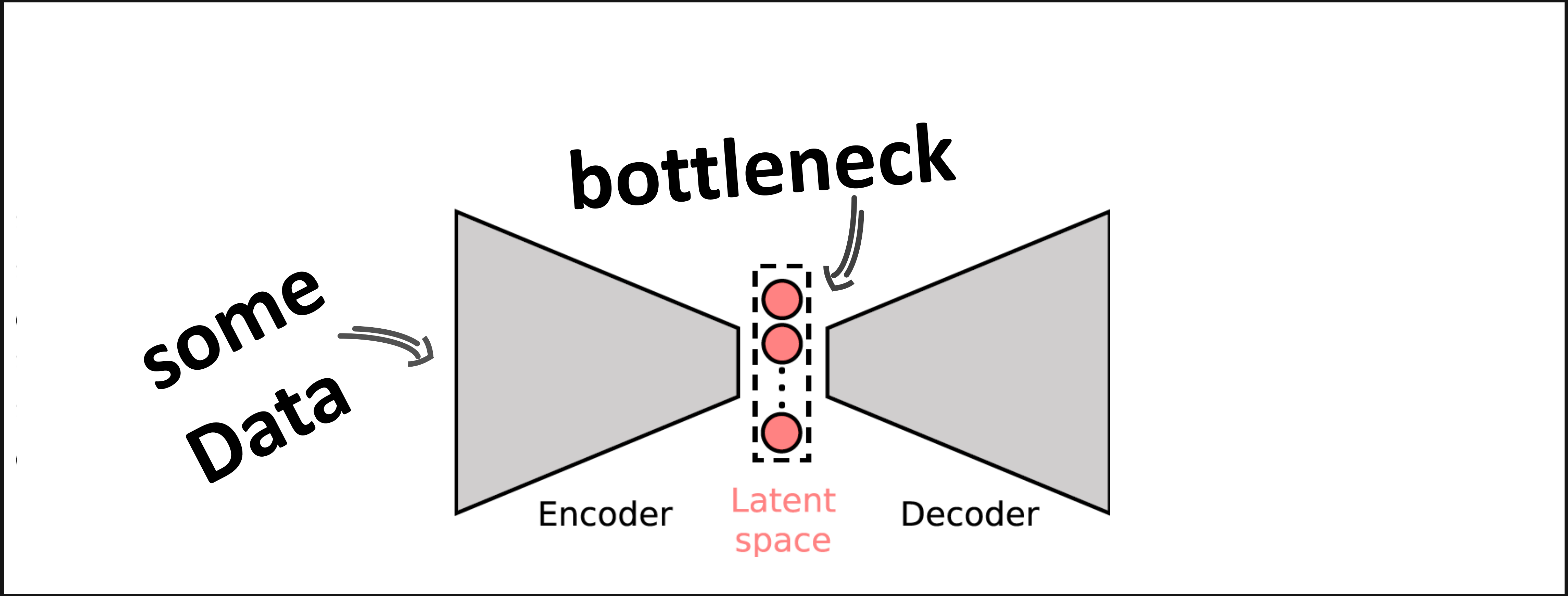
Representation learning



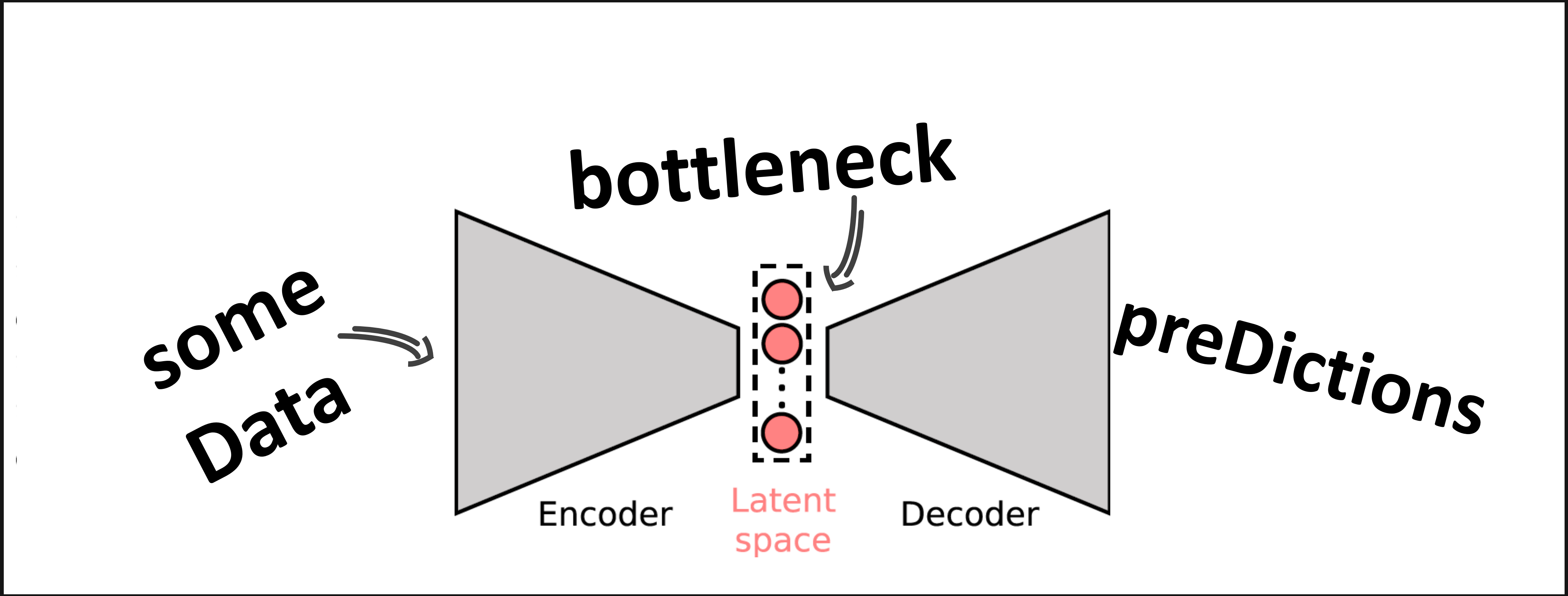
Representation learning



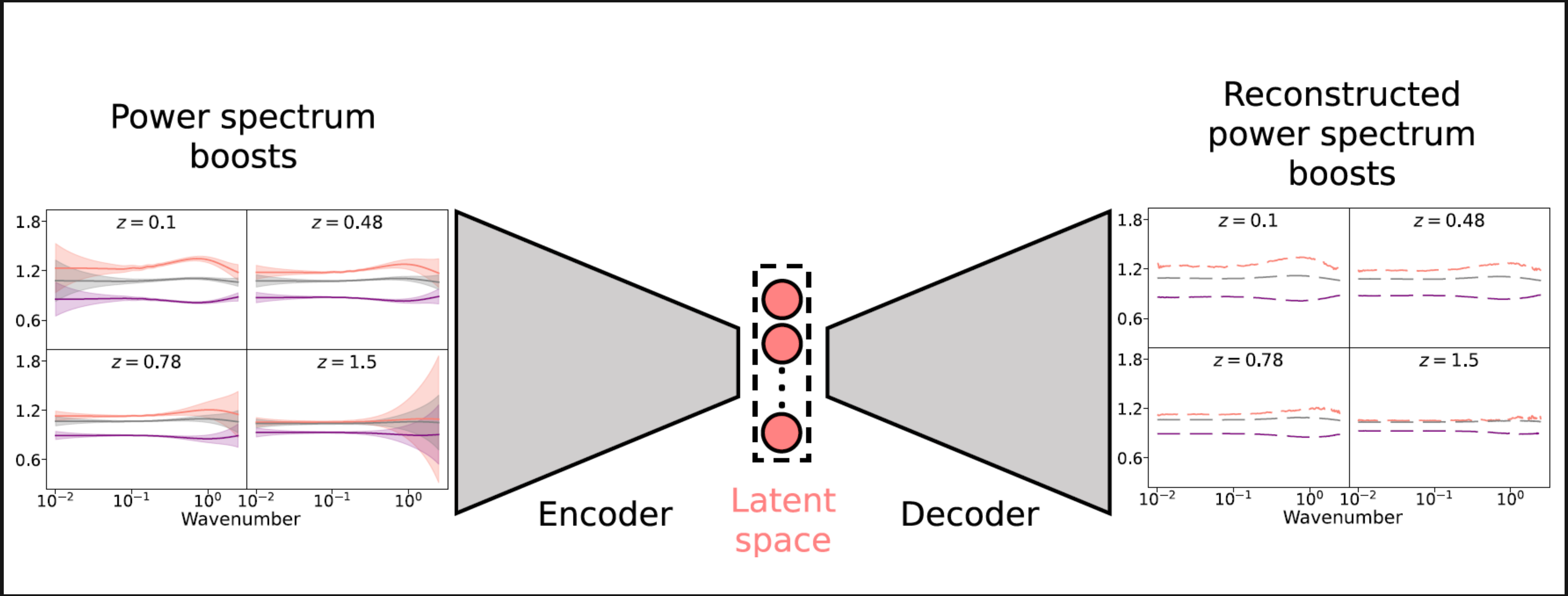
Representation learning



Representation learning



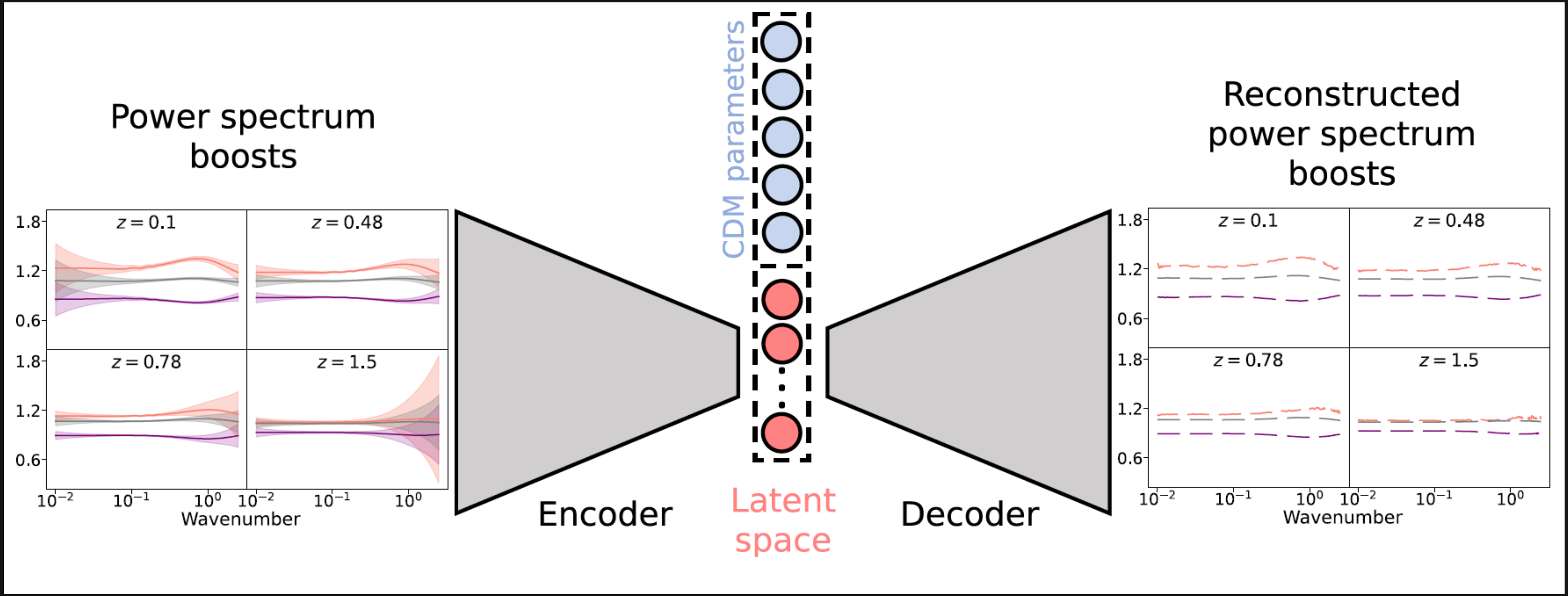
Representation learning



$$\text{Power spectrum boost} = \frac{\text{Power spectrum in extended model}}{\text{Power spectrum in } \Lambda\text{CDM model}}$$

Representation learning

Piras & Lombriser, arXiv 2310.10717



$$\text{Power spectrum boost} = \frac{\text{Power spectrum in extended model}}{\text{Power spectrum in } \Lambda\text{CDM model}}$$

Less is enough:

extending ΛCDM with representation learning

parameters
Less is enough:

extending ΛCDM with representation learning

An application to dark energy

- Apply our framework to single extension: w_0w_aCDM

An application to dark energy

- Apply our framework to single extension: w_0w_aCDM

- Two extra parameters: w_0 and w_a

$$w(a) = w_0 + (1-a)w_a$$

An application to dark energy

- Apply our framework to single extension: wCDM
- Two extra parameters: w_0 and w_a
- Expect two latent variables are needed…?

An application to dark energy

An application to dark energy

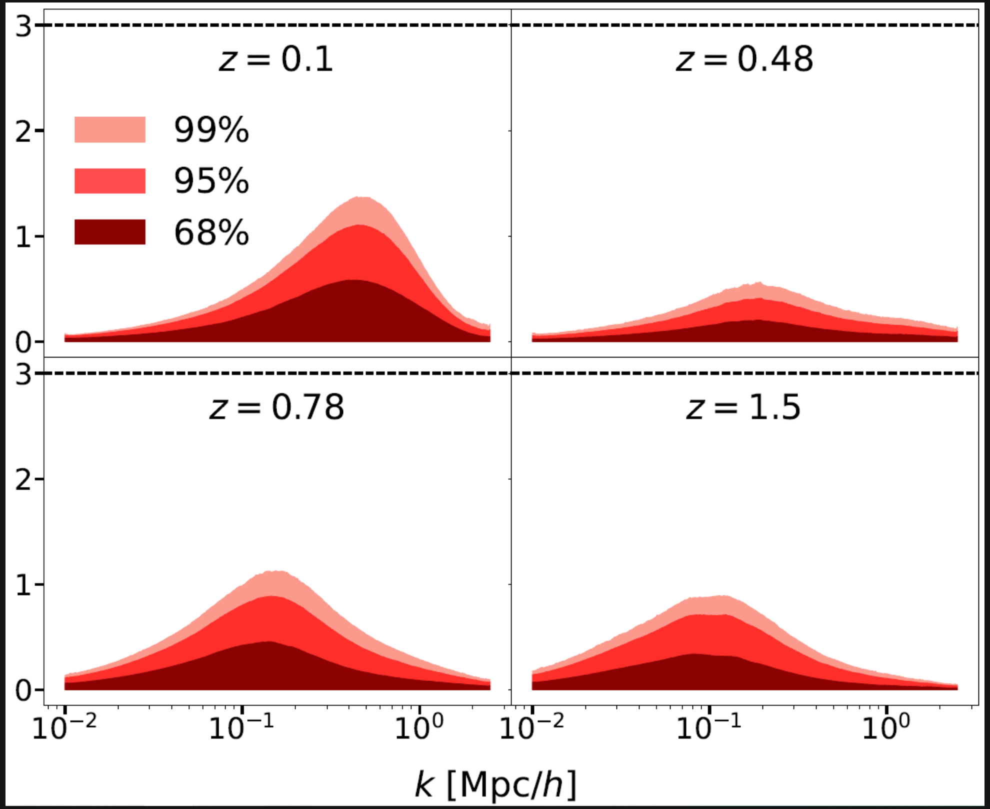
An application to dark energy

An application to dark energy

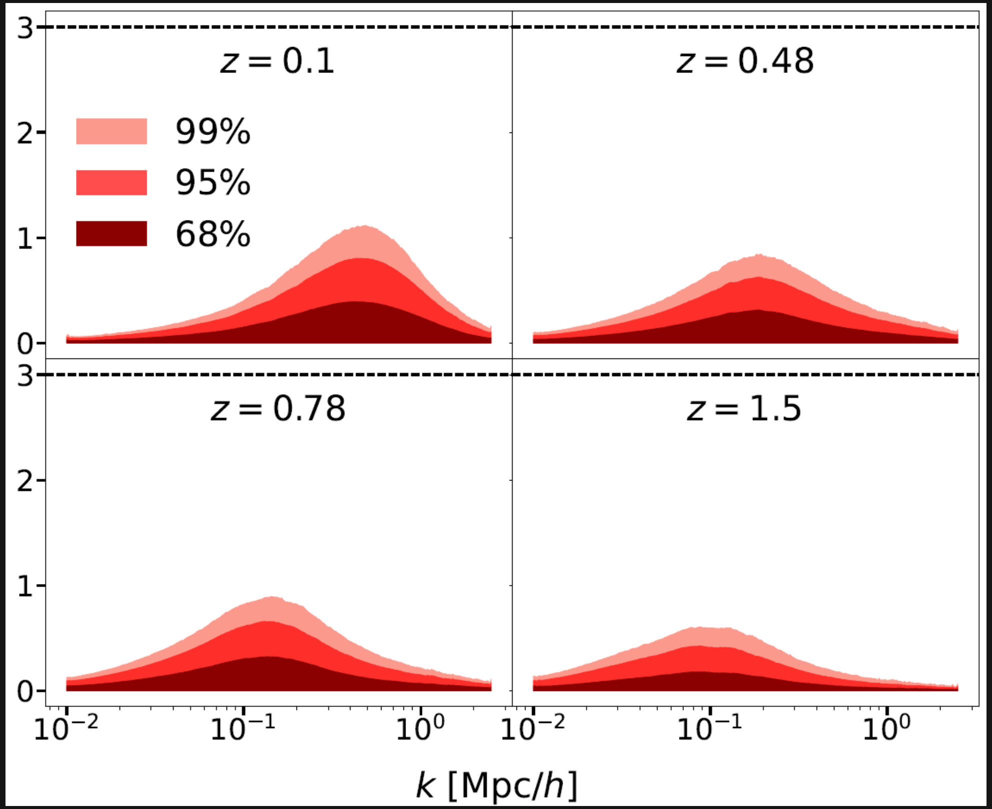
Results

Results

One latent variable

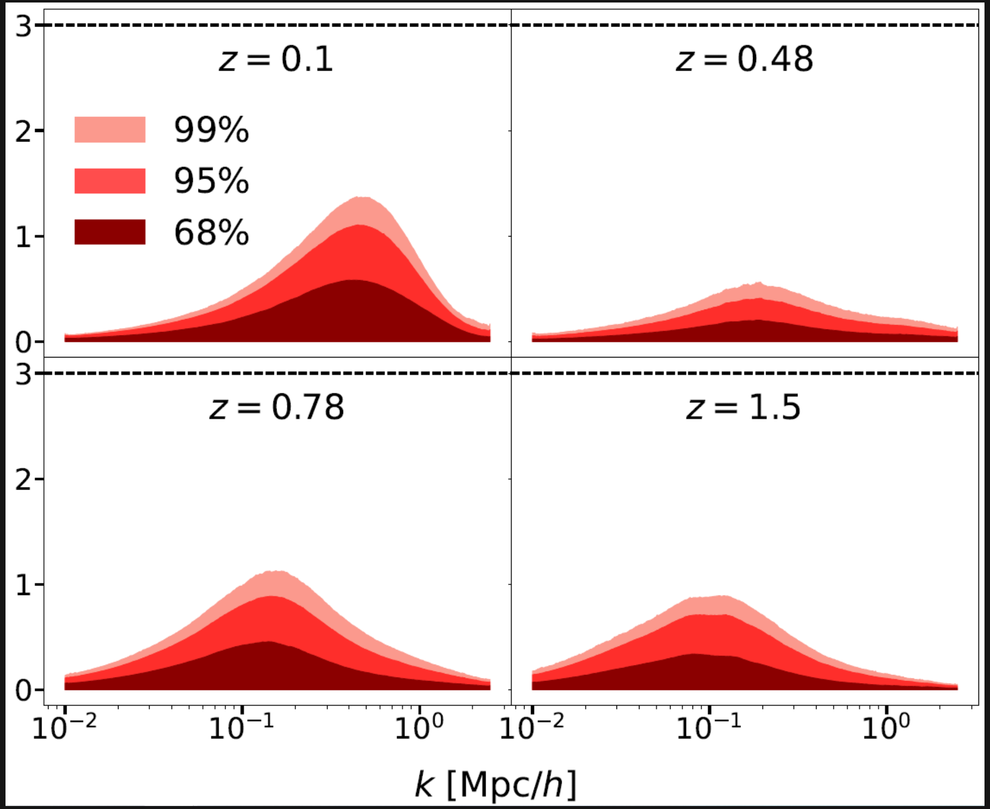


Two latent variables

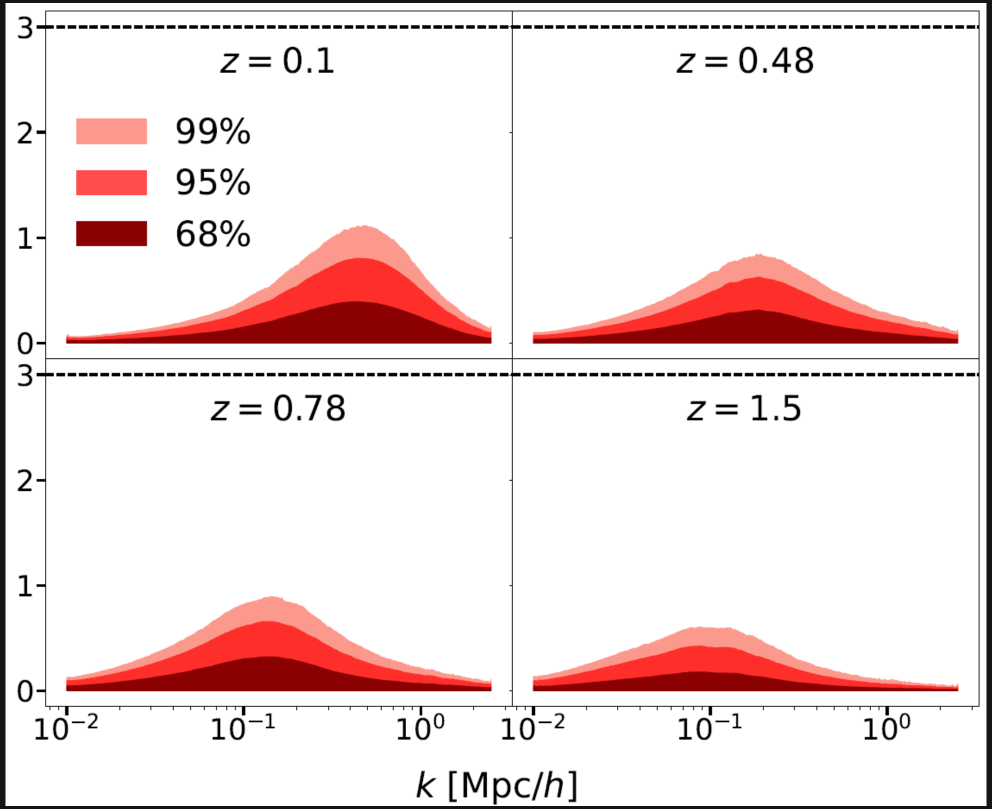


Results

One latent variable



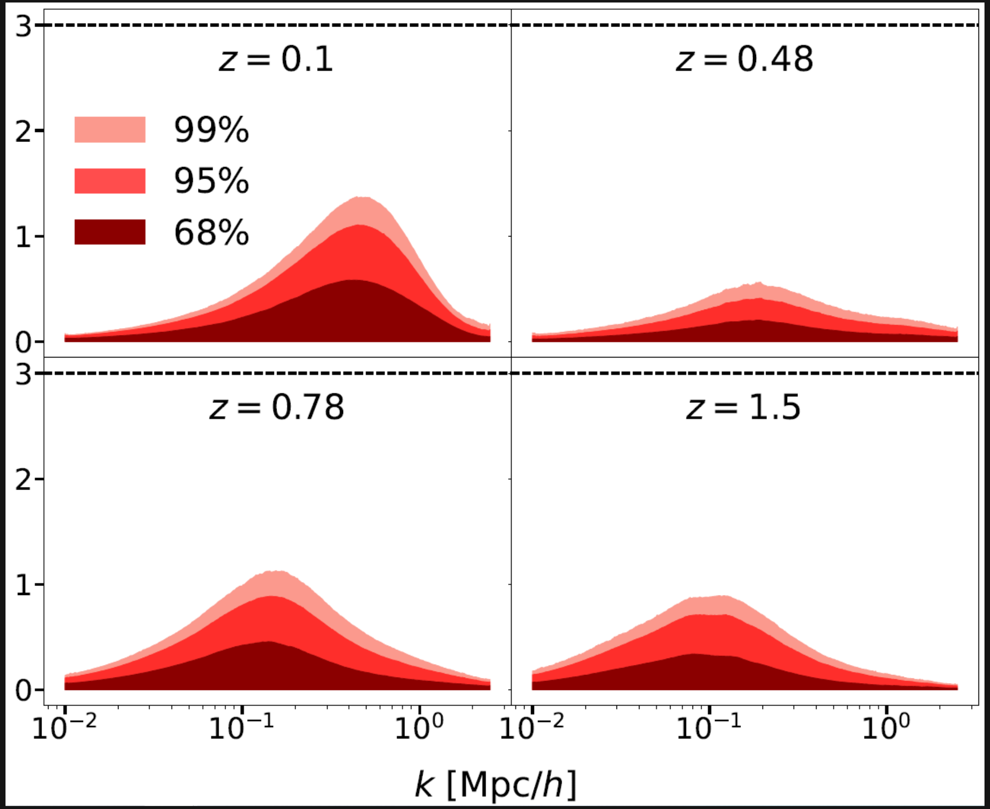
Two latent variables



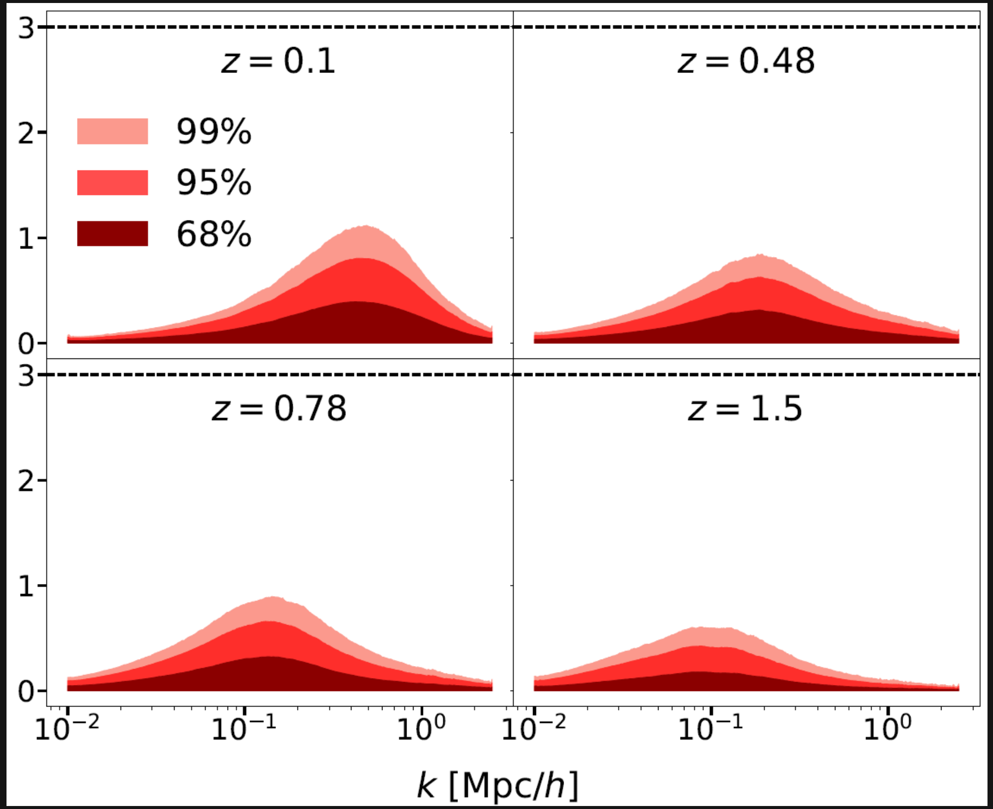
Results are pretty similar with one and two latents

Results

One latent variable



Two latent variables

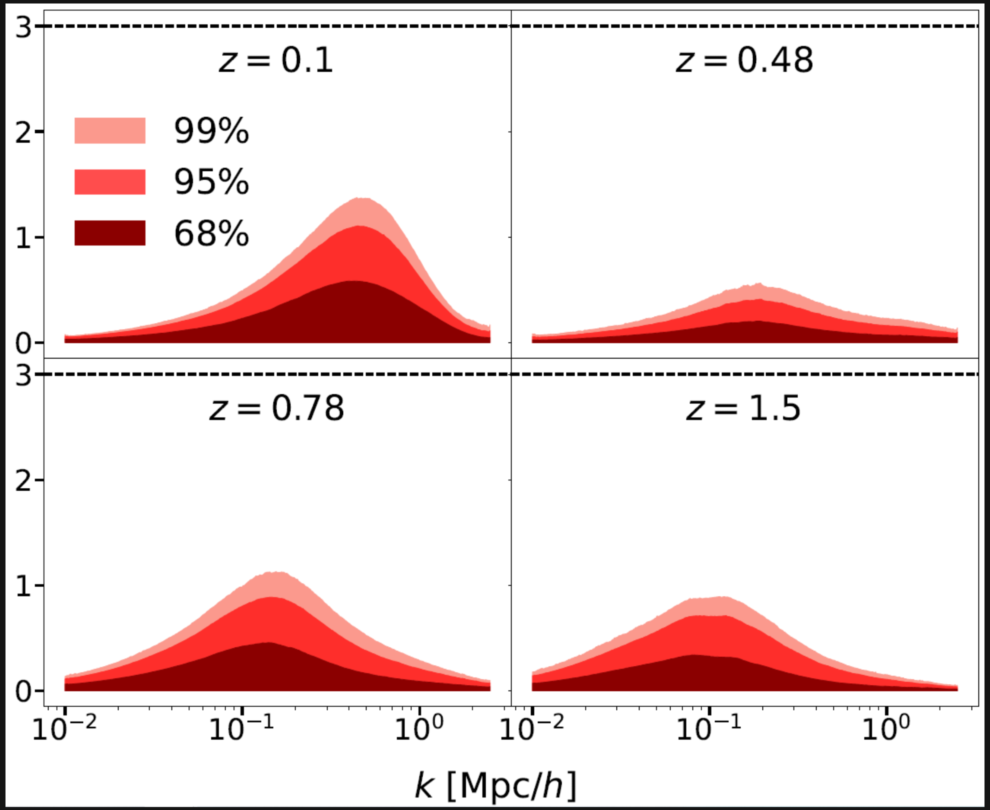


Results are pretty similar with one and two latents

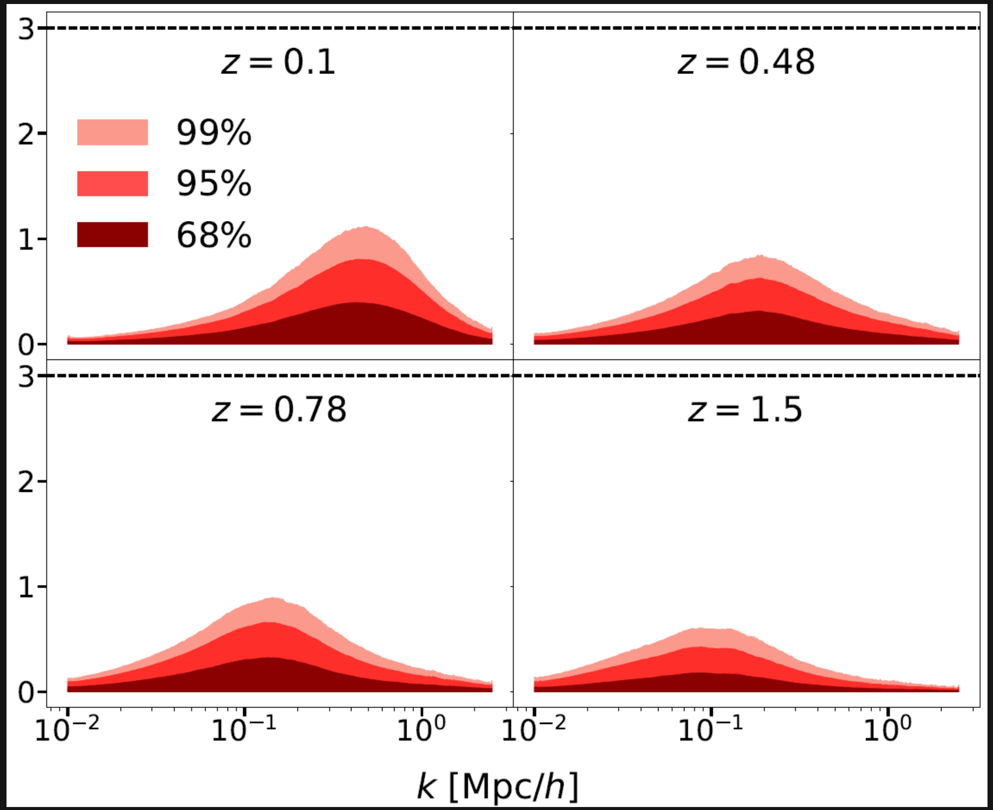
Vertical axis: error in the prediction of the power spectra (lower is better)

Results

One latent variable



Two latent variables



One variable
is enough for w_0w_aCDM!

How to analyse the latent space?

How to analyse the latent space?

- Mutual information

What is mutual information?

- Measures dependence between random variables (more general than Pearson, which measures correlation)

What is mutual information?

- Measures dependence between random variables (more general than Pearson, which measures correlation)

- Well-established in information theory

What is mutual information?

- Measures dependence between random variables (more general than Pearson, which measures correlation)
- Well-established in information theory
- Hard to estimate!

Estimating mutual information (MI)

- No available estimator returns uncertainty on MI

Estimating mutual information (MI)

- No available estimator returns uncertainty on MI

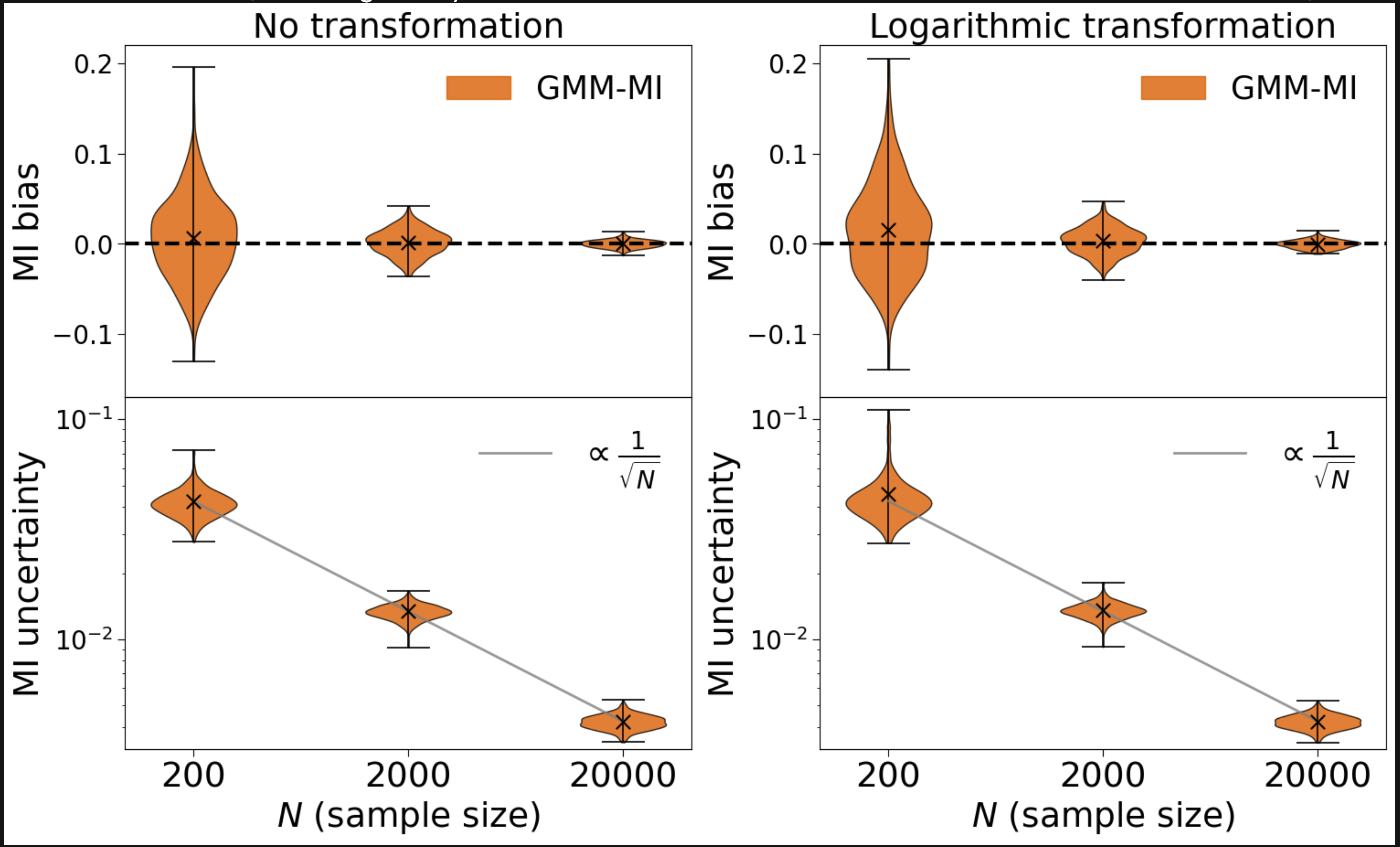
- **Solution:** density estimate with Gaussian mixture model

“Jimmie”

GMM-MI validation

Code

Piras et al. (including Hiranya Peiris, Andrew Pontzen, Luisa Lucie-Smith, Lillian Guo, Brian Nord), MLST



How we use mutual information (MI)

- Calculate MI between latent variables (are they disentangled?)

Latent A $\longleftrightarrow$ Latent B

How we use mutual information (MI)

- Calculate MI between latent variables (are they disentangled?)

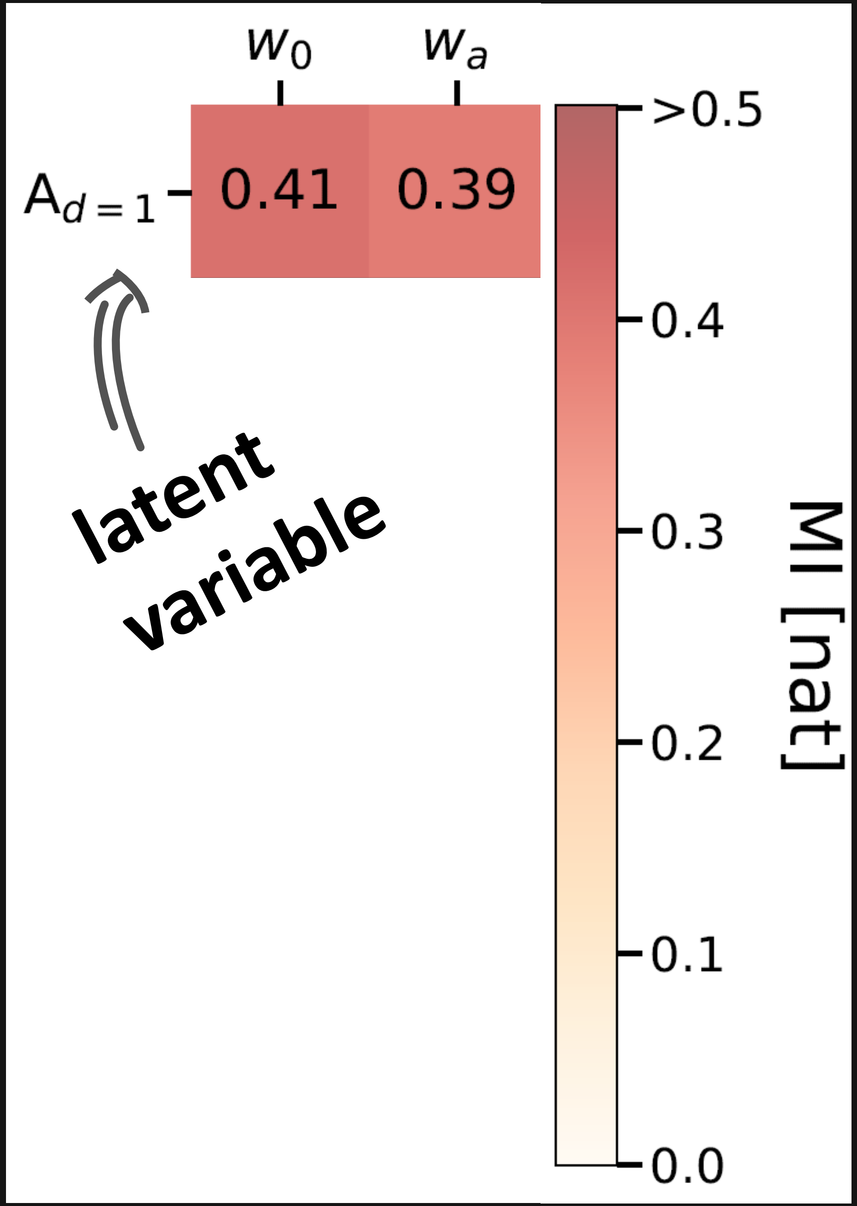
Latent A $\longleftrightarrow$ Latent B

- Calculate MI between a latent variable and model parameters

Latent A $\longleftrightarrow$ w_0, w_a

Mutual information in latent space

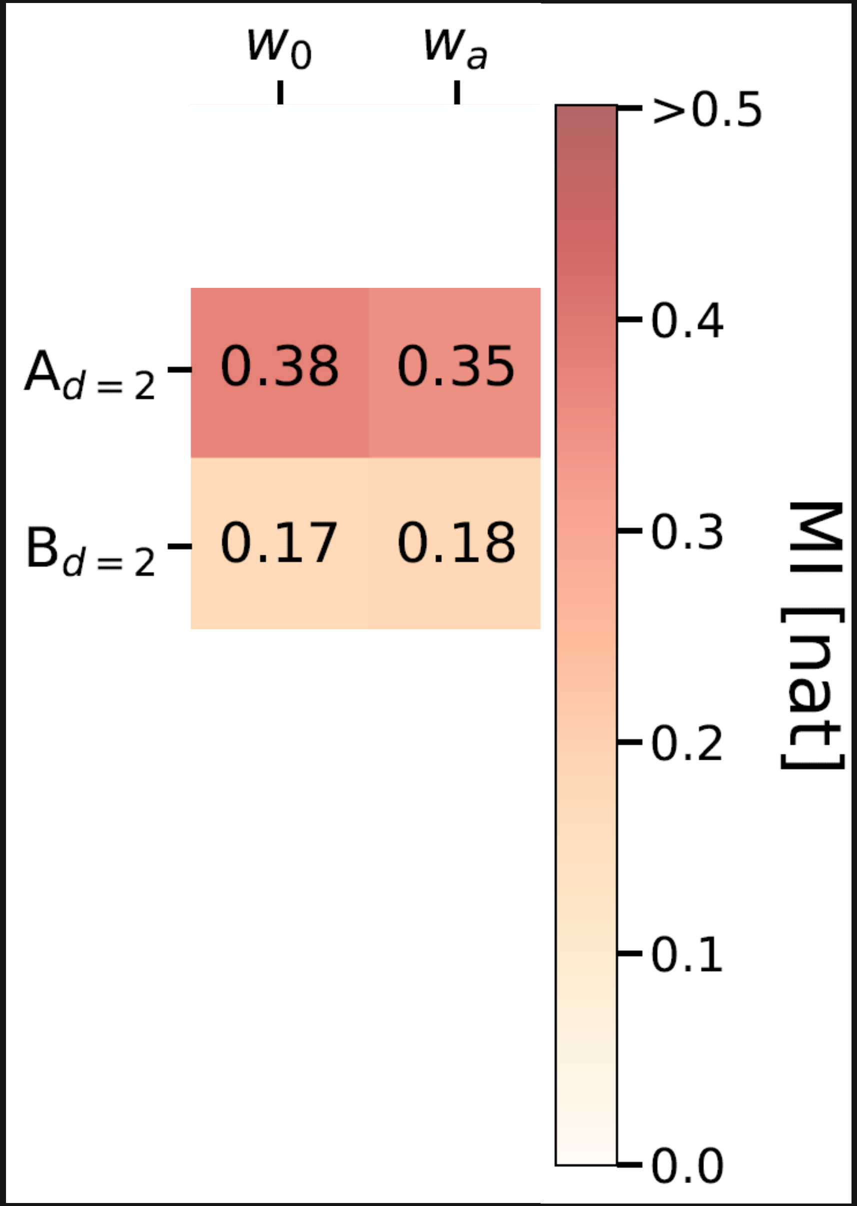
Latent variable has significant MI with w_0 and w_a



Mutual information in latent space

Latent variable has significant MI with w_0 and w_a

Little changes with two latent variables

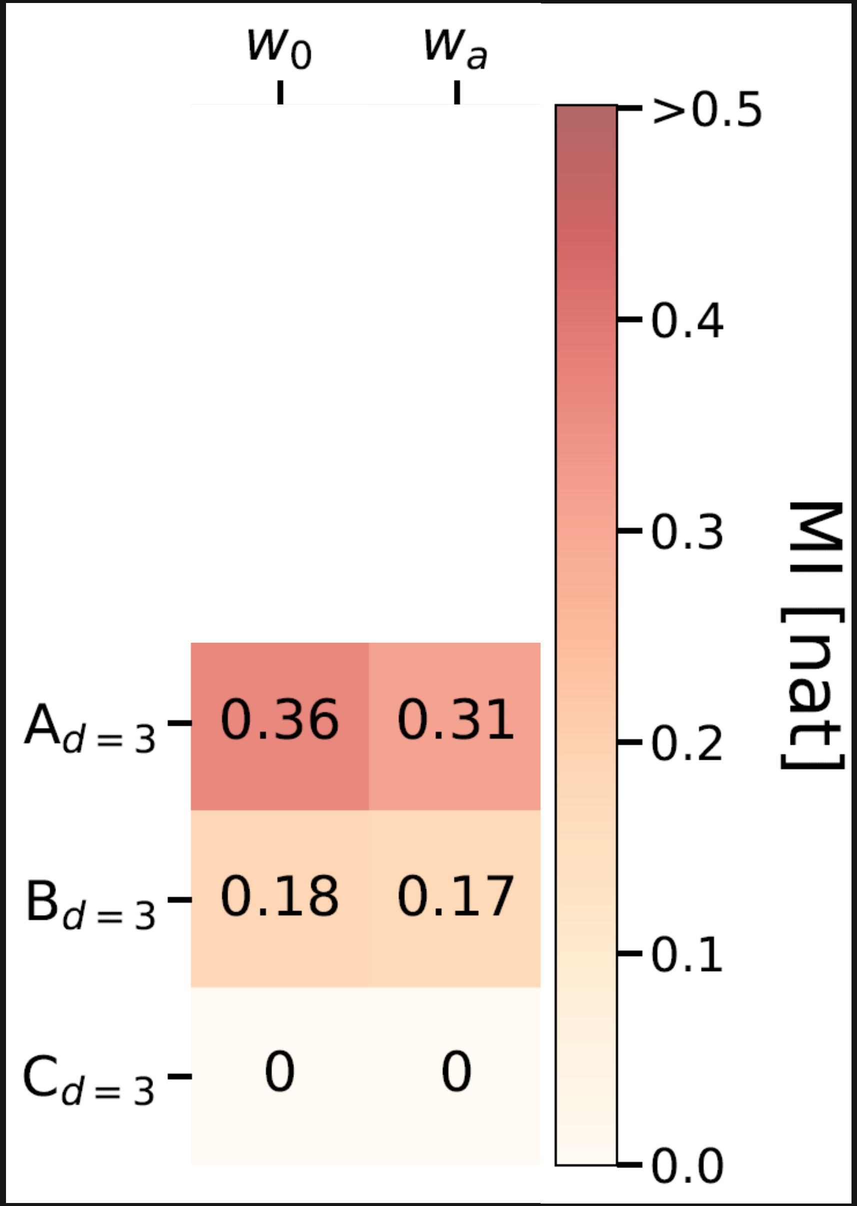


Mutual information in latent space

Latent variable has significant MI with w_0 and w_a

Little changes with two latent variables

Third latent variable is unused



How to analyse the latent space?

- Mutual information
- Symbolic regression

What is symbolic regression?

In a nutshell: find analytic equation between variables

Symbolic regression in latent space

- Link latent variable with w_0 and w_a

Symbolic regression in latent space

- Link latent variable with w_0 and w_a

$$\mathrm{A}_{d=1}(w_0, w_a) = w_0^2 + \frac{e^{w_a + \cos(w_0)}}{w_0}$$

latent variable

Symbolic regression in latent space

- Link latent variable with w_0 and w_a

$$\mathrm{A}_{d=1}(w_0, w_a) = w_0^2 + \frac{e^{w_a + \cos(w_0)}}{w_0}$$

- Analogous to $S_8 = \sigma_8(\Omega_m/0.3)^{0.5}$ …?

Conclusions

- Only need one variable to describe $w_0 w_a$CDM nonlinear matter power spectra

Conclusions

- Only need one variable to describe $w_0 w_a$CDM nonlinear matter power spectra

- Can use mutual information and symbolic regression to interpret latent space

Conclusions

- Only need one variable to describe w_0w_aCDM nonlinear matter power spectra
- Can use mutual information and symbolic regression to interpret latent space
- Will apply our framework to multiple extensions and different summaries

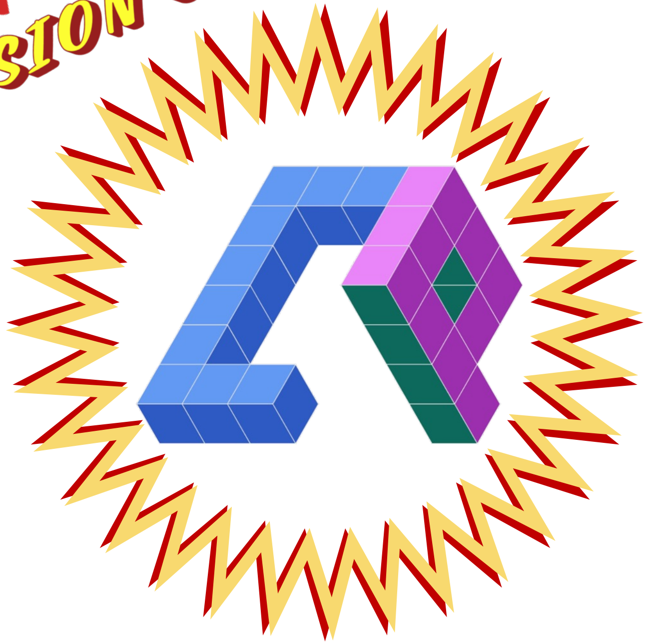
LIKELIHOOD CALL TOO SLOW? • TOO MANY PARAMETERS TO SAMPLE? • RUNNING OUT OF EXCUSES WITH YOUR SUPERVISOR?

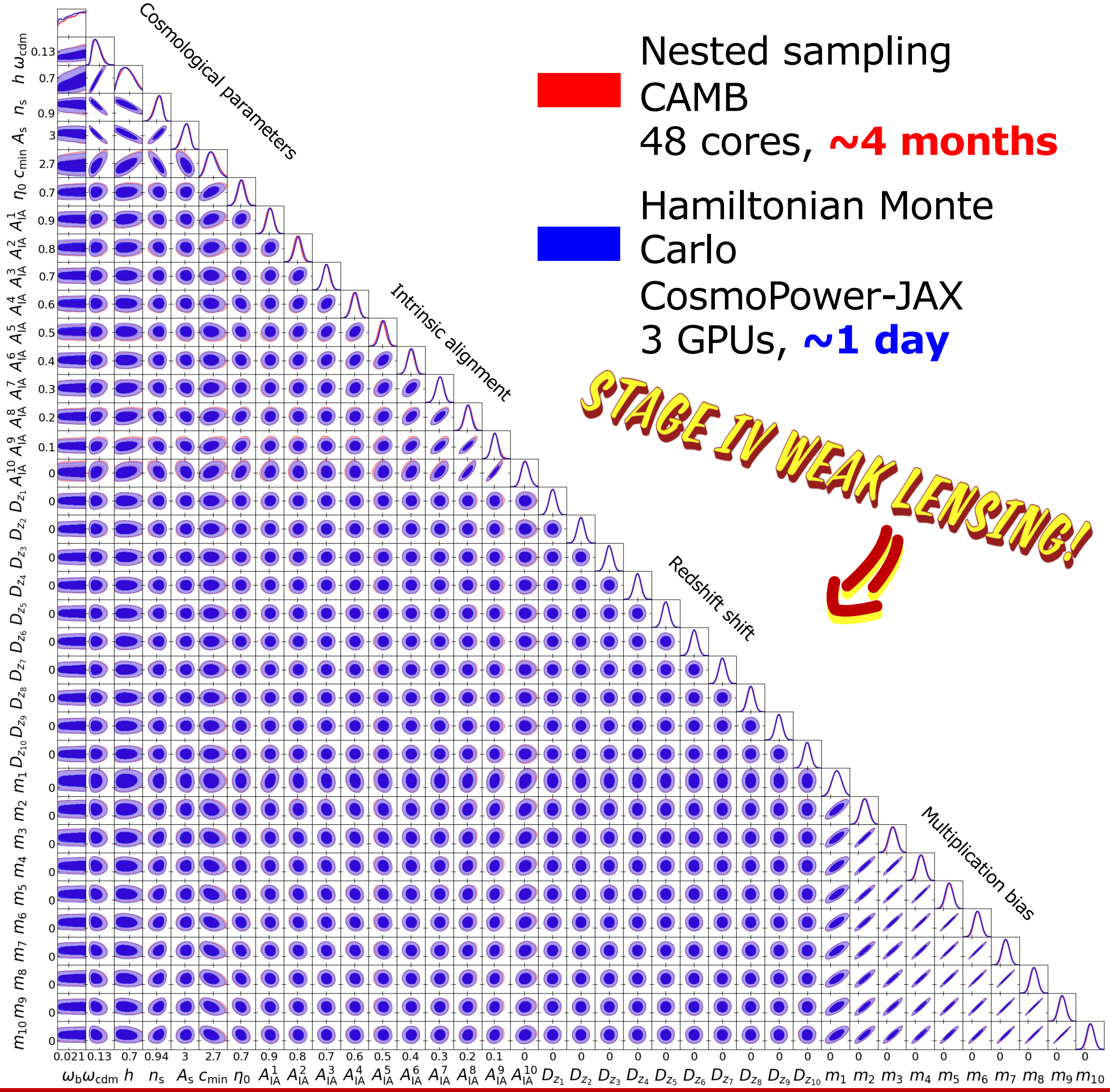

DAVIDE PIRAS
davide.piras@unige.ch

ALESSIO SPURIO MANCINI

Better Use CosmoPower-JAX!

THE JAX VERSION OF COSMOPOWER





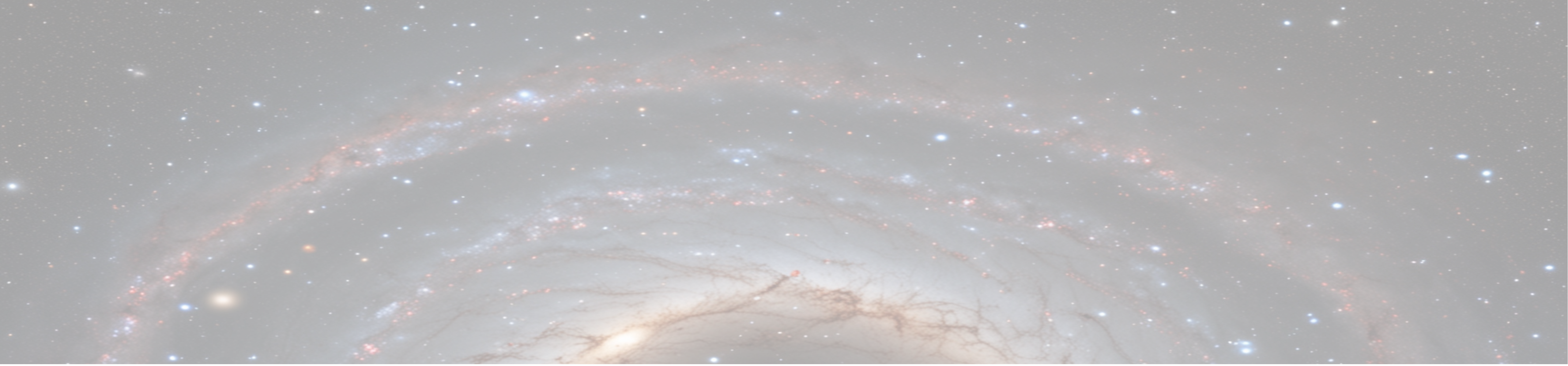
STAGE IV WEAK LENSING!

>1000x SPEED-UP WITH NEURAL EMULATORS • SCALES TO >100 PARAMETERS

WATCH MORE!

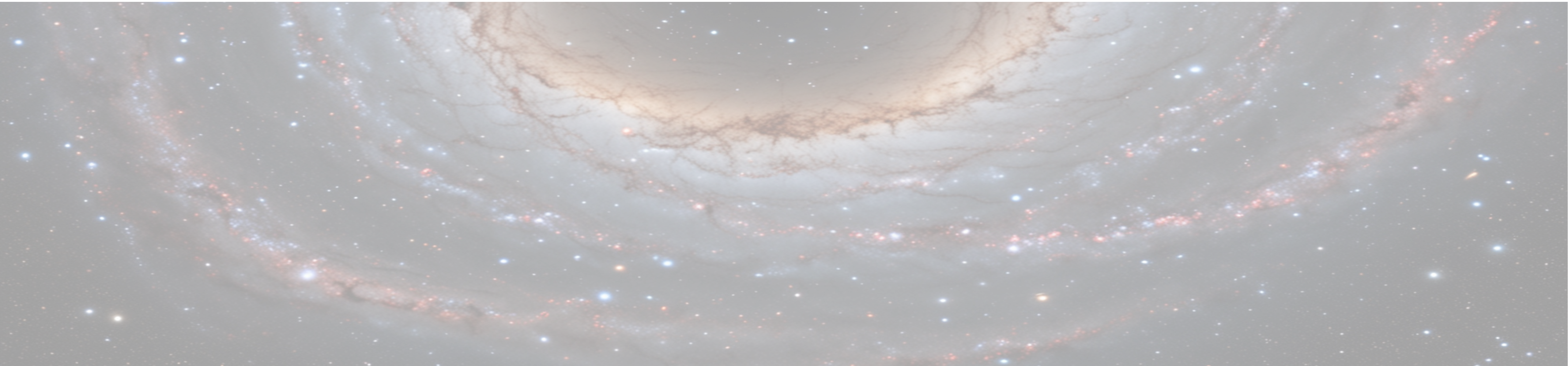
“Speedy Inference For You!”

CODE HERE!



THE FUTURE OF COSMOLOGICAL LIKELIHOOD-BASED INFERENCE: ACCELERATED HIGH-DIMENSIONAL PARAMETER ESTIMATION AND MODEL COMPARISON

Davide Piras[⋆†1,2], Alicja Polanska[†3], Alessio Spurio Mancini[4,3], Matthew A. Price[3], Jason D. McEwen[3,5]



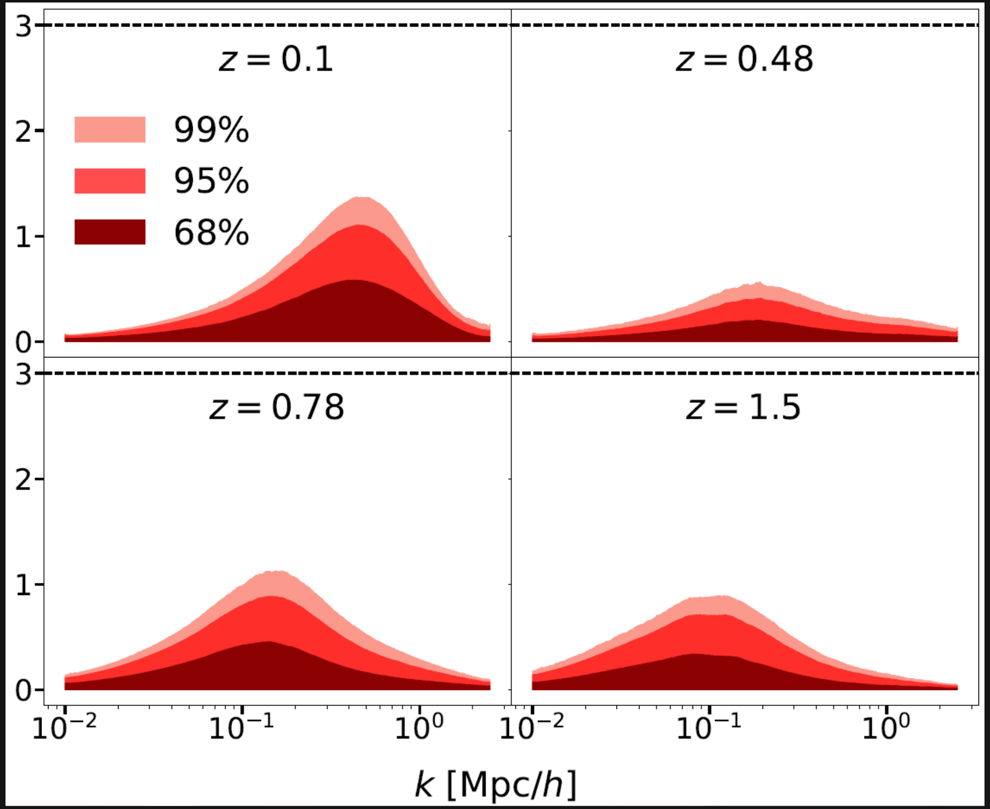
Conclusions

- Only need one variable to describe w_0w_aCDM nonlinear matter power spectra
- Can use mutual information and symbolic regression to interpret latent space
- Will apply our framework to multiple extensions and different summaries

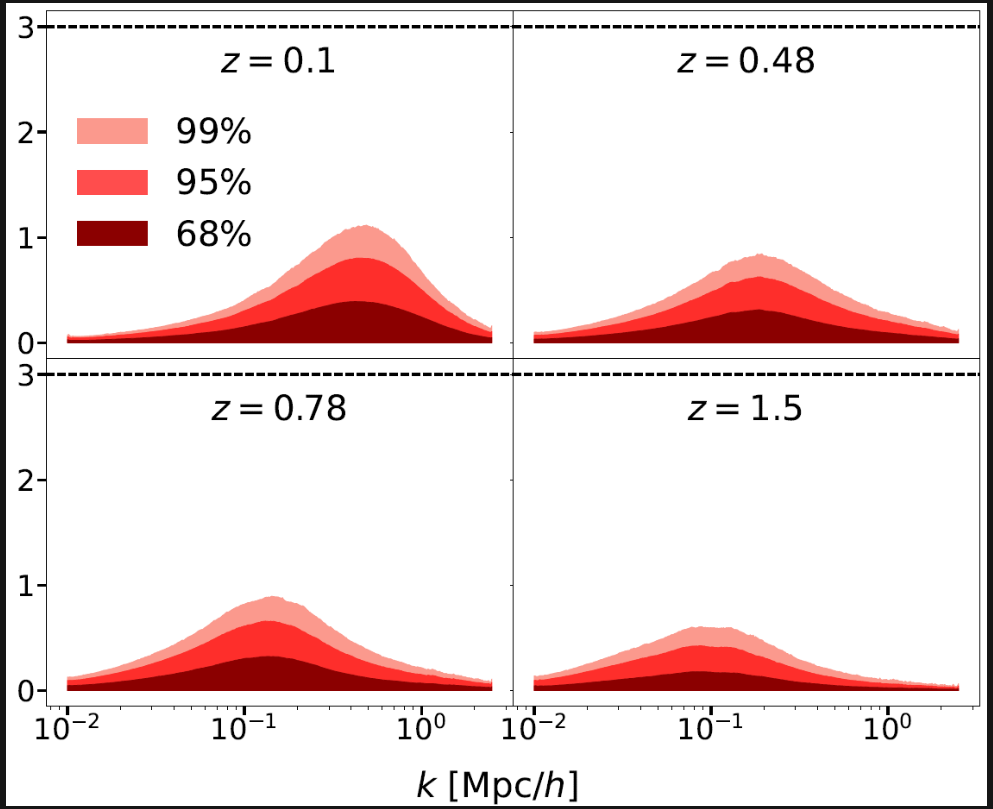
Extra slides
(and memes)

Results

One latent variable

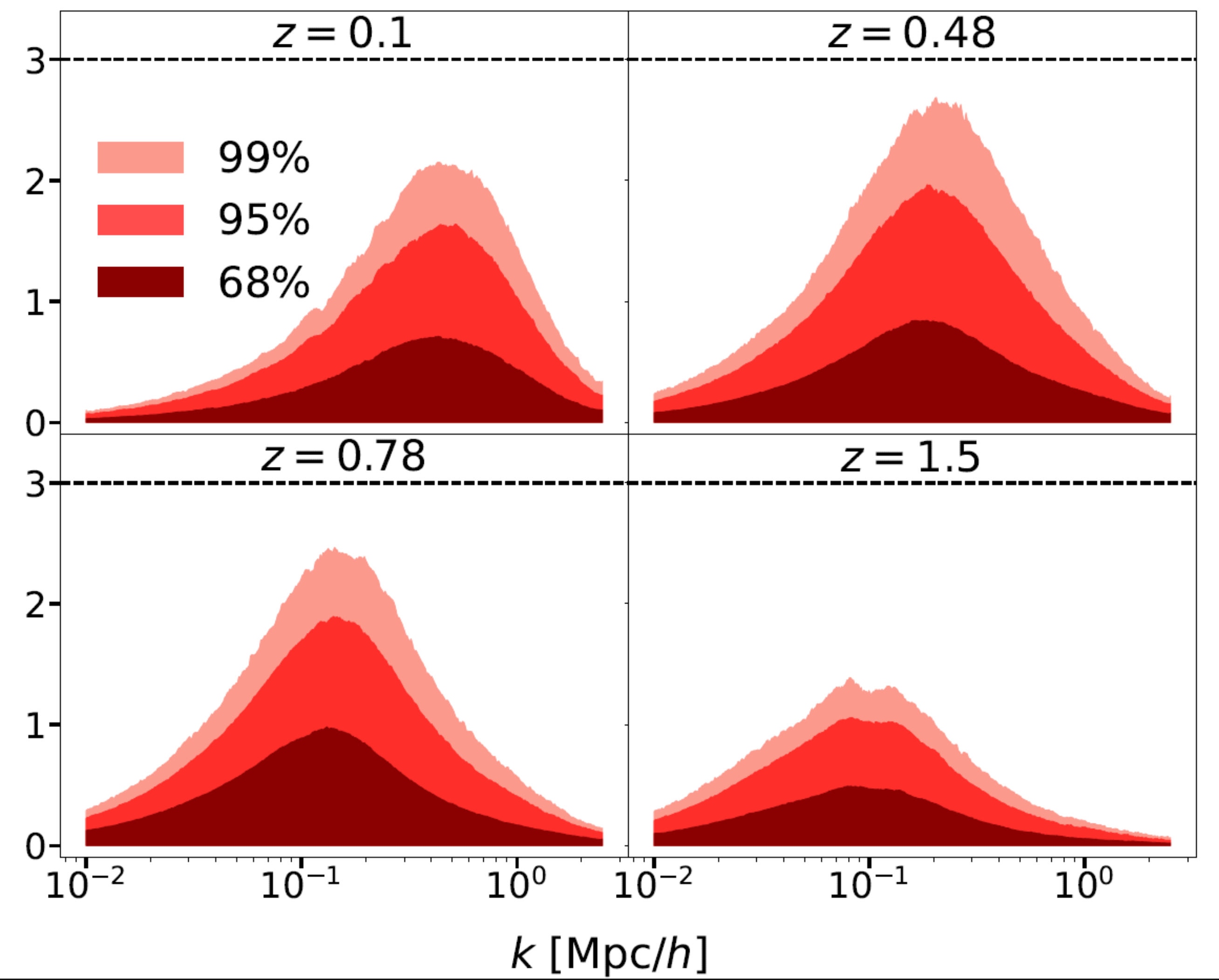


Two latent variables



$$\sigma(k,z) = \sqrt{\frac{4\pi^2}{k^2 \Delta k V(z)} \left(P_{\delta\delta}(k,z) + \frac{1}{\bar{n}(z)} \right)^2 + \sigma_{\rm sys}^2}$$

Symbolic regression results



What is mutual information?

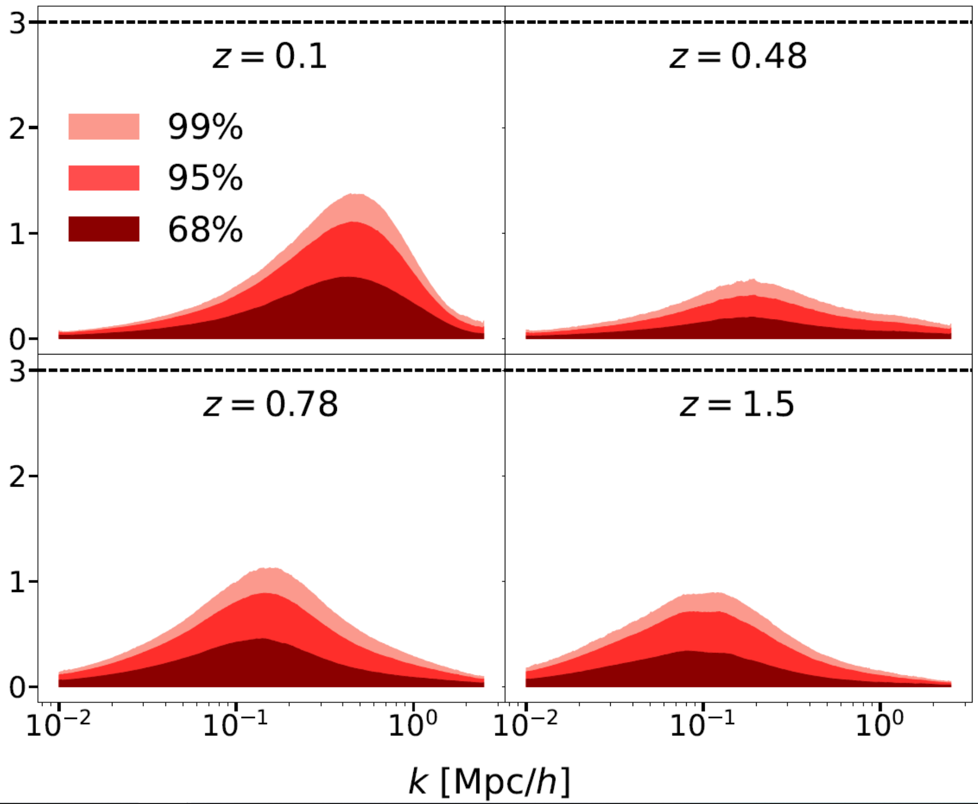
- Measures dependence between random variables (more general than Pearson, which measures correlation)
- Well-established in information theory
- Defined by :

$$\boxed{\mathrm{MI}(X,Y) = \int p(x,y)\log\frac{p(x,y)}{p(x)p(y)}\mathrm{d}x\mathrm{d}y}$$

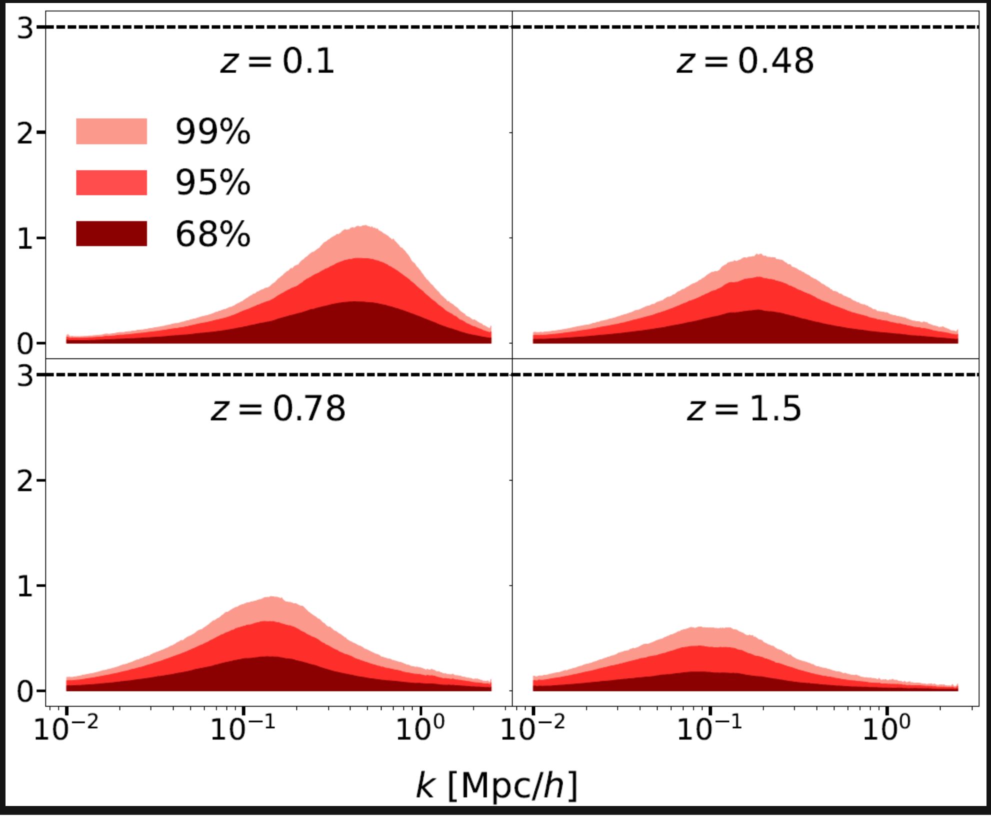
$\to \mathrm{MI}(X,Y) = 0$ if and only if X and Y are independent

Results

One latent variable



Two latent variables




arXiv > cs > arXiv:1506.02640

Computer Science > Computer Vision and Pattern Recognition

[Submitted on 8 Jun 2015 (v1), last revised 9 May 2016 (this version, v5)]

You Only Look Once: Unified, Real-Time Object Detection

Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi

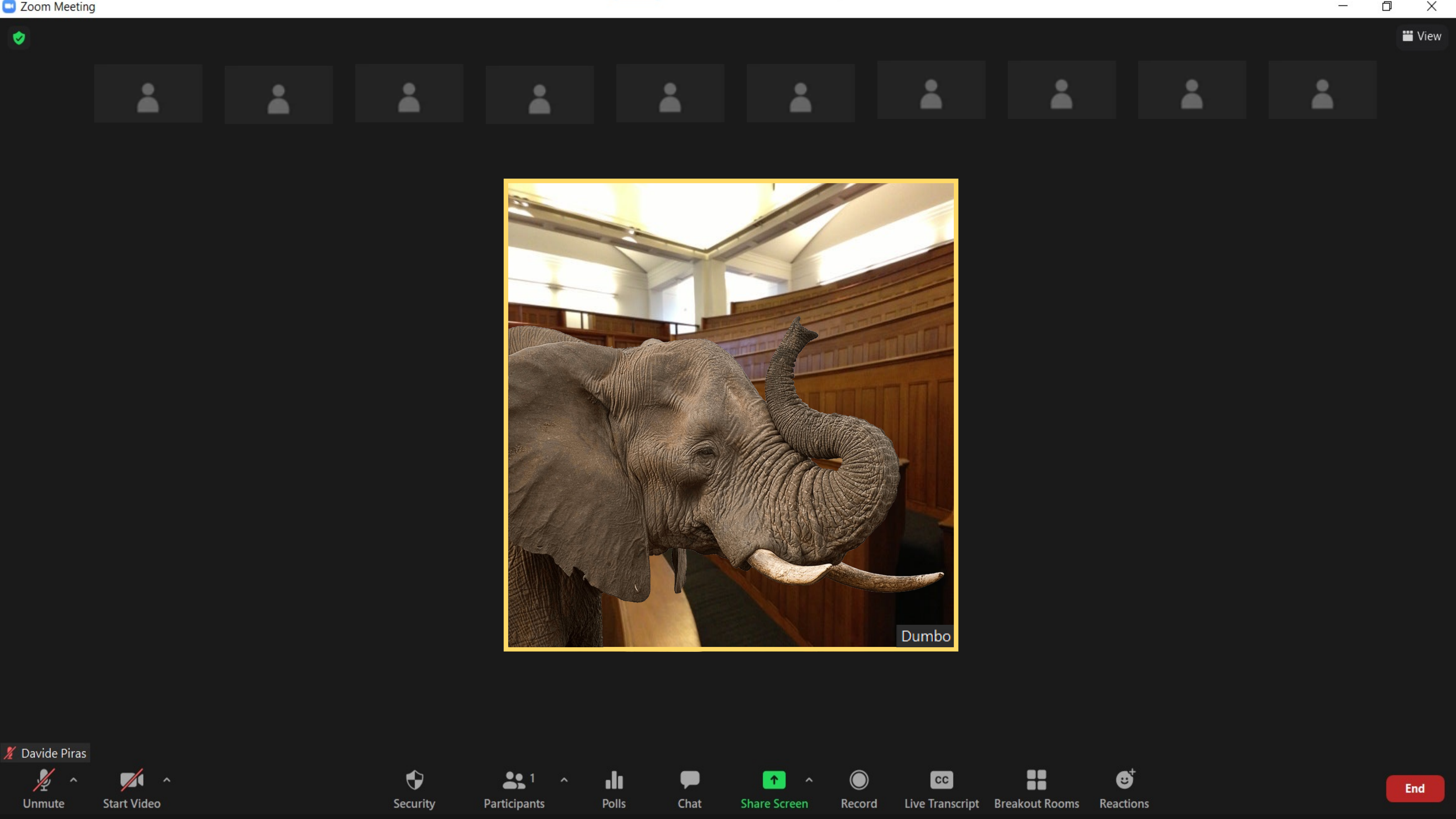
YONOV: You Only Need One Variable

An application to dark energy

- Expect two latent variables are needed…?

An application to dark energy

- Expect two latent variables are needed...?



Zoom Meeting
View
Dumbo
Davide Piras
Unmute
Start Video
Security
Participants 1
Polls
Chat
Share Screen
Record
Live Transcript
Breakout Rooms
Reactions
End

GMM-MI: a robust estimator of mutual information

- Cross-validation and multiple initialisations to optimise fit

GMM-MI: a robust estimator of mutual information

- Cross-validation and multiple initialisations to optimise fit

- Works with continuous and discrete variables

GMM-MI

GMM-MI: a robust estimator of mutual information

- Cross-validation and multiple initialisations to optimise fit
- Works with continuous and discrete variables
- GMM-MI returns **uncertainty** on MI through bootstrapping

GMM-MI at work

```
(gmm_mi) davide@crash:~$
```

GMM-MI at work

```
(gmm_mi) davide@crash:~$ pip install gmm-mi
```

```
In [1]:
   ...:
```

GMM-MI at work

```
(gmm_mi) davide@crash:~$ pip install gmm-mi
```

```
In [1]: import numpy as np
   ...: from gmm_mi.mi import EstimateMI
   ...:

In [2]:
   ...:
   ...:
   ...:
   ...:
   ...:
```

GMM-MI at work

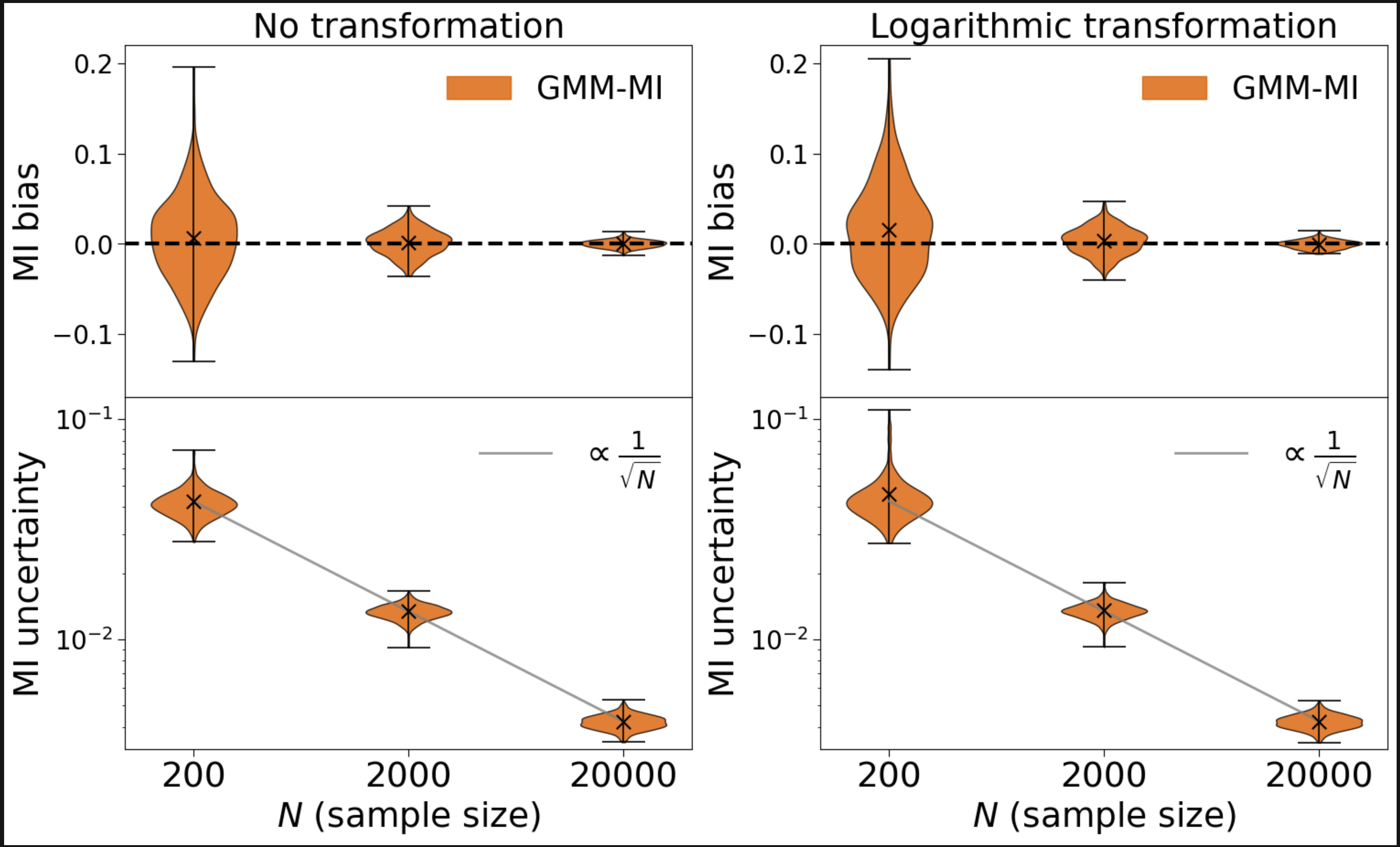
```
(gmm_mi) davide@crash:~$ pip install gmm-mi
```

```
In [1]: import numpy as np
   ...: from gmm_mi.mi import EstimateMI
   ...:

In [2]: # create bivariate Gaussian data
   ...: mean = np.array([0, 0])
   ...: cov = np.array([[1, 0.6], [0.6, 1]])
   ...: rng = np.random.default_rng(0)
   ...: X = rng.multivariate_normal(mean, cov, 200)
   ...:

In [3]:
   ...:
   ...:
```

GMM-MI validation

- GMM-MI is unbiased

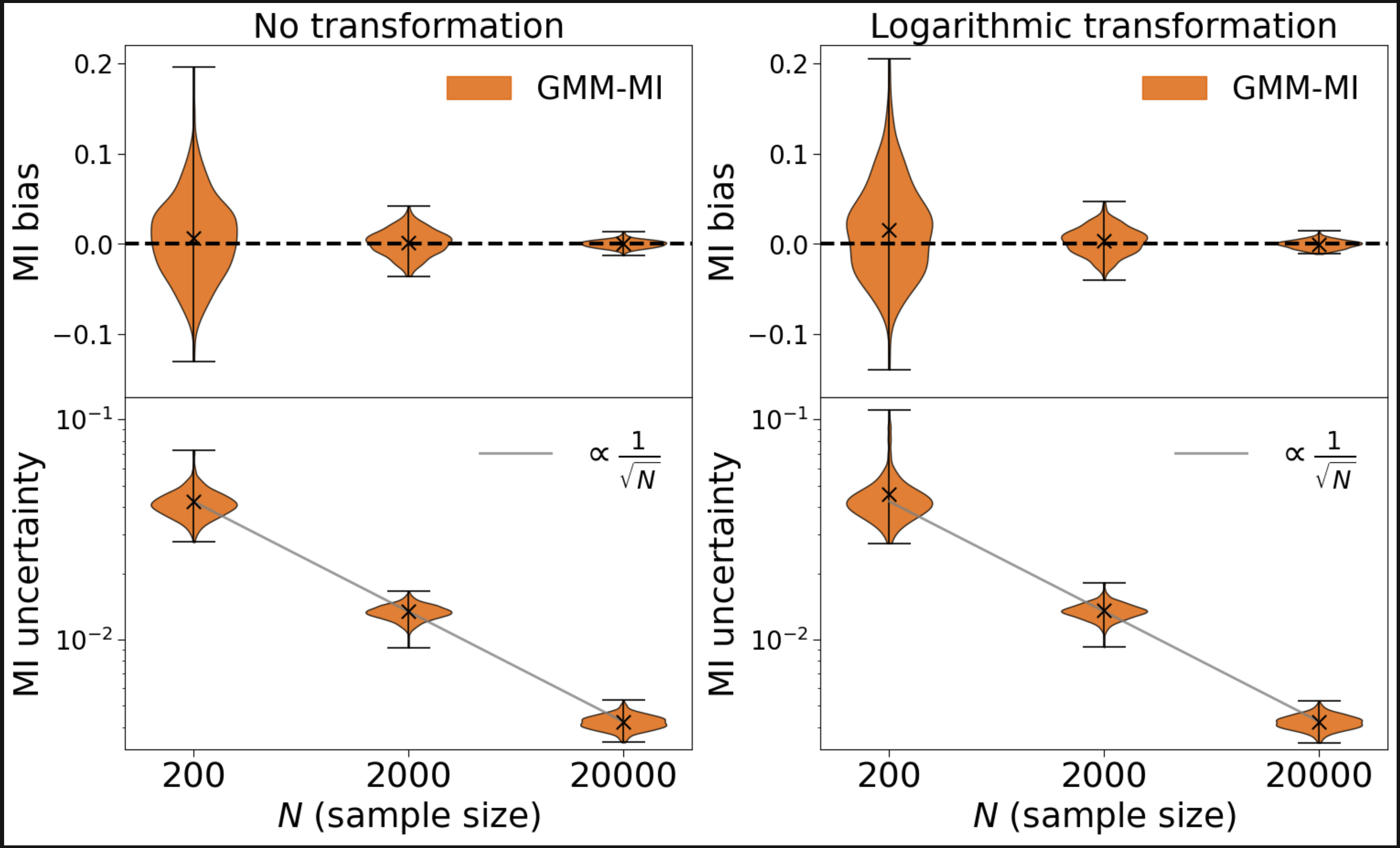
Piras et al., MLST



GMM-MI validation

- GMM-MI is unbiased
- GMM-MI respects MI invariance

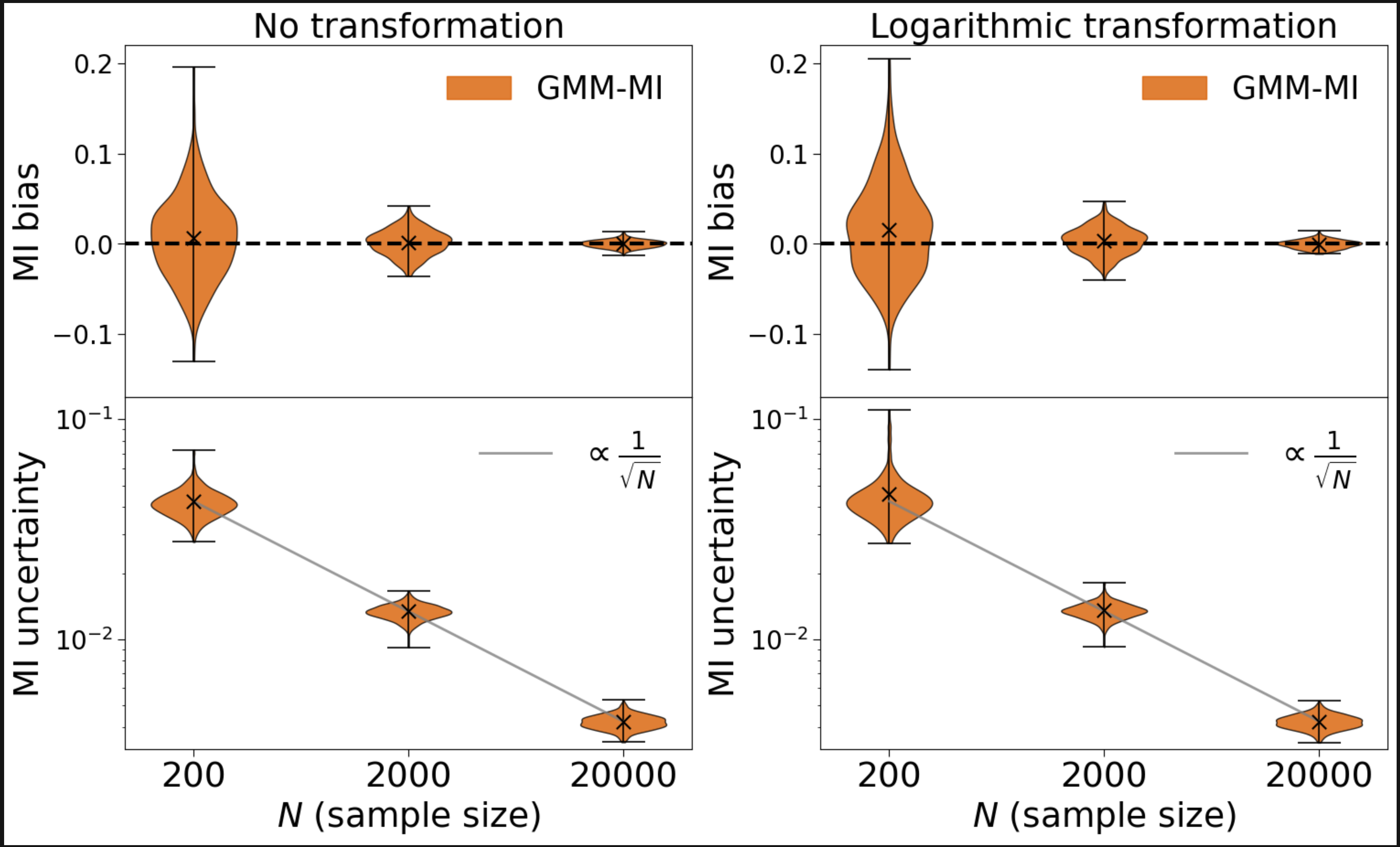
Piras et al., MLST



GMM-MI validation

- GMM-MI is unbiased
- GMM-MI respects MI invariance
- GMM-MI errors scale as expected

Piras et al., MLST



What is symbolic regression?

- Finds analytic equation linking variables

What is symbolic regression?

- Finds analytic equation linking variables
- Less accurate, but more interpretable (?)

What is symbolic regression?

- Finds analytic equation linking variables
- Less accurate, but more interpretable (?)
- Many implementations available