Neural Networks as Classifiers of Cosmological Models

Lisa Goh, Indira Ocampo, Savvas Nesseris, Valeria Pettorino
CosmoStat, CEA Paris-Saclay
lisa.goh@cea.fr

COSMO21, 21 May 2024, Chania, Crète

Context

The concordance ΛCDM model has been widely successful in describing our Universe. However, pertinent questions remain:

- What is the true nature of dark energy?
- How can we explain the Hubble tension ($\sim 5\sigma$ discrepancy between Planck and SH0ES) and the S_8 tension ($\sim 2-3\sigma$ discrepancy between Planck and low-redshift probes)?

Is the ΛCDM model sufficient?

Beyond ΛCDM

Quintessence Models: Coupled Dark Energy

- First hypothesised in the 1990s as a solution to the coincidence problem *(Wetterich 1995, Amendola 2000)*
- Has also shown to be able to relieve the H_0 tension *(Pettorino 2013, Di Valentino et al. Review 2020)*, while still compatible with data
- Still actively studied at the background and perturbation levels, Nbody simulations, spherical collapse models

Coupled Dark Energy

Introducing a coupling between fermionic dark matter (DM) particles and dark energy, assumed to take on the form of a scalar field ϕ

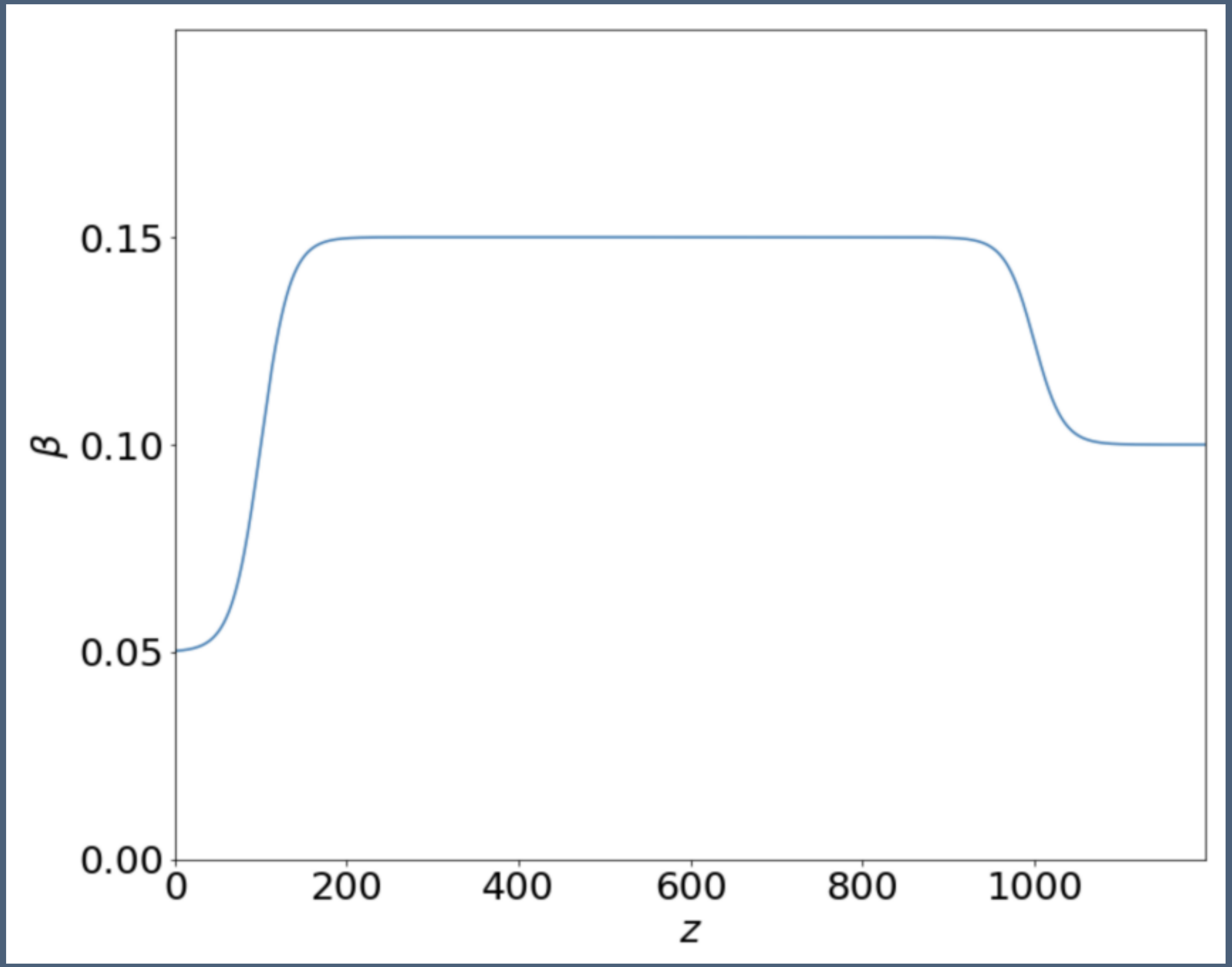
$$\nabla^\mu T^{\phi}_{\mu\nu} = \kappa\beta T^{\mathrm{cdm}} \nabla_\nu \phi \quad ; \quad \nabla^\mu T^{\mathrm{cdm}}_{\mu\nu} = -\kappa\beta T^{\mathrm{cdm}} \nabla_\nu \phi$$

Where the coupling strength parameter β is a function of redshift:

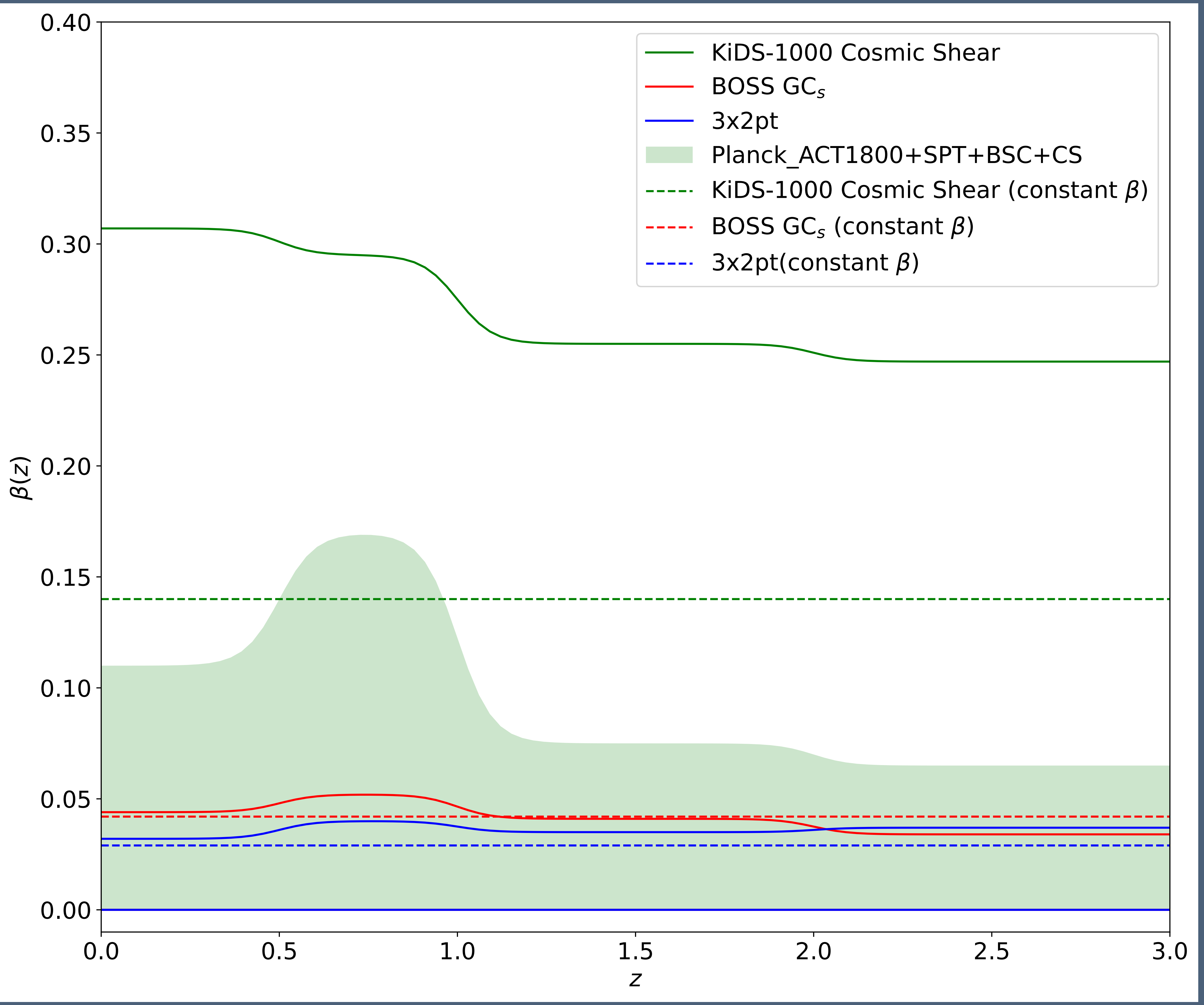
$$\beta(z) = \frac{\beta_1 + \beta_n}{2} + \frac{1}{2}\sum_{i=1}^{n-1} (\beta_{i+1} - \beta_i)\ \tanh[s_i(z - z_i)]$$

Coupled Dark Energy

$$z = \{0, 100, 1000\}$$



Constraining tomographic CDE with 3x2pt probes



Source: Goh et al. 2023

Galaxy Clustering and 3x2pt very effective at constraining coupling strength β at low redshifts!

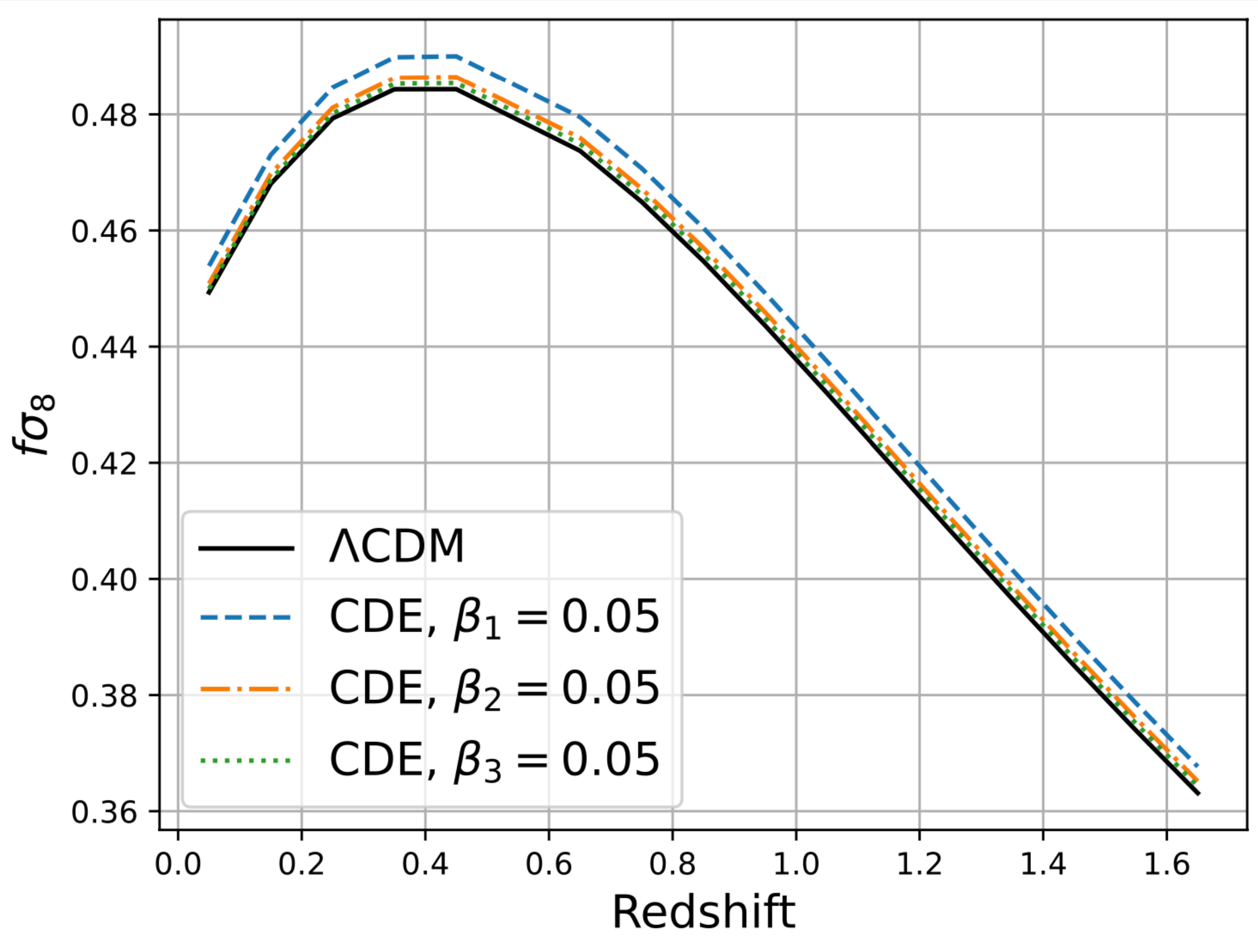
Neural Networks as Model Classifiers

Can we go beyond Bayesian constraints?

- We investigate if a neural network (NN) is able to differentiate between models — ΛCDM and tomographic CDE — based on data
- Generate mock data to train our NN, using it on a sample test set to classify between both models
- Using $f\sigma_8(z)$ and its uncertainty as our observable to probe our cosmological model

Methodology

Generating mock $f\sigma_8(z)$



- Generate values of $f\sigma_8(z)$ for 16 redshift bins, using the modified CLASS code of Goh et al. 2023 (https://github.com/LisaGoh/CDE)
- ΛCDM: $\omega_m = [0.01, 0.7]$
  CDE: $\omega_m = [0.01, 0.7]$, $\beta_i = [0.001, 0.5]$
- Assume a 3-bin parameterisation with $z = \{0, 100, 1000\}$

Methodology

Simulating Stage IV survey-like specifications

z	$\frac{dN_{BGS}}{dz\ \mathrm{ddeg}^2}$
0.05	1165
0.15	3074
0.25	1909
0.35	732
0.45	120

z	$\frac{dN_{ELG}}{dz\ \mathrm{ddeg}^2}$	$\frac{dN_{LRG}}{dz\ \mathrm{ddeg}^2}$
0.65	309	832
0.75	2269	986
0.85	1923	662
0.95	2094	272
1.05	1441	51
1.15	1353	17
1.25	1337	0
1.35	523	0
1.45	466	0
1.55	329	0
1.65	126	0
1.75	0	0
1.85	0	0

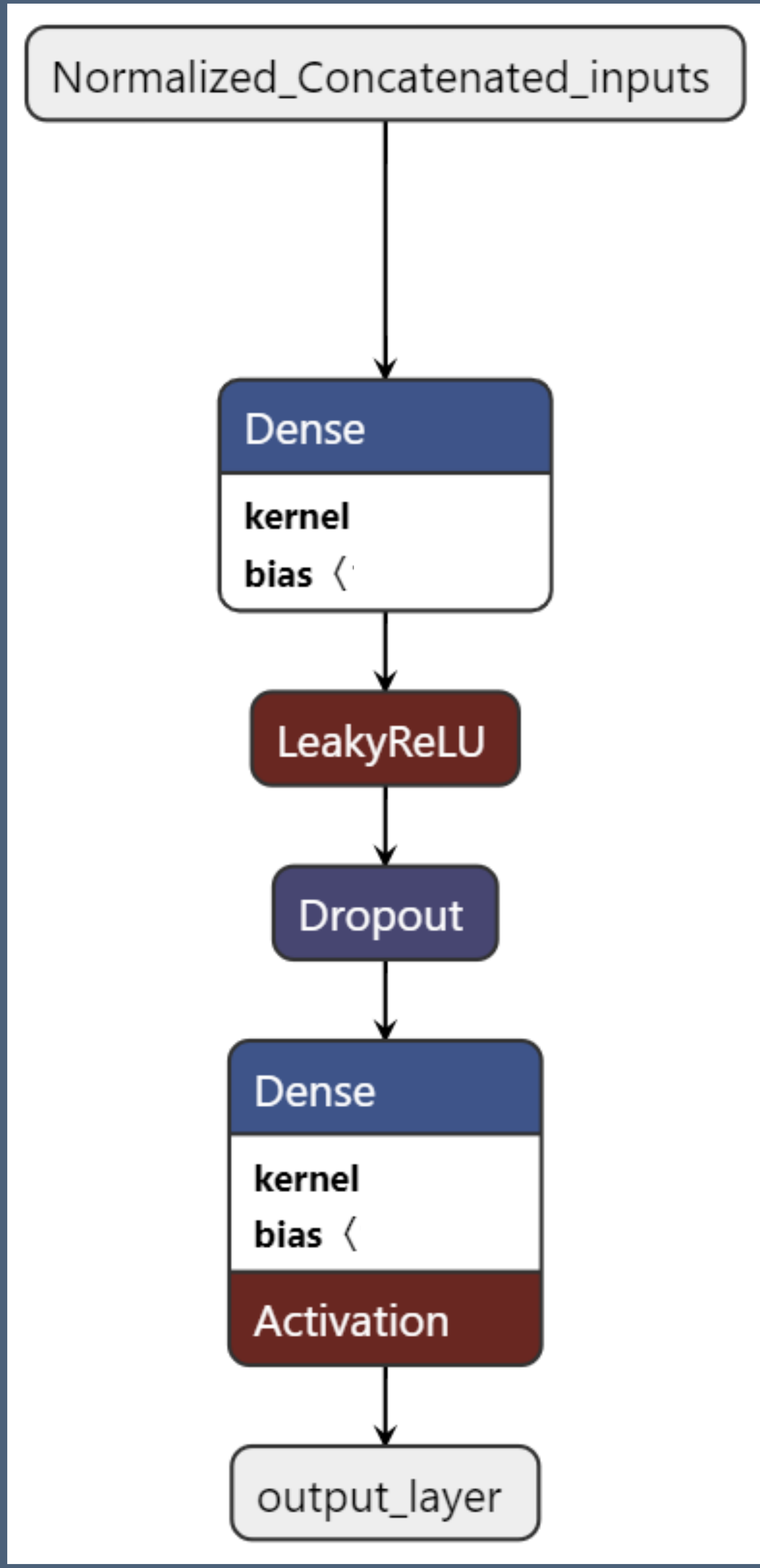
Source: DESI Collaboration et al. 2016

- Following the redshift distribution, galaxy bias and number densities from DESI Bright Galaxy Survey (BGS), Emission Line Galaxies (ELG) and Luminous Red Galaxies (LRG) (DESI Collaboration et al. 2016)
- Adding noise: Calculate the Fisher matrix to obtain the covariance matrix C_{ij} of $f\sigma_8(z)$ between redshift bins

Training the NN

Tuning NN Hyperparameters with Optuna

- We use Optuna to optimise the NN architecture: varying the number of hidden layers, nodes, type of optimiser and dropout rate
- Minimise the loss function
- Use early stopping to prevent overfitting of the data



Training the NN

Tuning NN Hyperparameters with Optuna

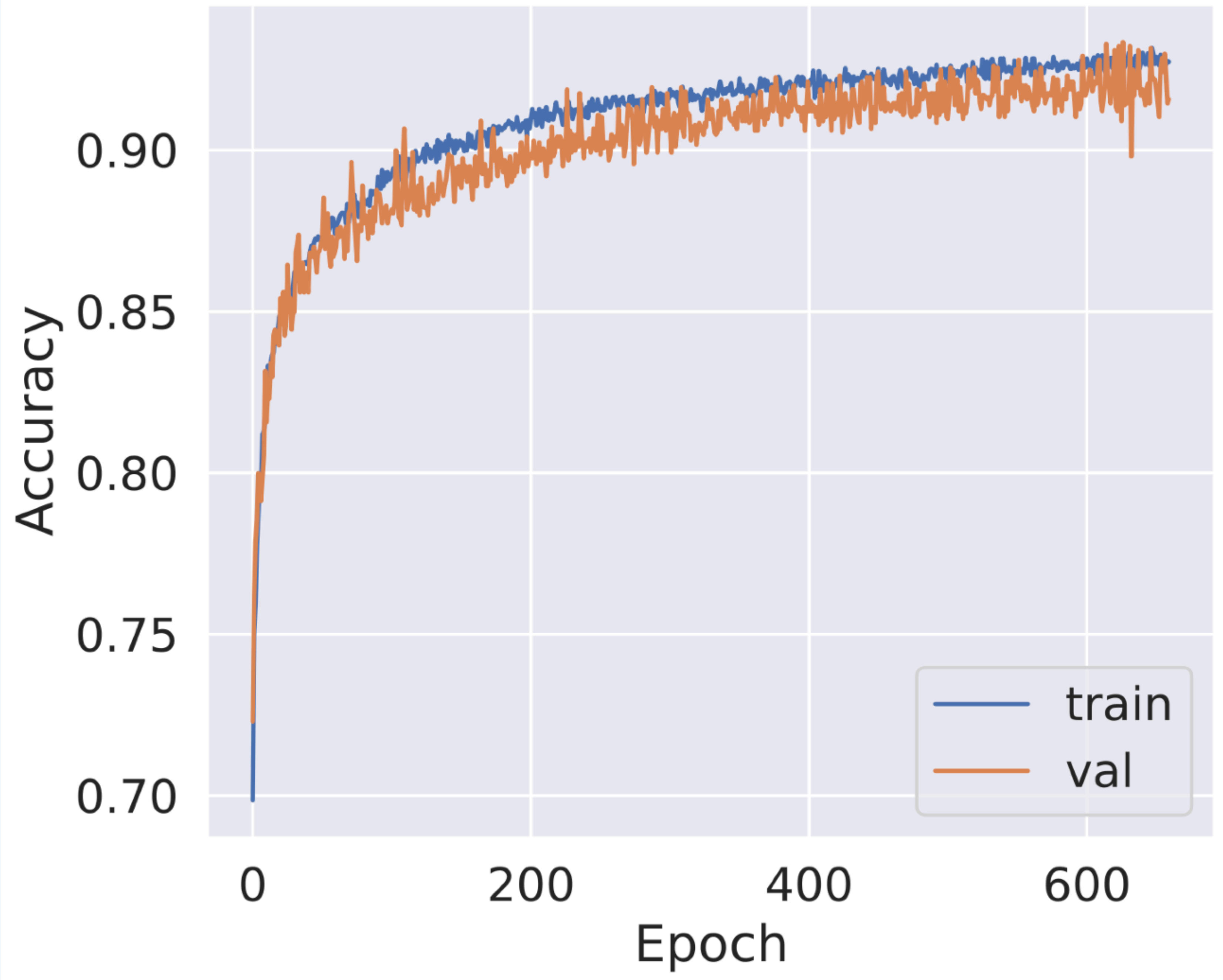
- We use Optuna to optimise the NN architecture: varying the number of hidden layers, nodes, type of optimiser and dropout rate
- Minimise the loss function
- Use early stopping to prevent overfitting of the data

Model	Hyperparameter			
	Hidden layers	Nodes	Dropout rate	Training epochs
β_1	1	38	0.224	660
β_2	1	116	0.218	683
β_3	1	82	0.215	673

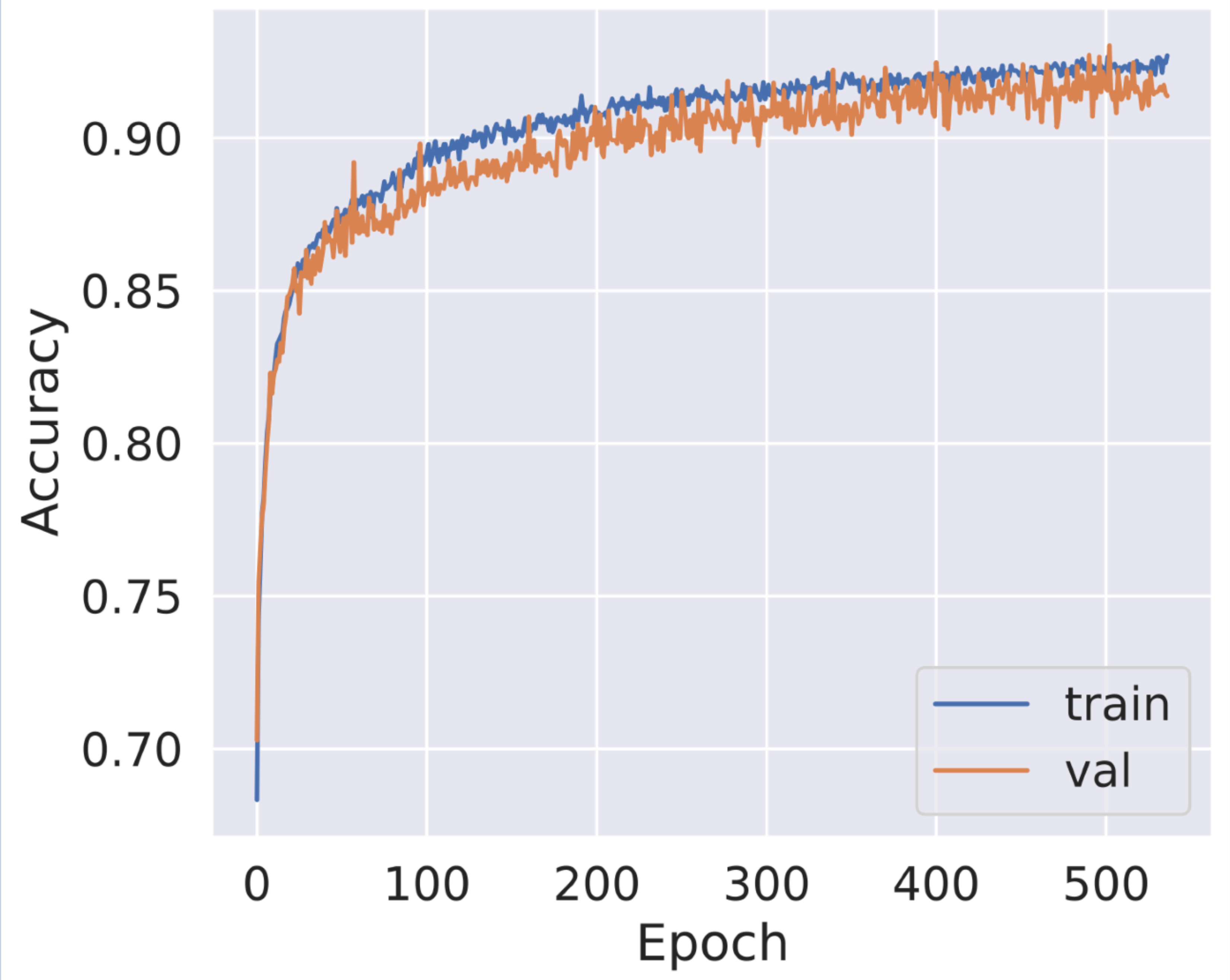
Training the NN

Accuracy Curves

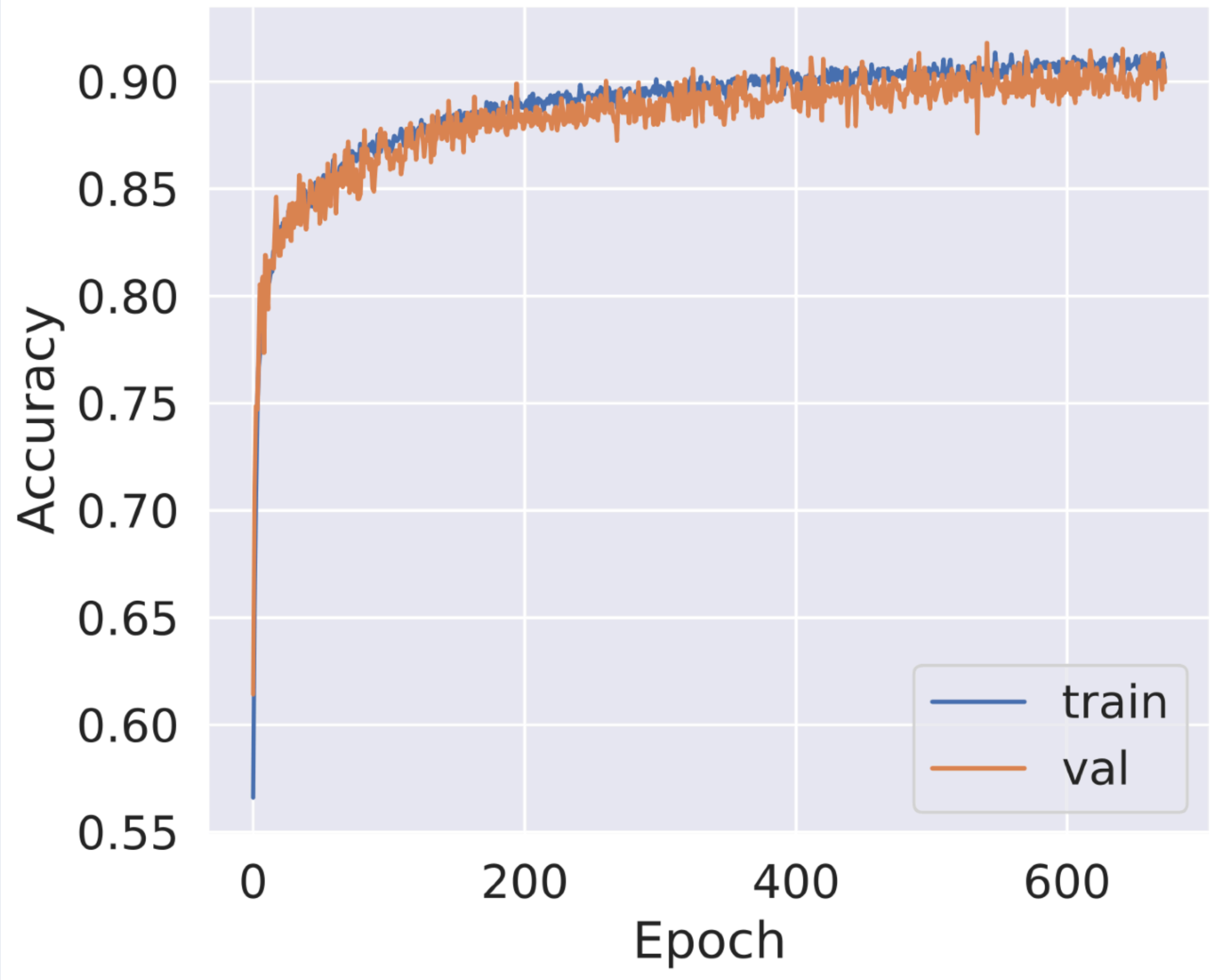
β_1



β_2



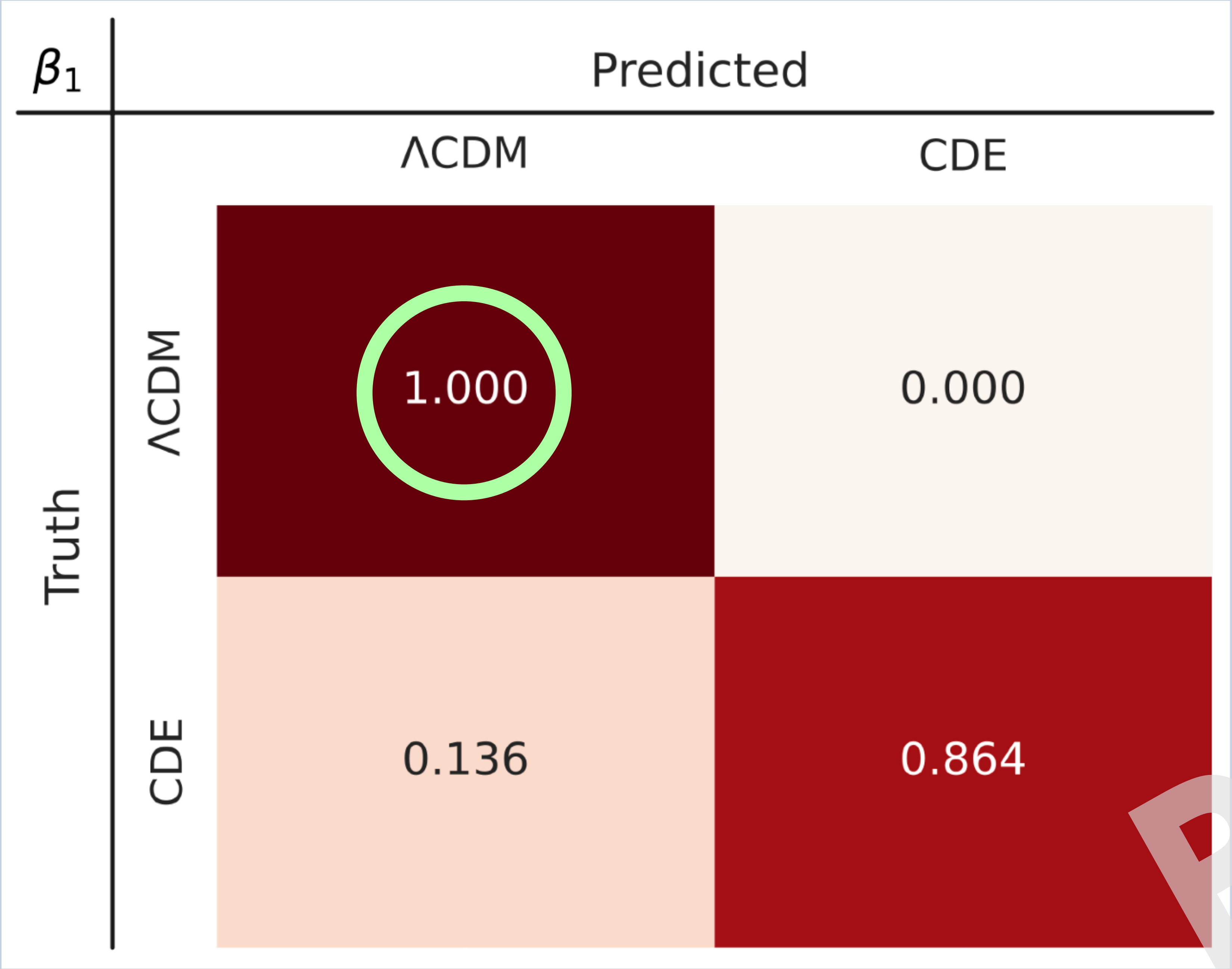
β_3



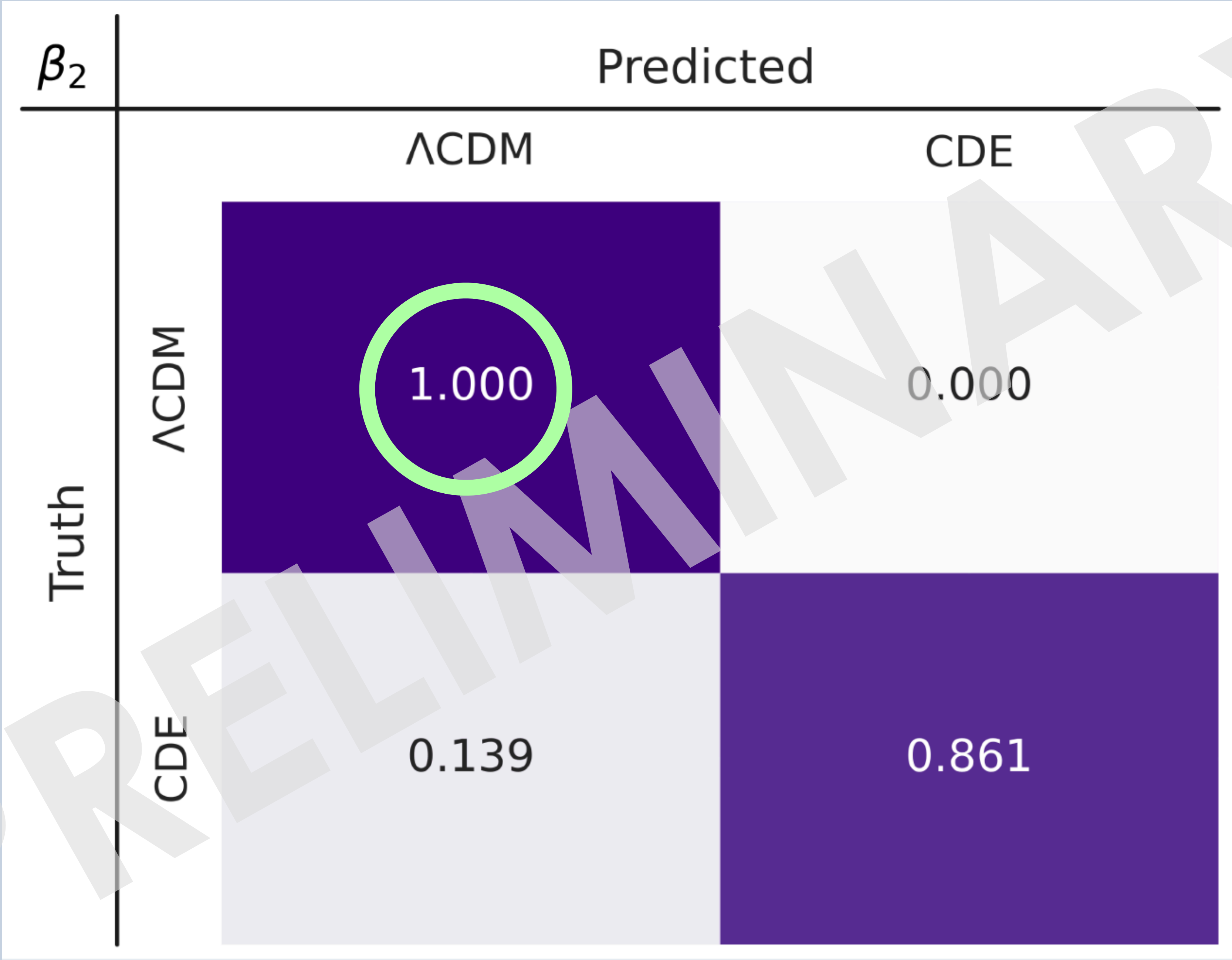
Results

Confusion Matrices

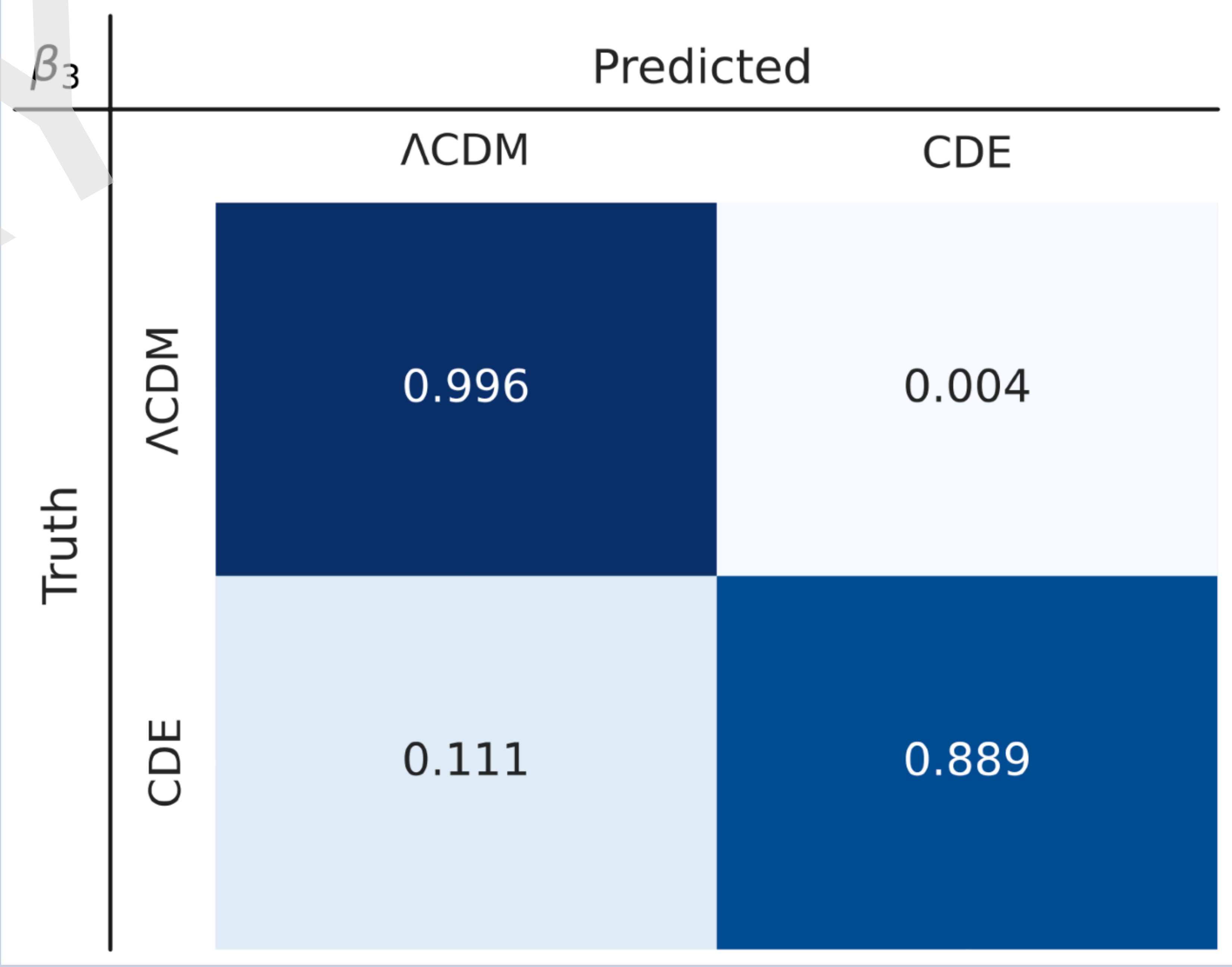
β_1



β_2



β_3



Multiclass Classification

Generalising between coupling epochs

Can we differentiate between early and late time coupling?

SORT OF!

Truth \ Predicted	ΛCDM	β_1	β_2	β_3
ΛCDM	0.99 ±0.01	0.13 ±0.02	0.15 ±0.03	0.18 ±0.03
β_1	0.00 ±0.00	0.76 ±0.04	0.01 ±0.02	0.01 ±0.02
β_2	0.00 ±0.00	0.03 ±0.02	0.20 ±0.06	0.12 ±0.06
β_3	0.01 ±0.01	0.08 ±0.04	0.63 ±0.06	0.69 ±0.06

Conclusions

- We demonstrate the feasibility of our NN in differentiating between ΛCDM and tomographic CDE models, using simulated stage-IV survey data of $f\sigma_8$'s
- With binary classification, the network performs similarly well regardless of epoch in which coupling is activated
- In multi-class classification, models with late-time coupling are most accurately identified by the NN
- In light of upcoming surveys like *Euclid*, which will provide us with a wealth of spectroscopic clustering data, NNs could prove a robust cross-check of conventional Bayesian methods in constraining beyond-ΛCDM physics

Thank you!

Lisa Goh
CosmoStat, CEA Paris-Saclay
lisa.goh@cea.fr