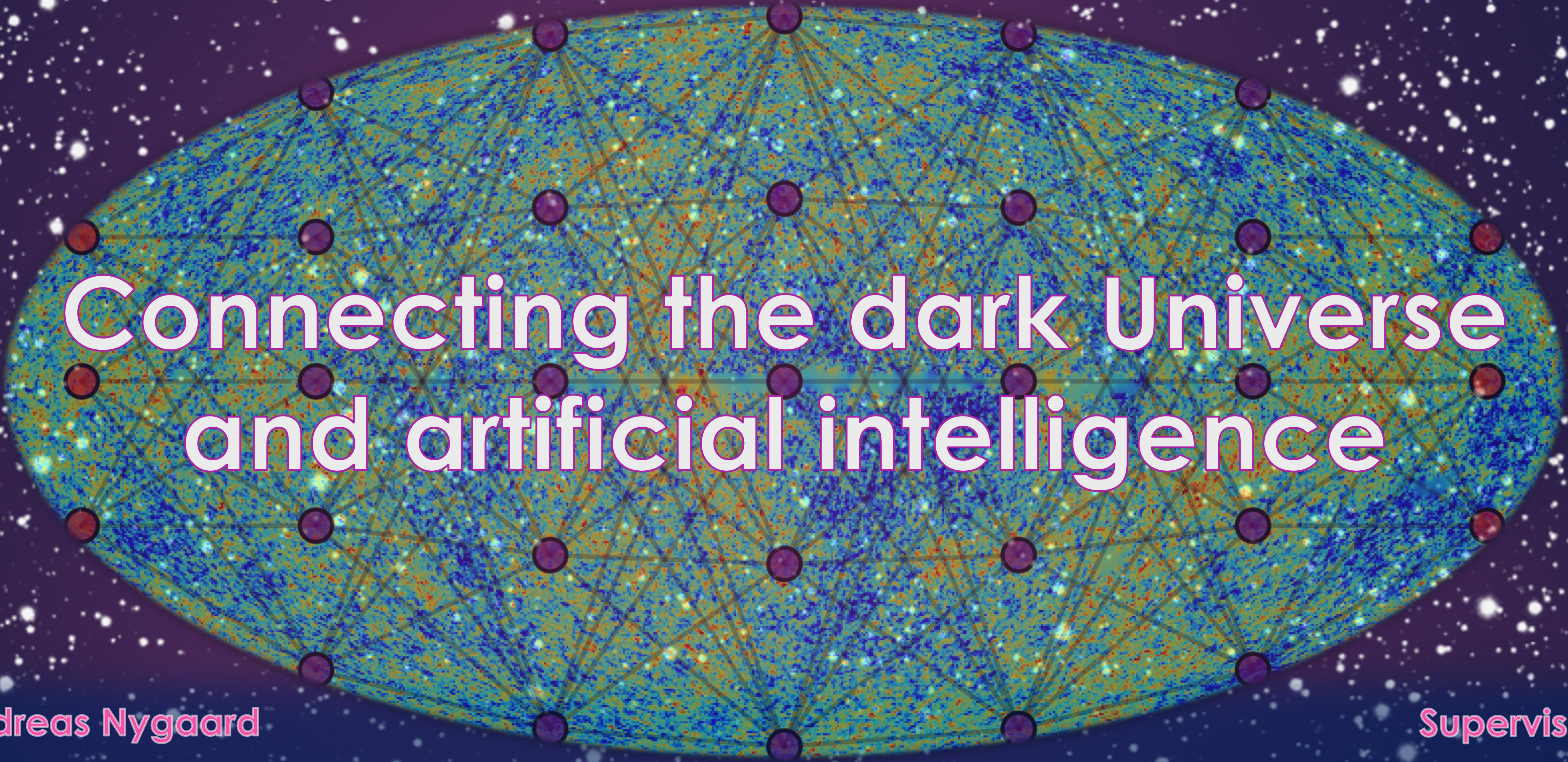
Connecting the dark Universe and artificial intelligence



Andreas Nygaard

andreas@phys.au.dk

1525-520, Aarhus University

Supervisors:

Thomas Tram,

Steen Hannestad

Contents

- Introduction and motivation | oo
- Machine learning （CONNECT） | oooooooooooooo
- Current and future developments | ooo

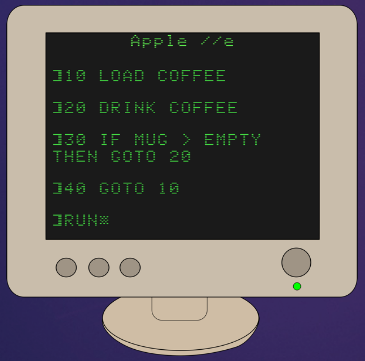
Introduction and motivation

Inference in cosmology

$\vec{\theta}$

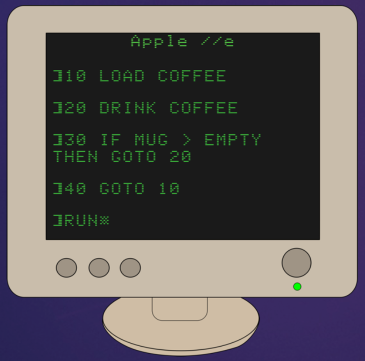
Inference in cosmology

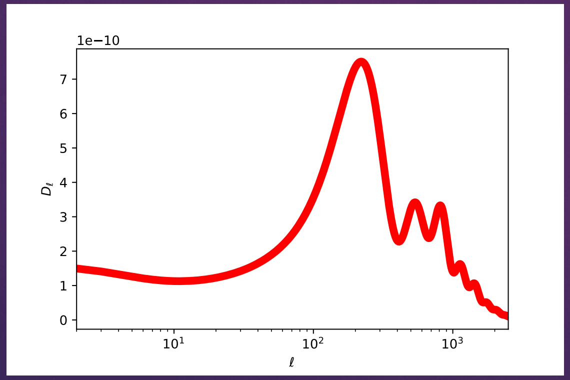
$\vec{\theta}$



Inference in cosmology

$\vec{\theta}$

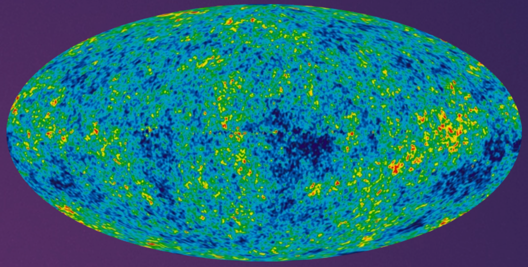


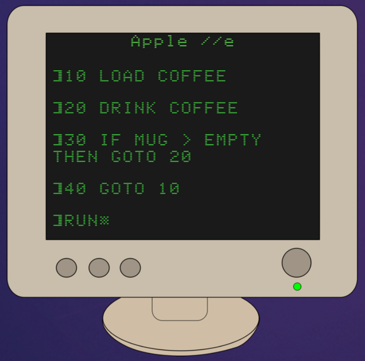


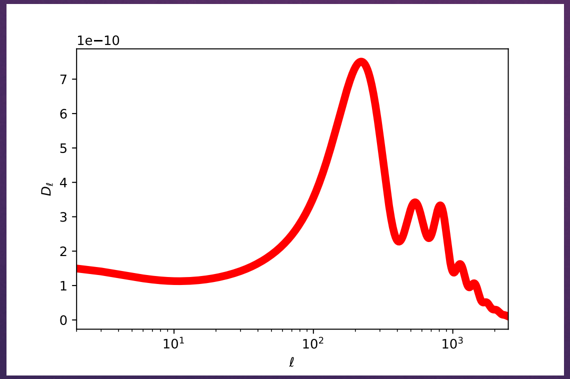
Inference in cosmology

$\vec{\theta}$

$\mathcal{L}$



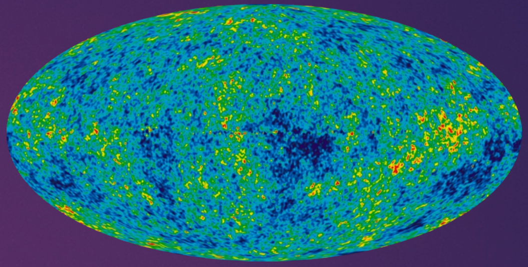




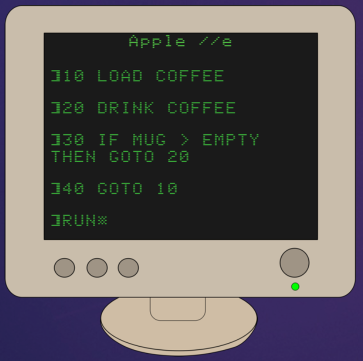
Inference in cosmology

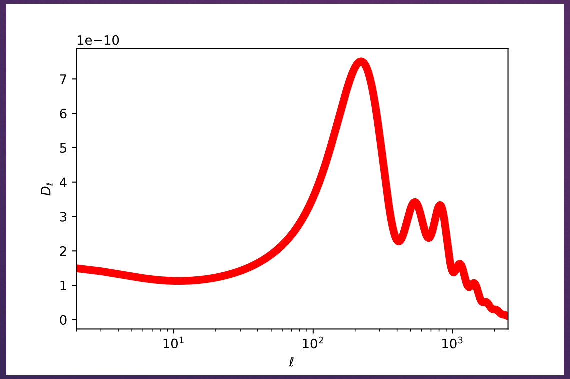
$\vec{\theta}$

$\mathcal{L}$



$\sim 10^5 - 10^6$

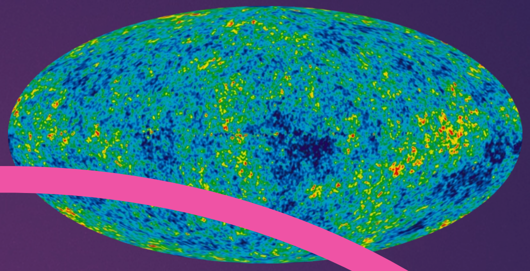




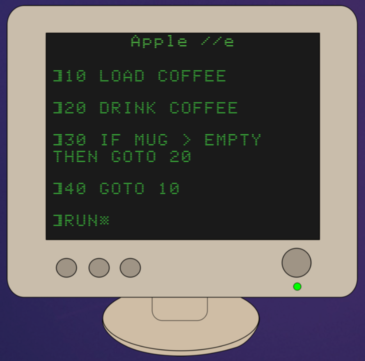
Inference in cosmology

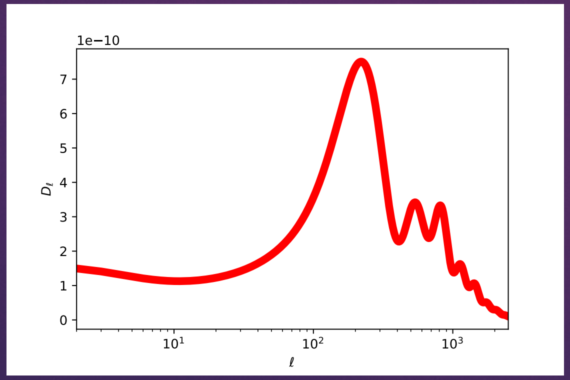
$\vec{\theta}$

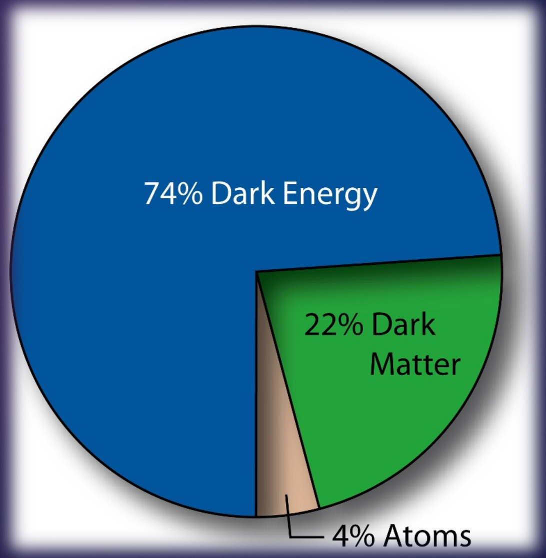
$\mathcal{L}$



$\sim 10^5 - 10^6$



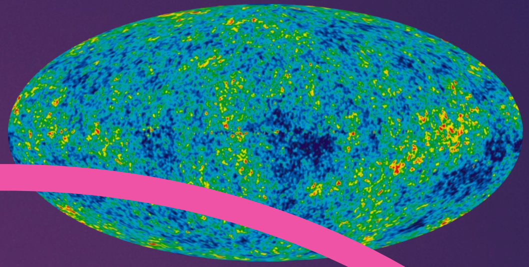




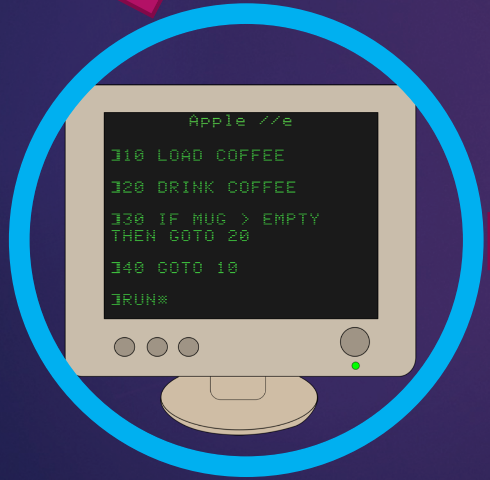
Inference in cosmology

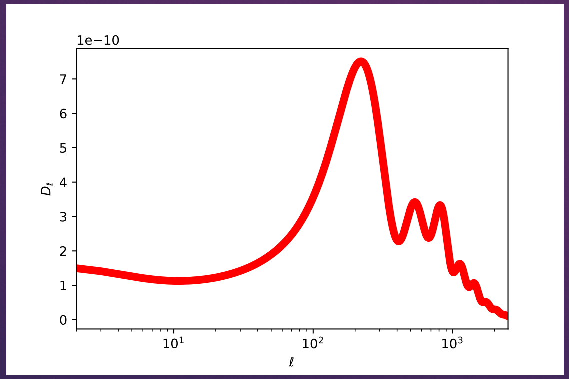
$\vec{\theta}$

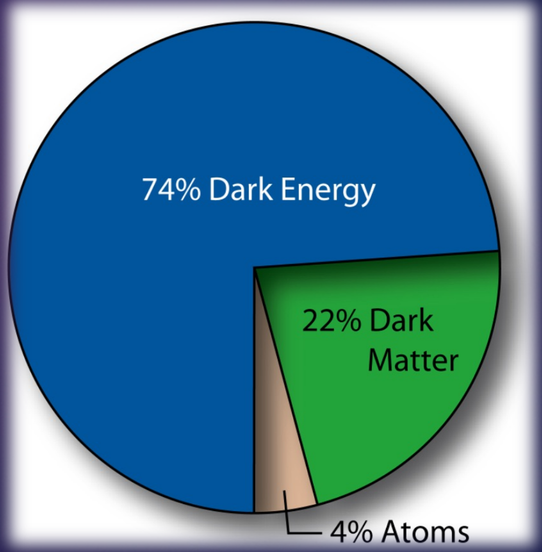
$\mathcal{L}$



$\sim 10^5 - 10^6$



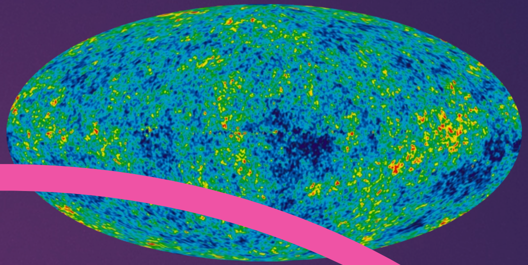




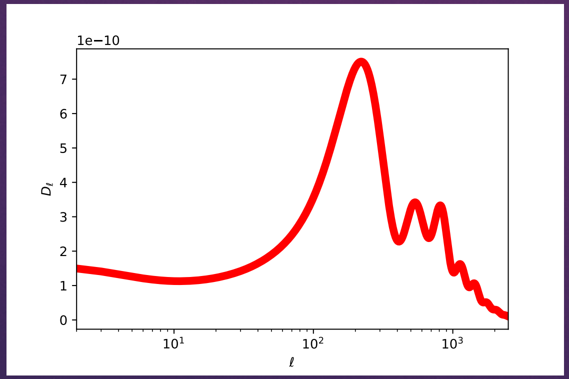
Inference in cosmology

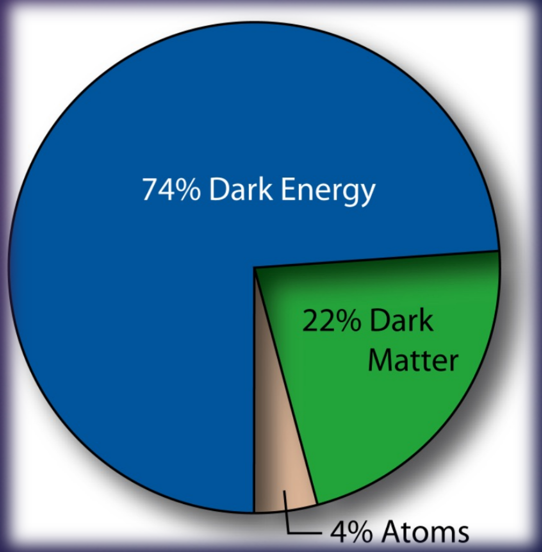
$\vec{\theta}$

$\mathcal{L}$



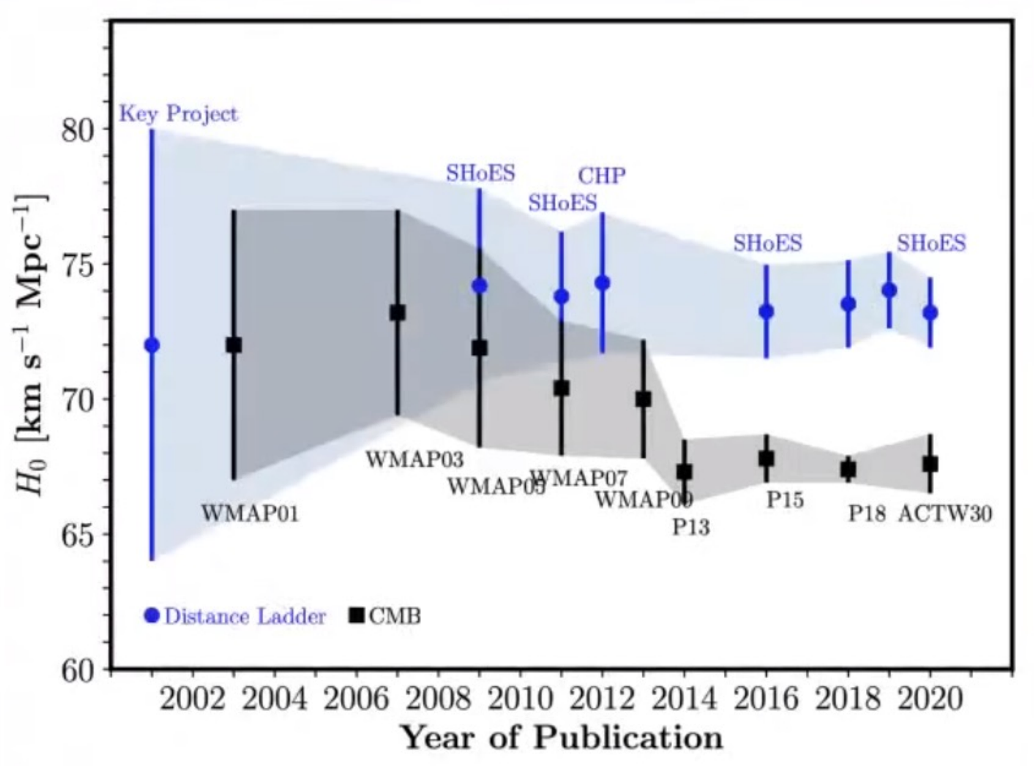
$\sim 10^5 - 10^6$





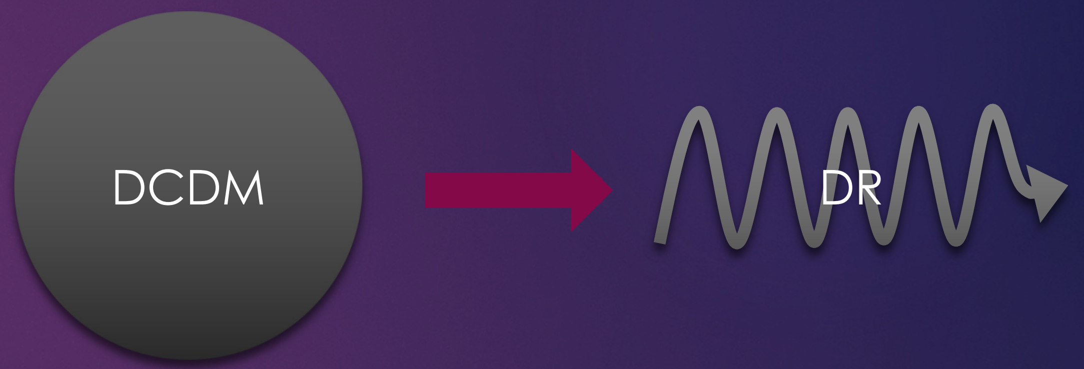
Decaying dark matter

- Why a decay?
  - Most particles in SM decay
  - Hubble tension
- Simple DCDM→DR model
  - $\omega_{\rm dcdm}^{\rm ini}, \quad f_{\rm dcdm}, \quad \Gamma_{\rm dcdm}$
- Two-body decay to WDM
- DWDM→WDM model



Decaying dark matter

- Why a decay?
  - Most particles in SM decay
  - Hubble tension
- Simple DCDM→DR model
  - $\omega_{\rm dcdm}^{\rm ini}, \quad f_{\rm dcdm}, \quad \Gamma_{\rm dcdm}$
- Two-body decay to WDM
- DWDM→WDM model

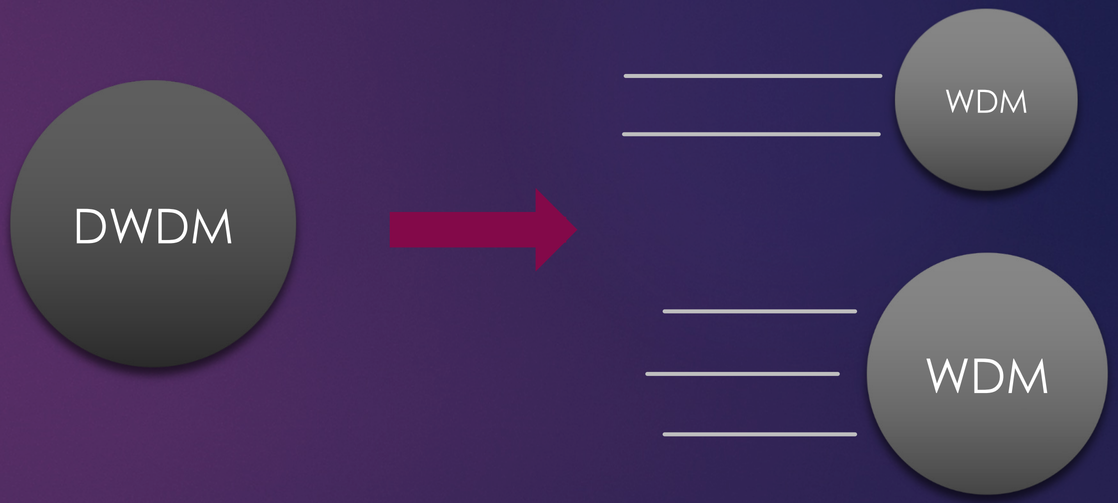


Decaying dark matter

- Why a decay?
  - Most particles in SM decay
  - Hubble tension
- Simple DCDM→DR model
  - $\omega_{\mathrm{dcdm}}^{\mathrm{ini}}, \quad f_{\mathrm{dcdm}}, \quad \Gamma_{\mathrm{dcdm}}$
- Two-body decay to WDM
- DWDM→WDM model

Decaying dark matter

- Why a decay?
  - Most particles in SM decay
  - Hubble tension
- Simple DCDM→DR model
  - $\omega_{\mathrm{dcdm}}^{\mathrm{ini}}, \quad f_{\mathrm{dcdm}}, \quad \Gamma_{\mathrm{dcdm}}$
- Two-body decay to WDM
- DWDM→WDM model



Machine learning

CONNECT

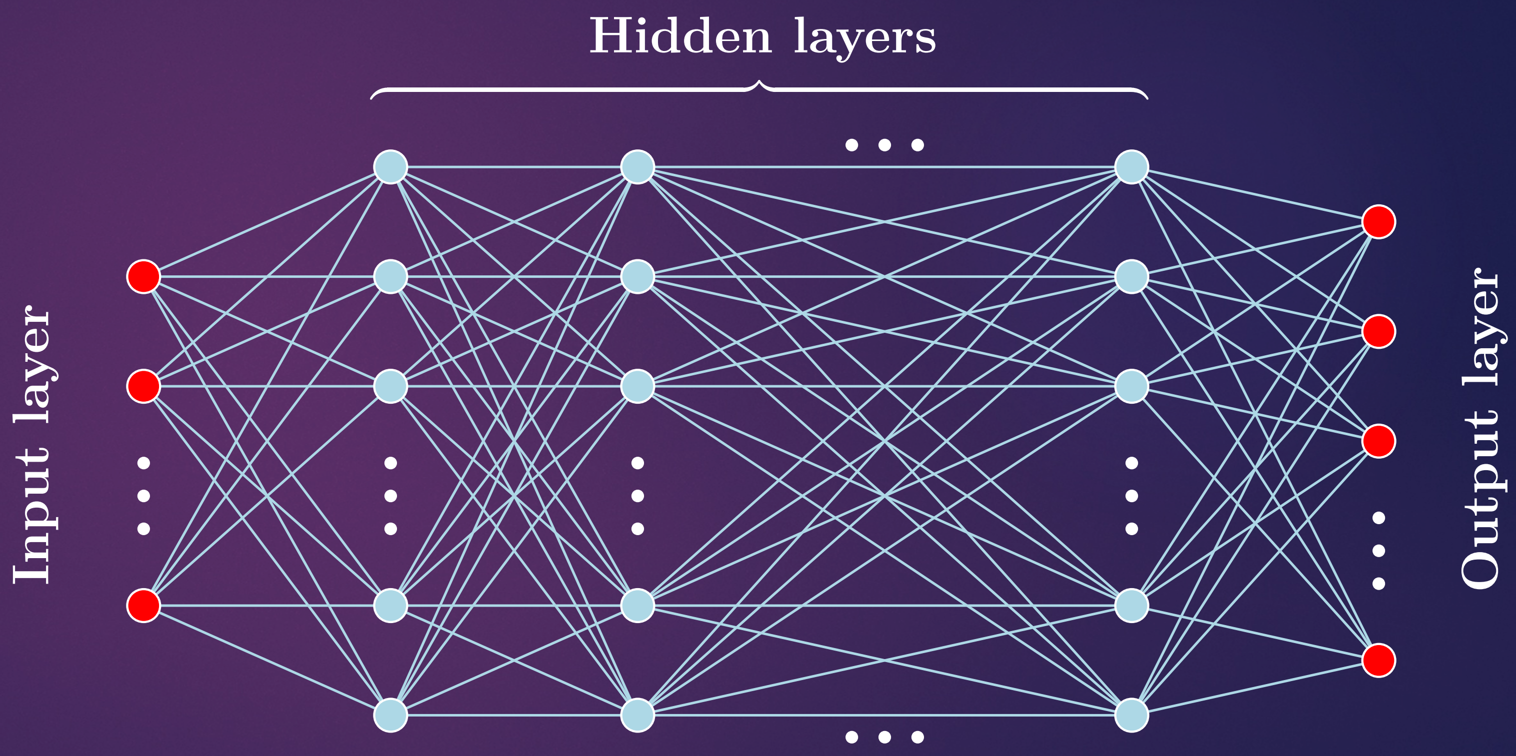
COsmological **N**eural **N**etwork **E**mulation of **C**LASS using **T**ensorflow

Neural network

- Emulation of $C_\ell s$
  - Weights
  - Bias
  - Activation

$$x_{n+1}^{i} = A\left(\vec{x}_n \cdot \vec{w}^{i} + b^{i}\right)$$

- Architecture
  - Fully connected
  - Hidden layers
  - Nodes



Choice of activation and loss

- Activation function
  - ReLU?
  - Parameterised ReLU with exponential smoothing – Alsing et al. (2019)

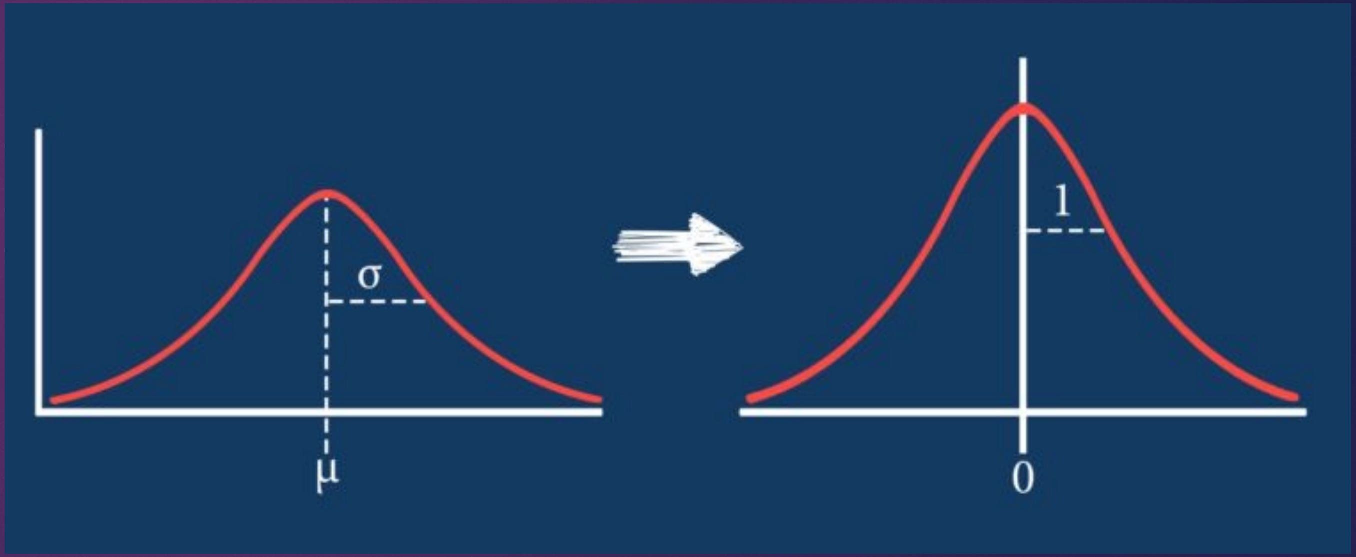
$$\vec{f}(\vec{x}) = \left(\vec{\gamma} + \left(1 + e^{-\vec{\beta}\odot\vec{x}}\right)^{-1} \odot (1 - \vec{\gamma})\right) \odot \vec{x}$$

- Loss function
  - MSE
  - Cosmic variance

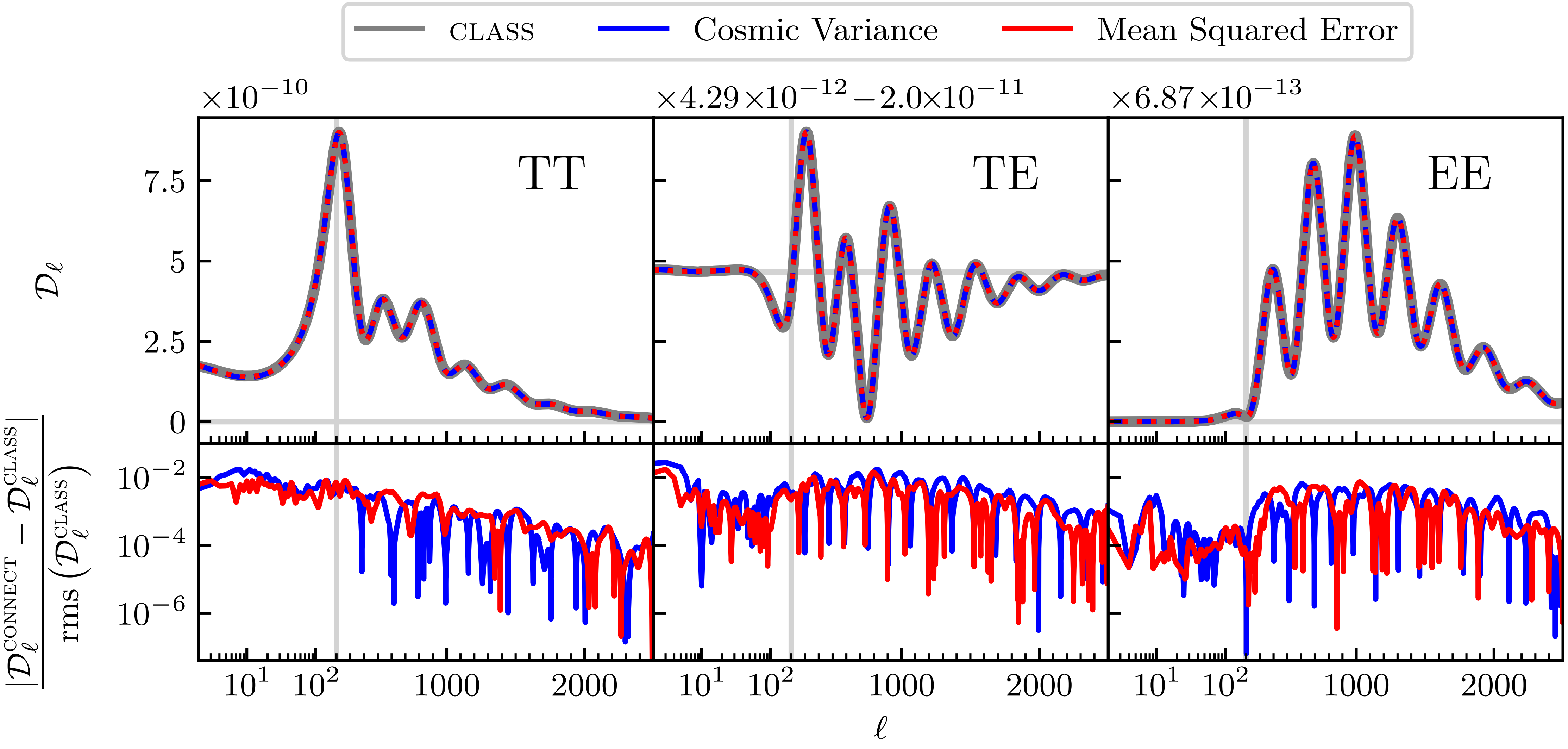
$$L(\vec{x}, \vec{y}) = \frac{1}{n}\sum_{i=1}^{n} a(i)\,(x_i - y_i)^2, \qquad a(i) = \begin{cases} 1, & MSE \\ \dfrac{2\ell_i + 1}{2}, & CV \end{cases}$$

Normalisation

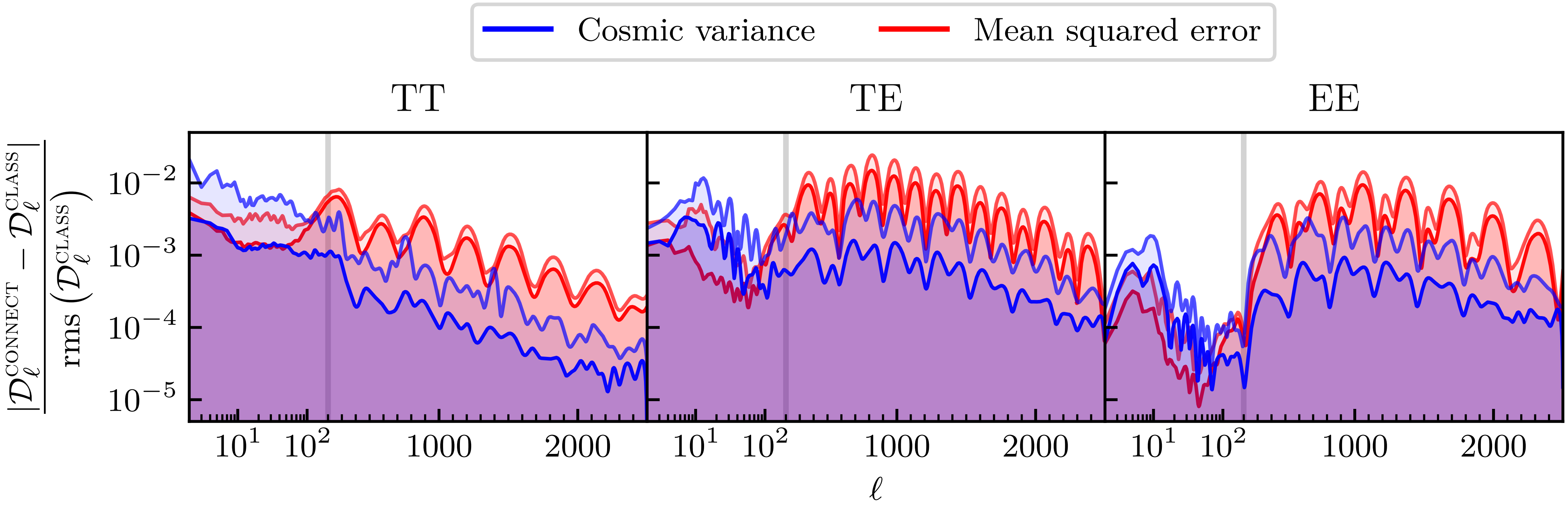
- $C_\ell s$ are quite small ($< 10^{-10}$)
- Normalization methods
  - Logarithmic
  - Min-Max
  - Standardisation
- Normalisation of inputs
  - TensorFlow routine
- Normalisation of outputs
  - Custom implementation



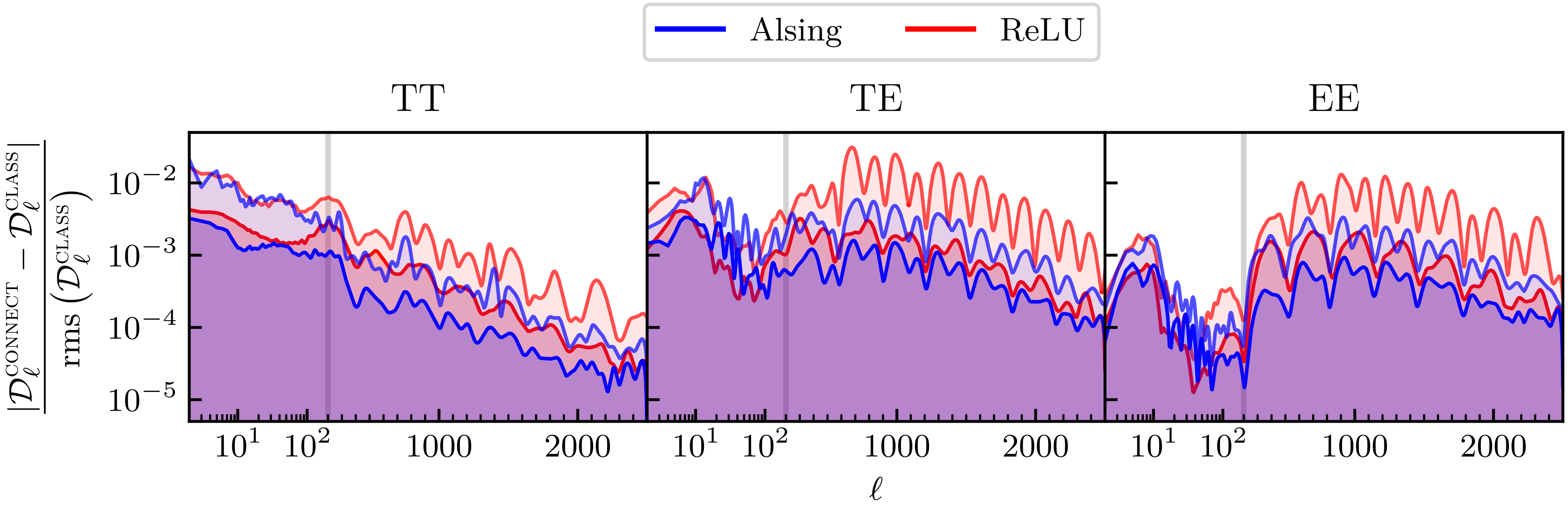
Precision


CLASS
Cosmic Variance
Mean Squared Error
×10^−10
×4.29×10^−12 −2.0×10^−11
×6.87×10^−13
TT
TE
EE
$\mathcal{D}_\ell$
7.5
5
2.5
0
$\frac{|\mathcal{D}_\ell^{\mathrm{CONNECT}} - \mathcal{D}_\ell^{\mathrm{CLASS}}|}{\mathrm{rms}(\mathcal{D}_\ell^{\mathrm{CLASS}})}$
10^{-2}
10^{-4}
10^{-6}
10^1
10^2
1000
2000
ℓ

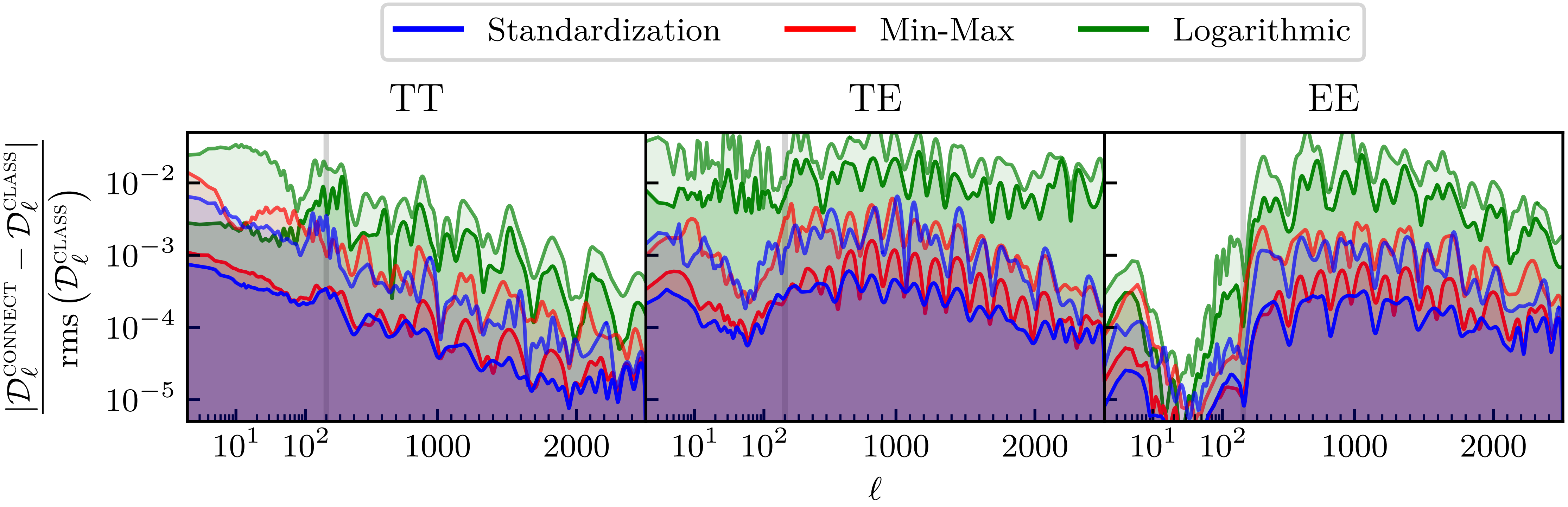
Precision


Cosmic variance
Mean squared error
TT
TE
EE
$\frac{\left|\mathcal{D}_\ell^{\mathrm{CONNECT}} - \mathcal{D}_\ell^{\mathrm{CLASS}}\right|}{\mathrm{rms}\left(\mathcal{D}_\ell^{\mathrm{CLASS}}\right)}$
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^1
10^2
1000
2000
ℓ

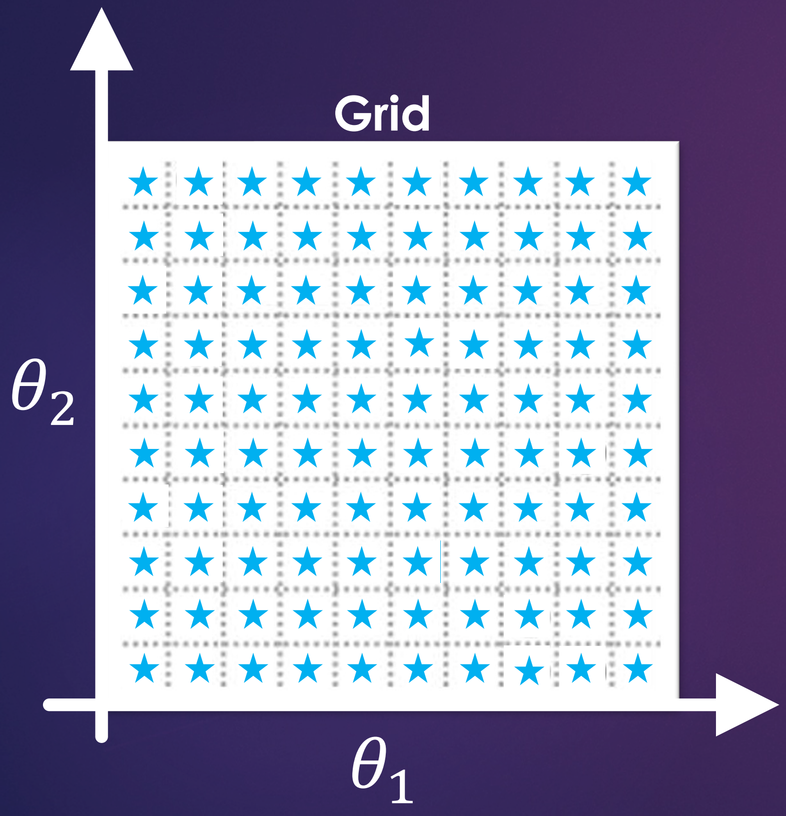
Precision

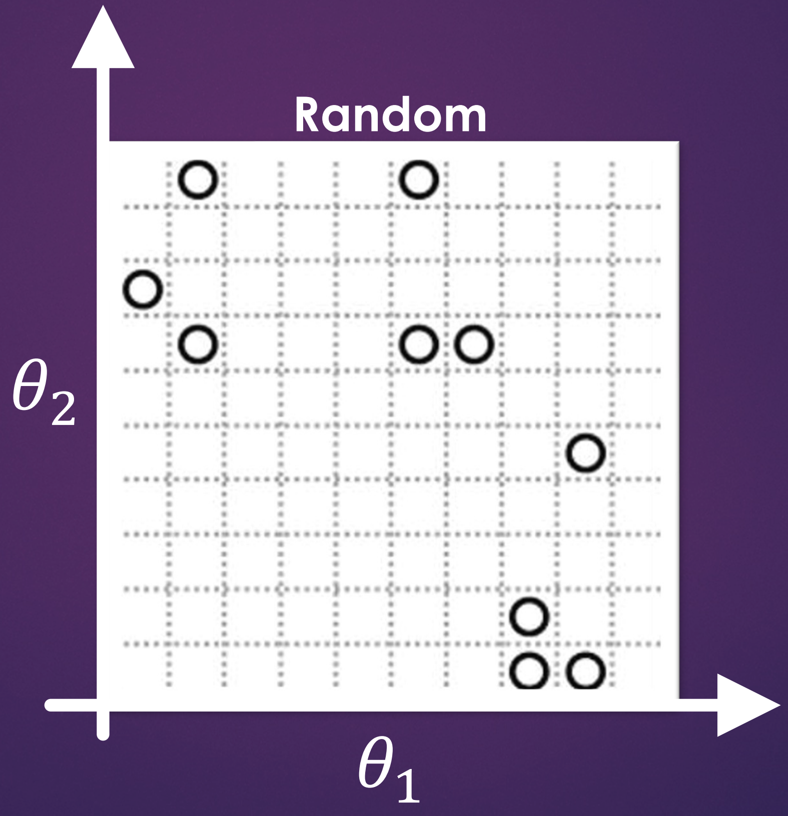

Alsing
ReLU
TT
TE
EE
$\frac{\left|\mathcal{D}_\ell^{\mathrm{CONNECT}} - \mathcal{D}_\ell^{\mathrm{CLASS}}\right|}{\mathrm{rms}\left(\mathcal{D}_\ell^{\mathrm{CLASS}}\right)}$
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{1} 10^{2} 1000 2000
ℓ

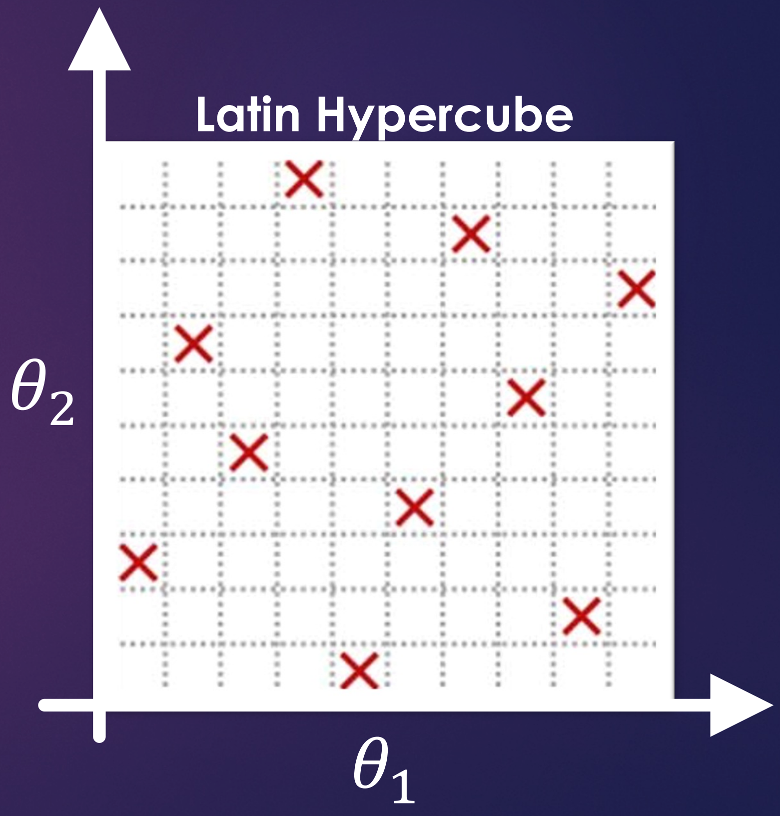
Precision


Standardization
Min-Max
Logarithmic
TT
TE
EE
$\frac{|\mathcal{D}_\ell^{\mathrm{CONNECT}} - \mathcal{D}_\ell^{\mathrm{CLASS}}|}{\mathrm{rms}(\mathcal{D}_\ell^{\mathrm{CLASS}})}$
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^1
10^2
1000
2000
ℓ

Sampling of training data





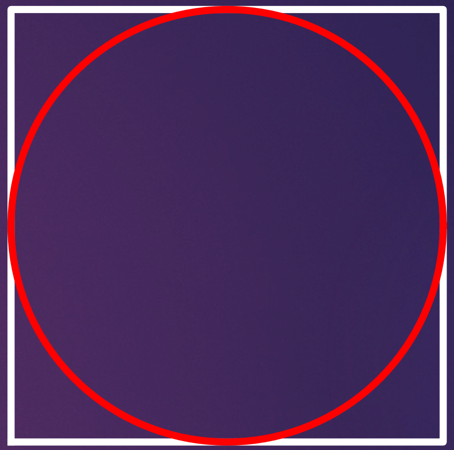


Problems with LHS

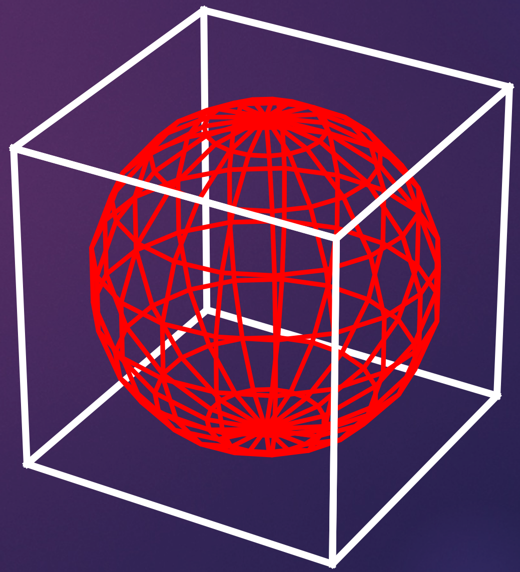
- Many non-relevant models
  - Volume ratio in high-dimensional space

$$r_n = \frac{V_n^{sphere}}{V_n^{cube}} = \frac{\pi^{n/2}}{2\,\Gamma\left(\frac{n}{2}+1\right)}$$

- High likelihood near a boundary
  - No models share any parameter
  - Only a single point on each boundary



$$r_2 = \frac{\pi}{4} \approx 0.785$$

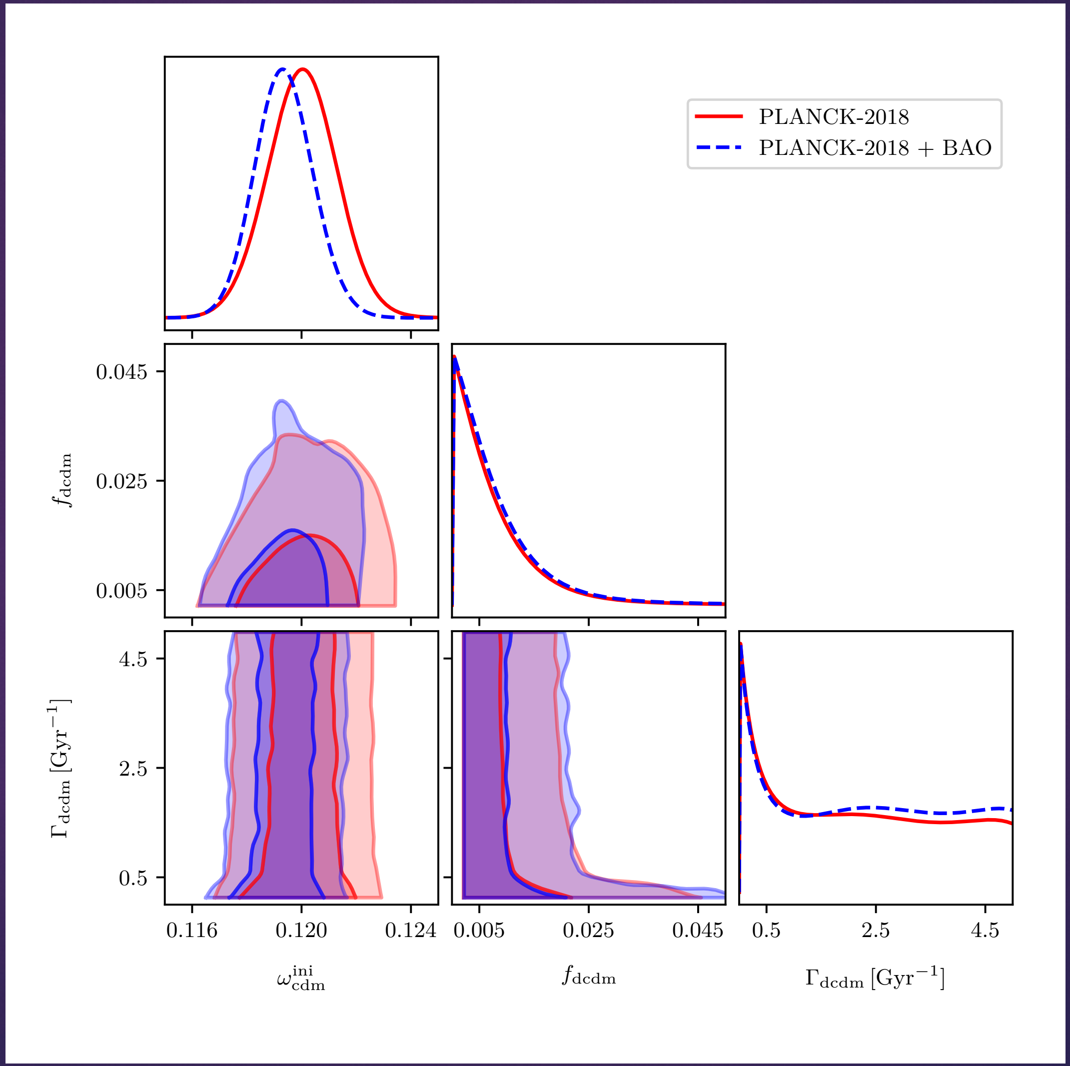


$$r_3 = \frac{\pi}{6} \approx 0.524$$

$$\vdots$$

$$r_8 = \frac{\pi^4}{6144} \approx 0.016$$

Problems with LHS



Nygaard, Tram, Hannestad (2020)

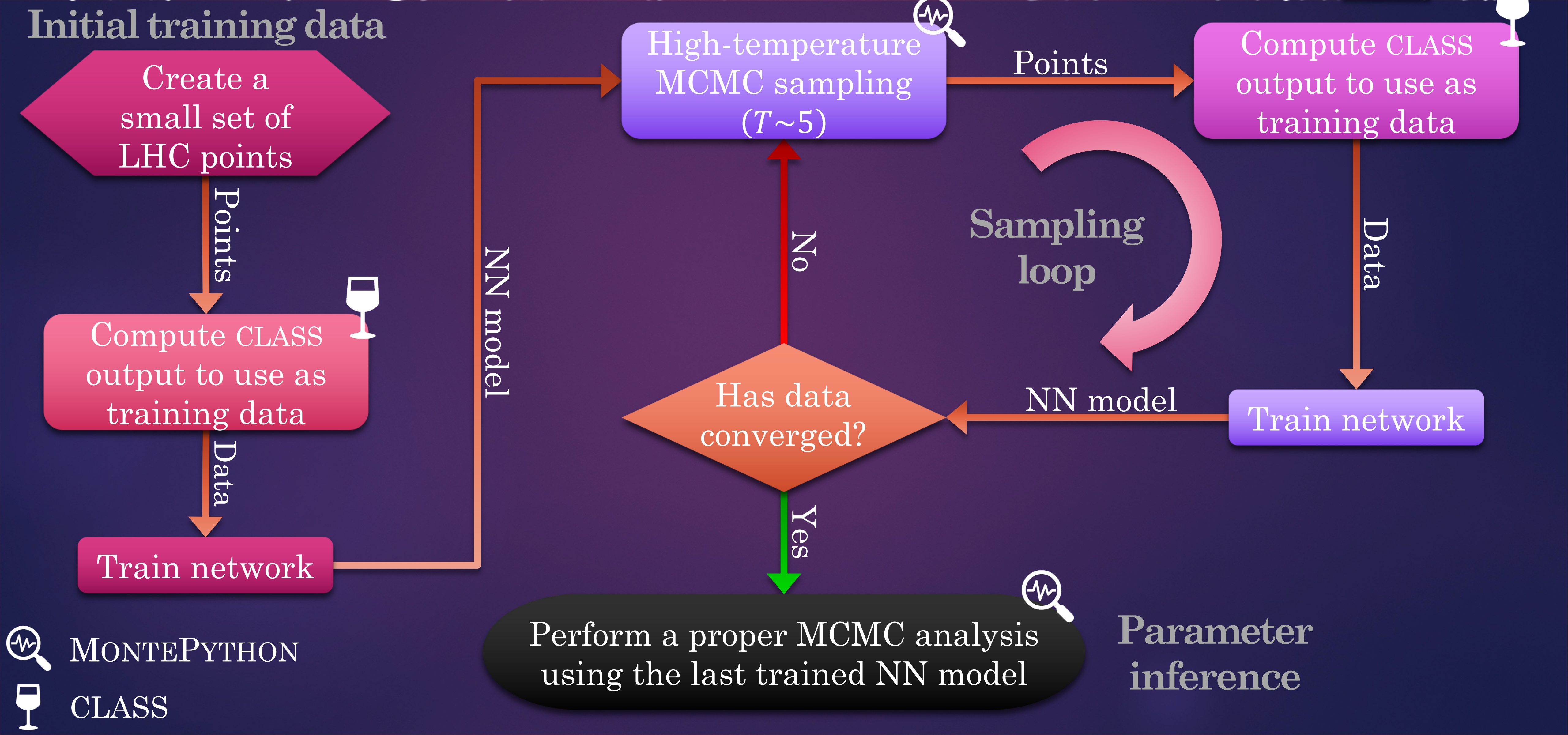
Sampling using MCMC

- Most optimal set of points
  - Fewer points
  - Cover the likelihood

- Drawbacks
  - Slow
  - Many wasted CLASS computations

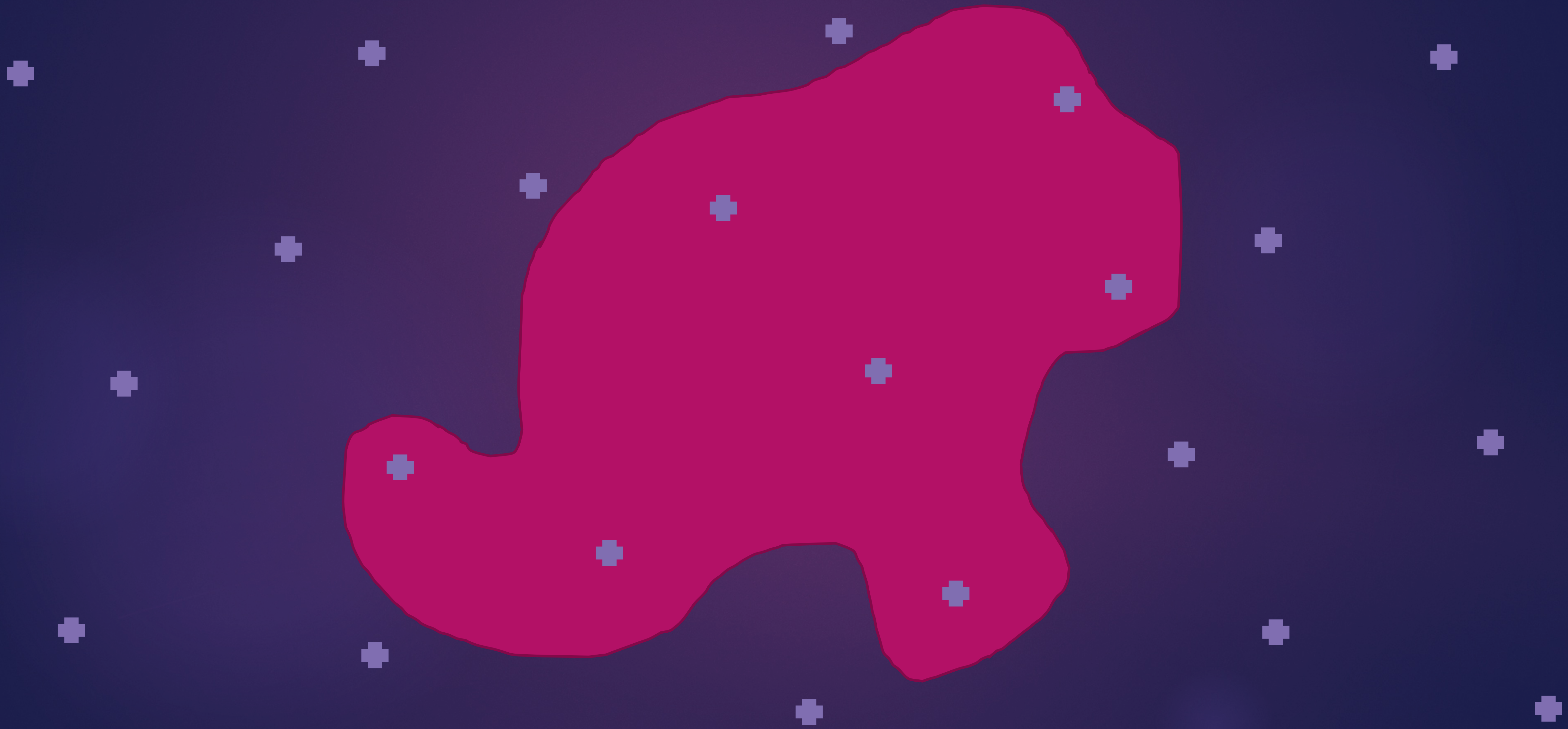
- New strategy – More networks!

Iterative sampling

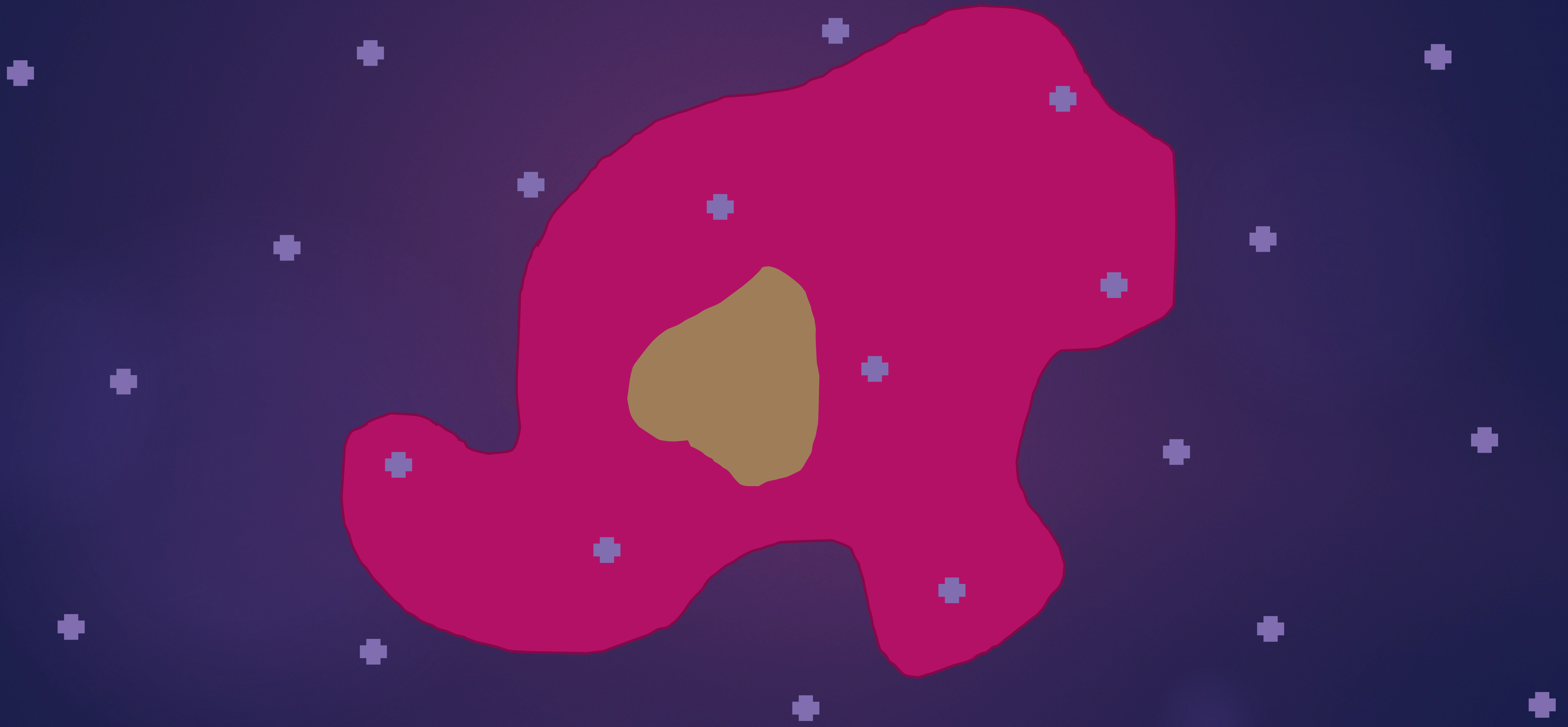


Iterative sampling

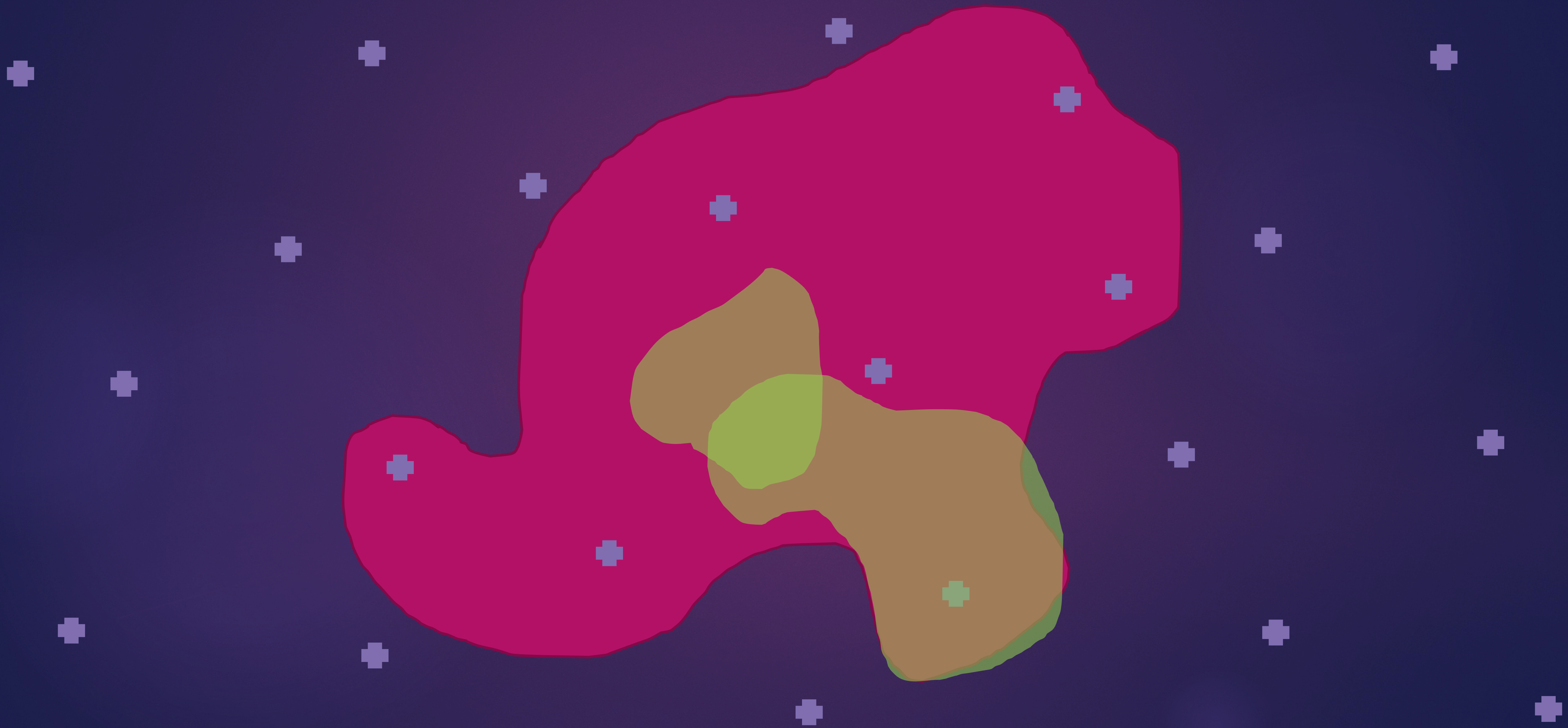
Iterative sampling



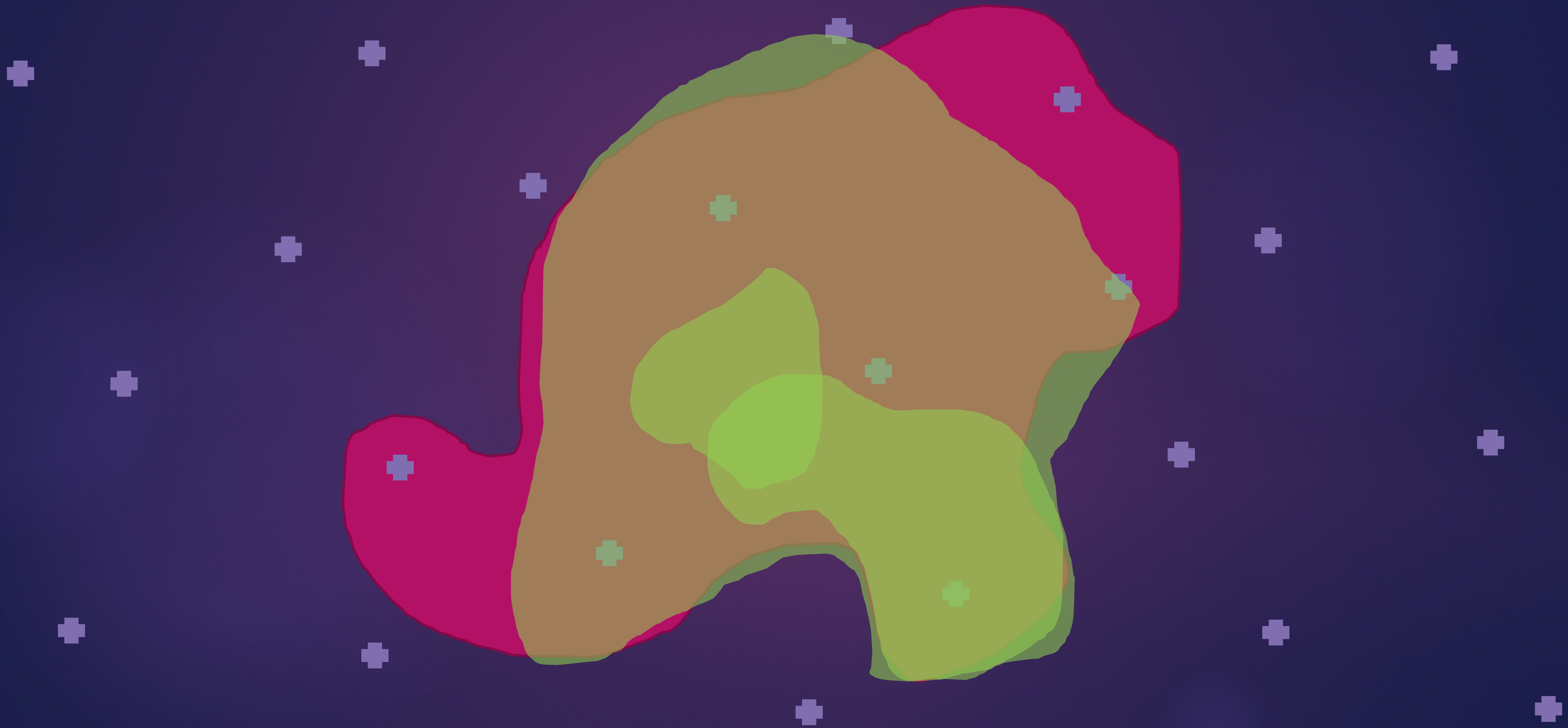
Iterative sampling



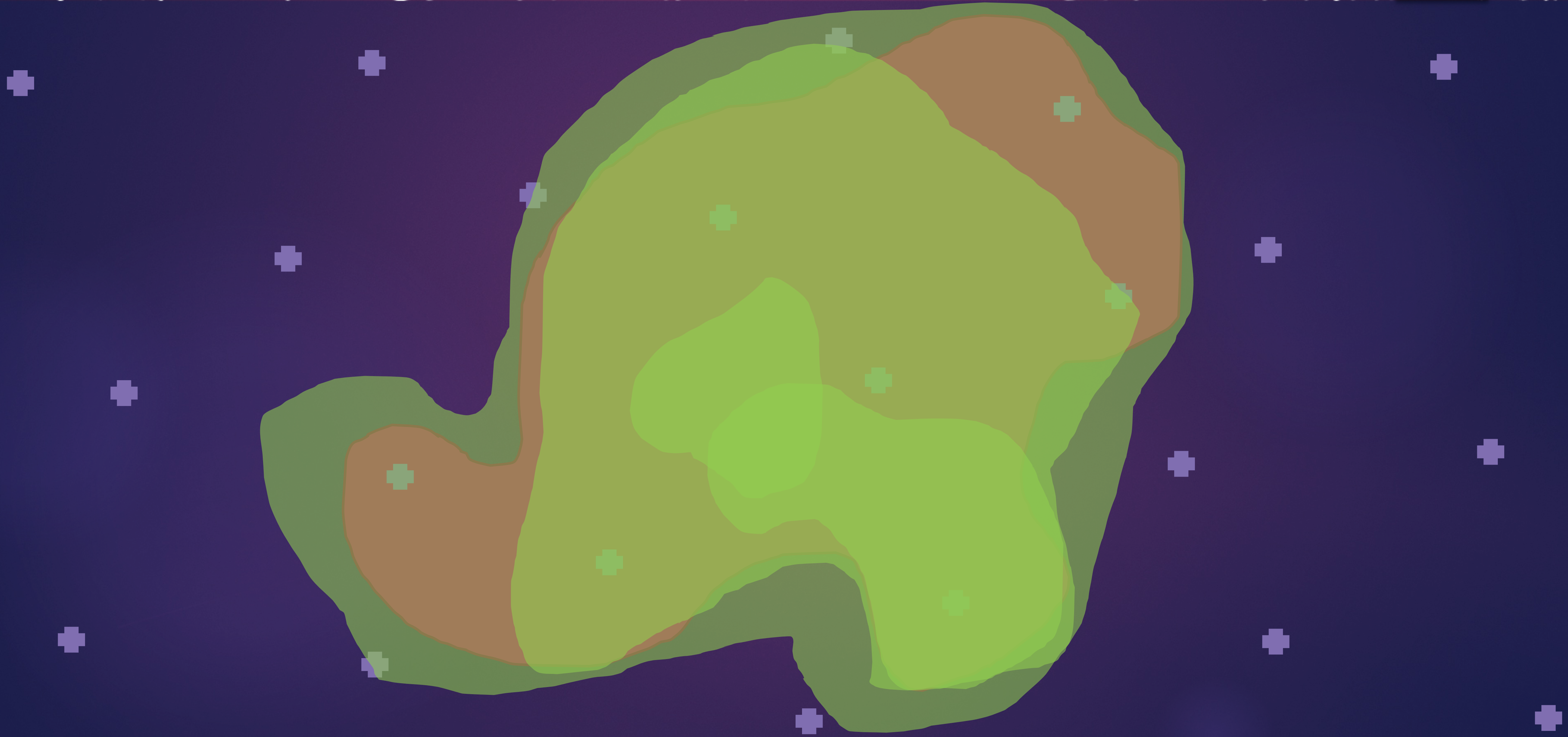
Iterative sampling



Iterative sampling

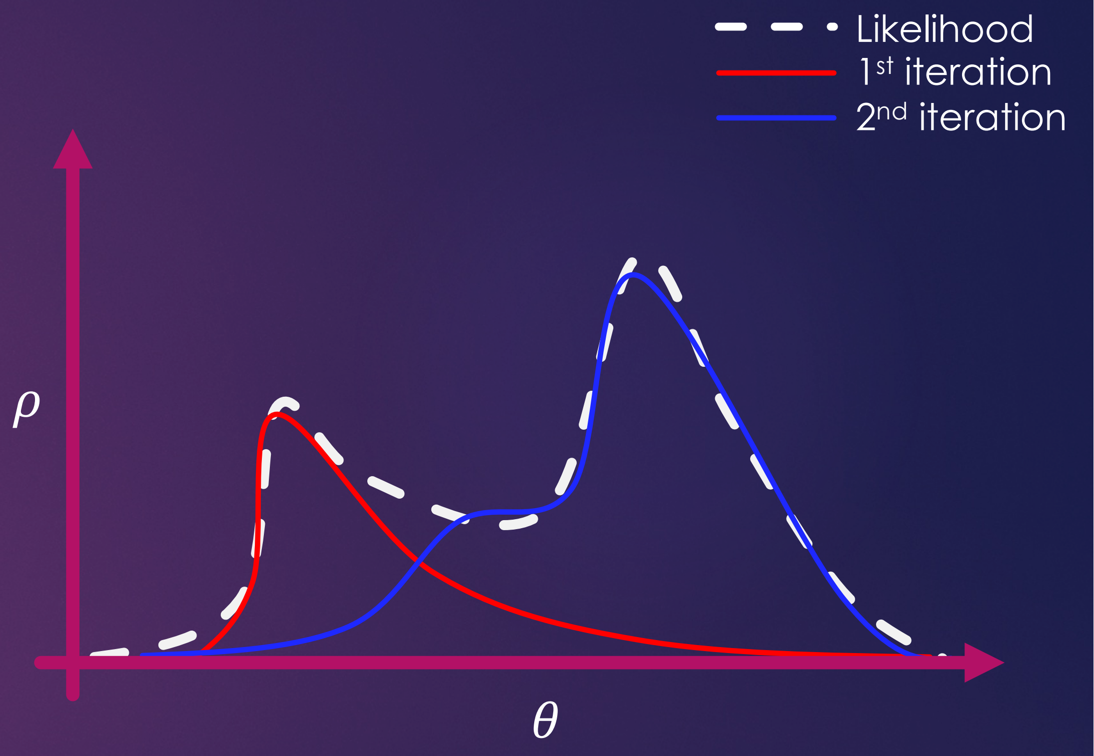


Iterative sampling



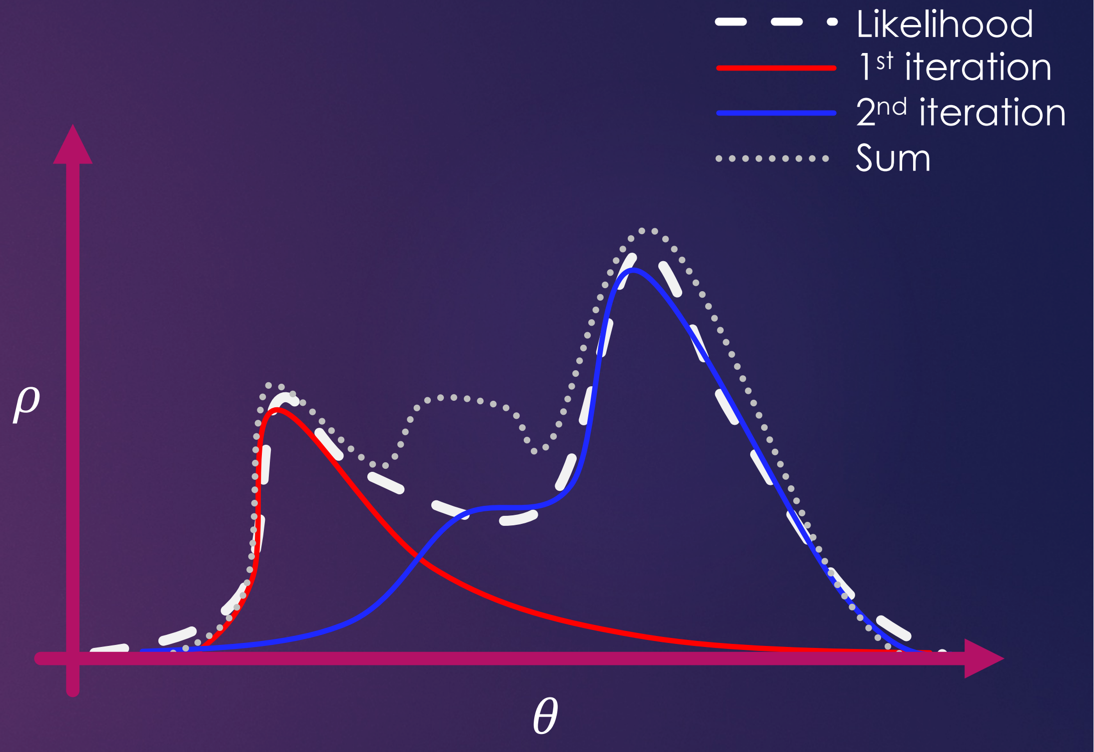
Oversampling of regions

- Overlap between iterations
  - Wrong point density
  - Bias in network
- Filtration of points
  - Accept point, x, from i^{th} iteration if $\rho_i(x) > \rho_{i+1}(x)$
- Problems
  - Very localised initial bias



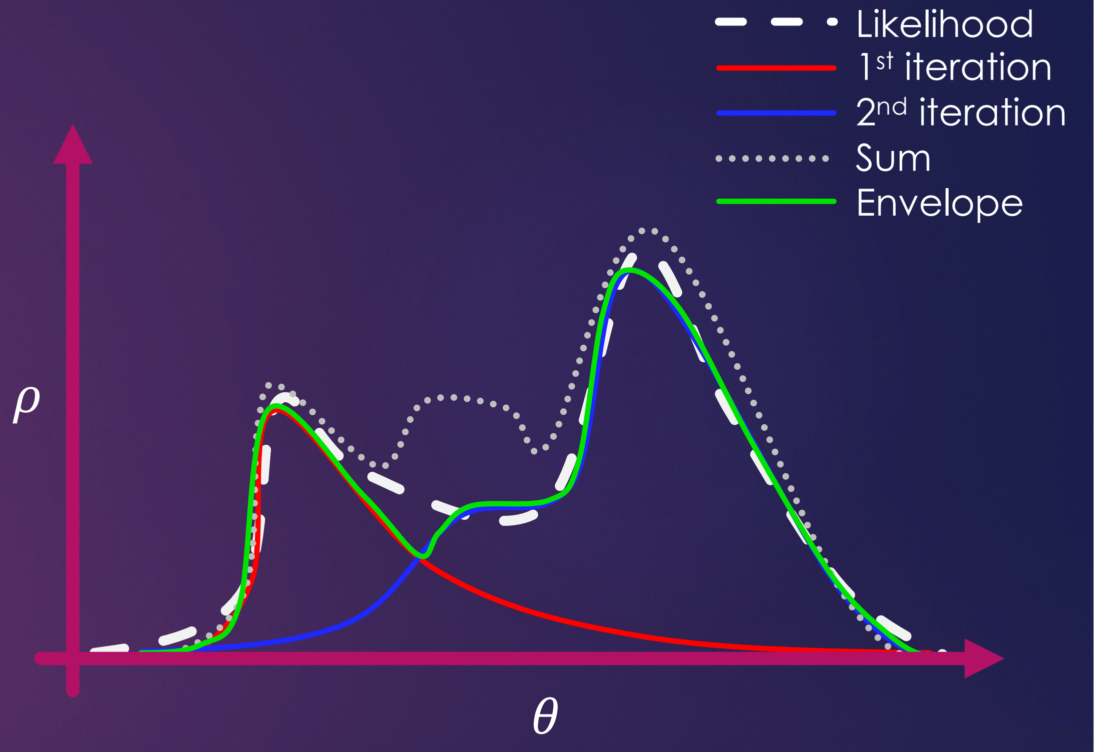
Oversampling of regions

- Overlap between iterations
  - Wrong point density
  - Bias in network
- Filtration of points
  - Accept point, x, from i^{th} iteration if $\rho_i(x) > \rho_{i+1}(x)$
- Problems
  - Very localised initial bias



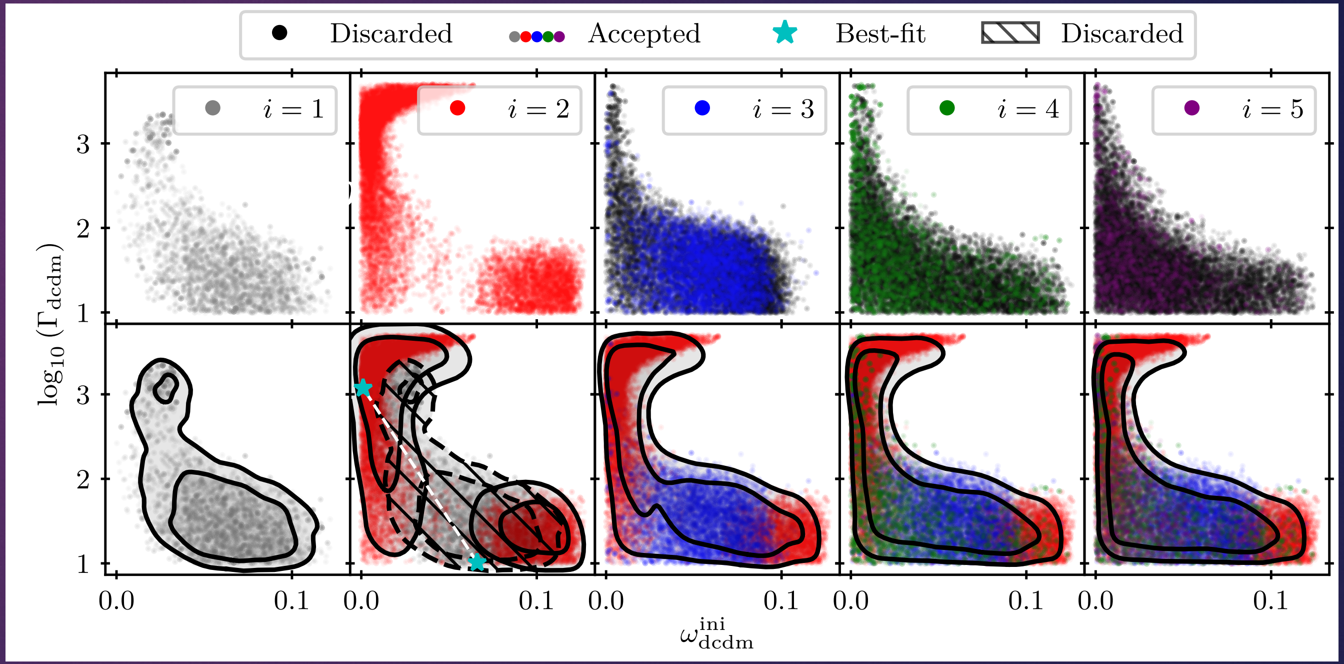
Oversampling of regions

- Overlap between iterations
  - Wrong point density
  - Bias in network
- Filtration of points
  - Accept point, x, from i^{th} iteration if $\rho_i(x) > \rho_{i+1}(x)$
- Problems
  - Very localised initial bias



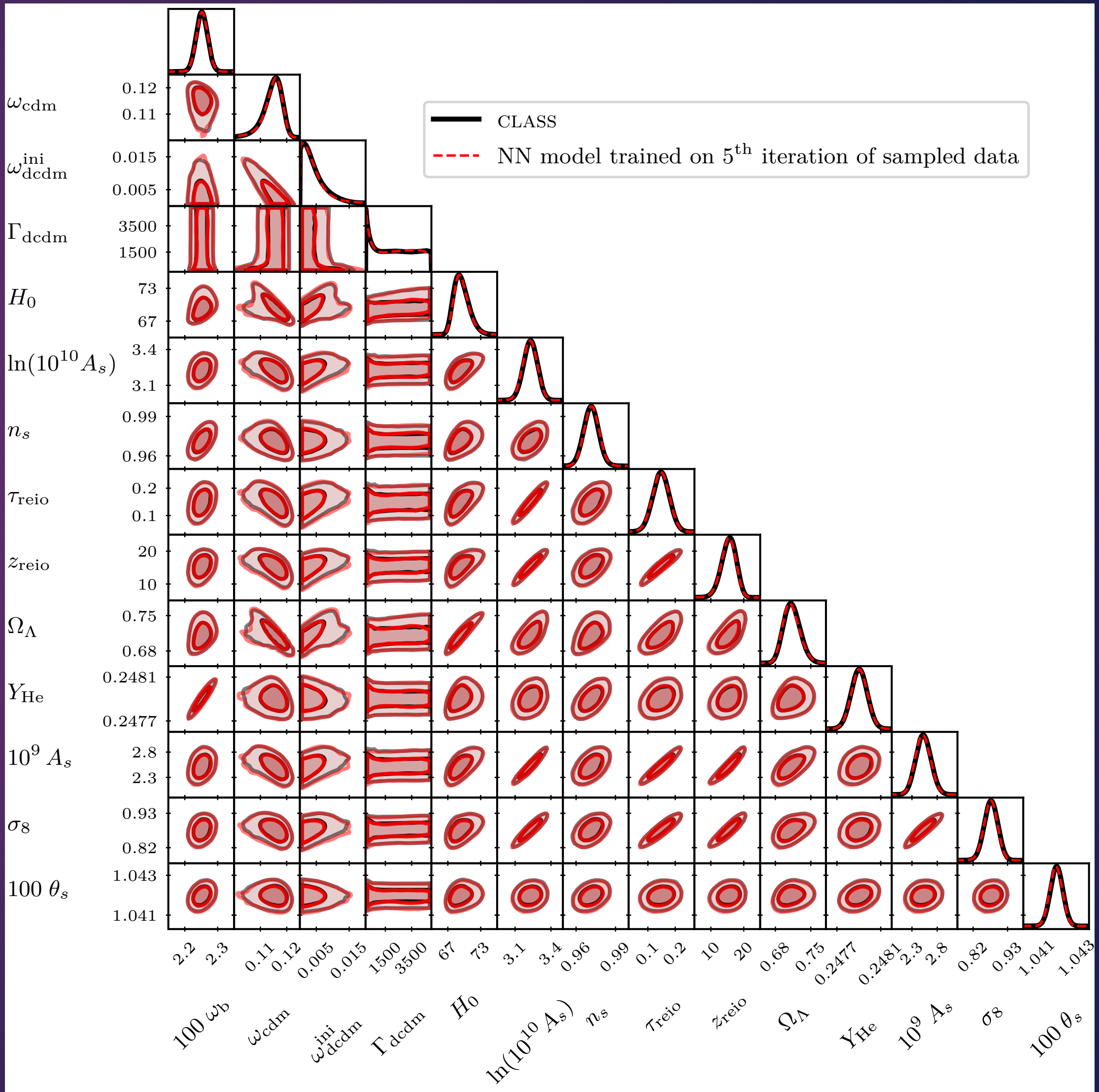
Oversampling of regions

- Overlap between iterations
  - Wrong point density
  - Bias in network
- Filtration of points
  - Accept point, x, from i^{th} iteration if $\rho_i(x) > \rho_{i+1}(x)$
- Problems
  - Very localised initial bias



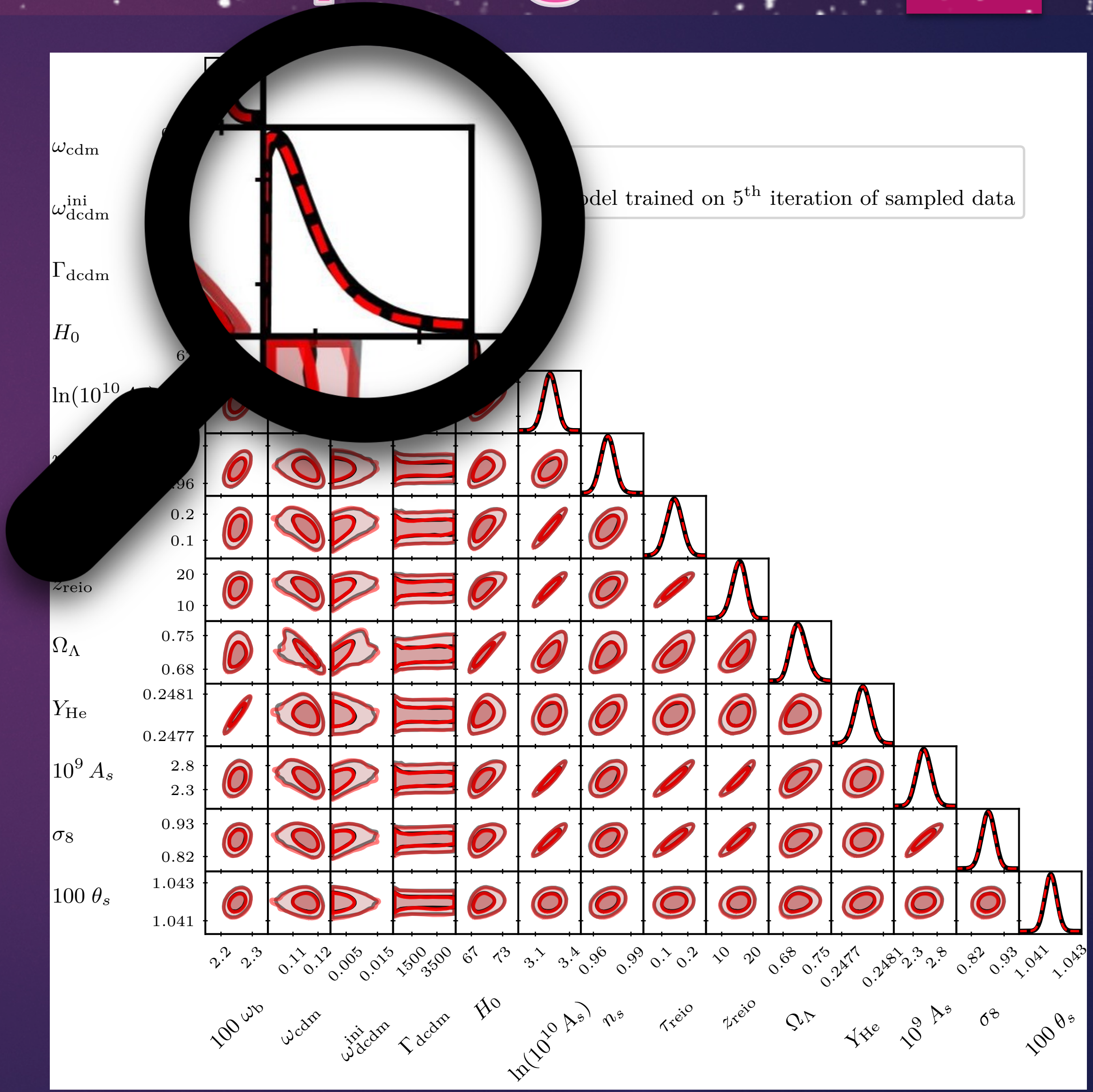
Results of iterative sampling

- DCDM model
  - Near perfect overlap!



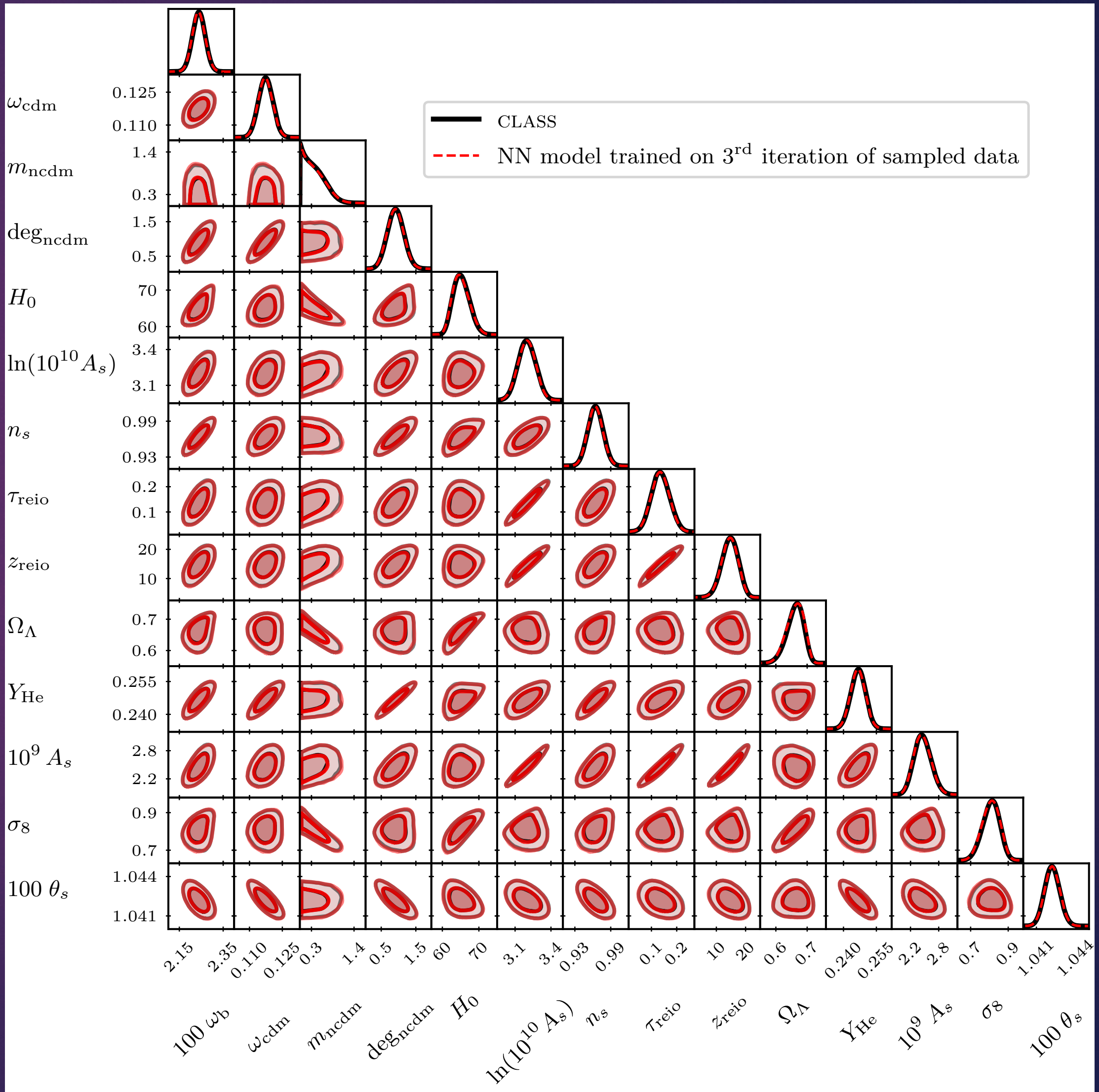
Results of iterative sampling

- DCDM model
  - Near perfect overlap!



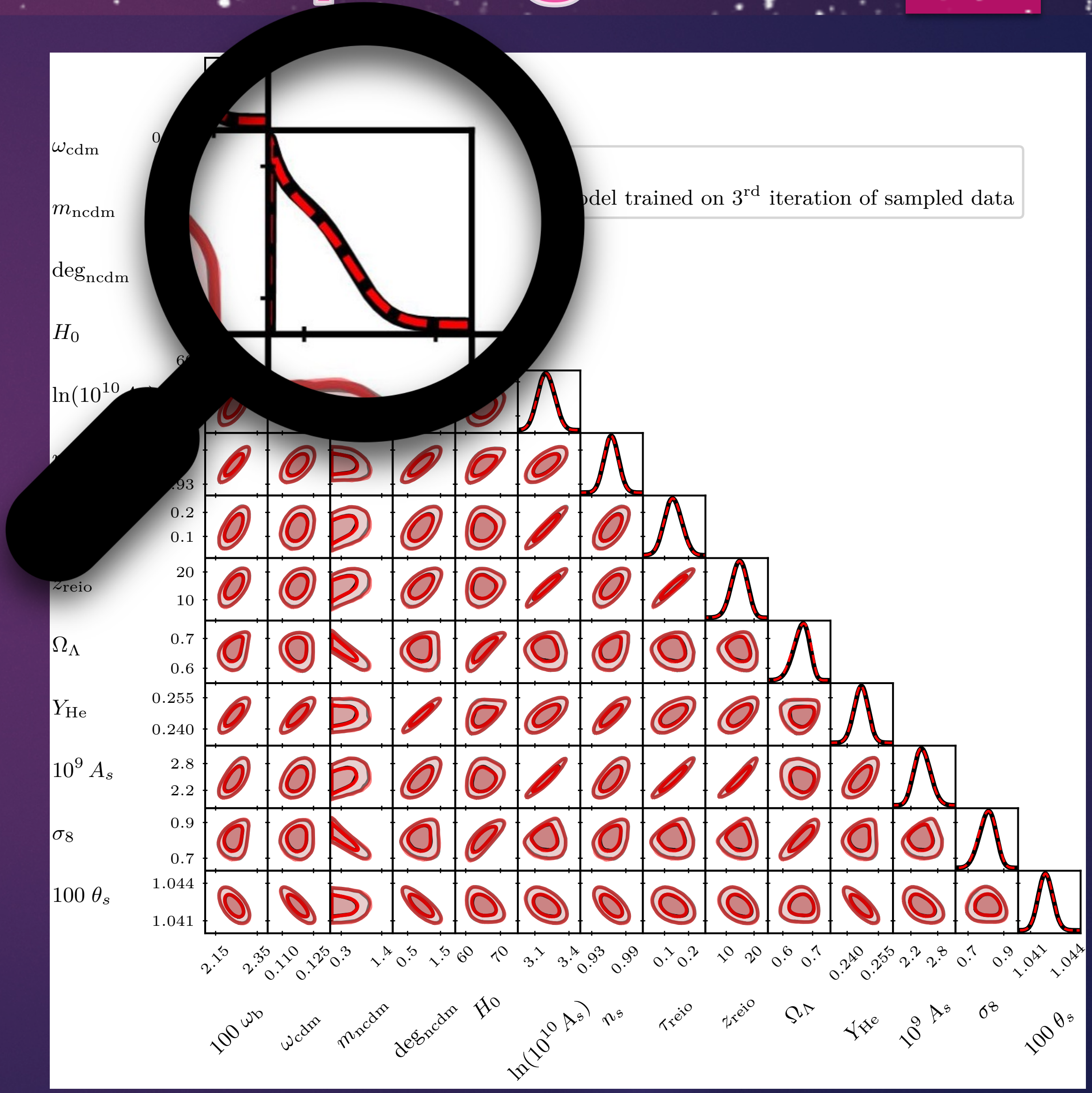
Results of iterative sampling

- DCDM model
  - Near perfect overlap!
- Massive neutrinos
  - Same excellence!

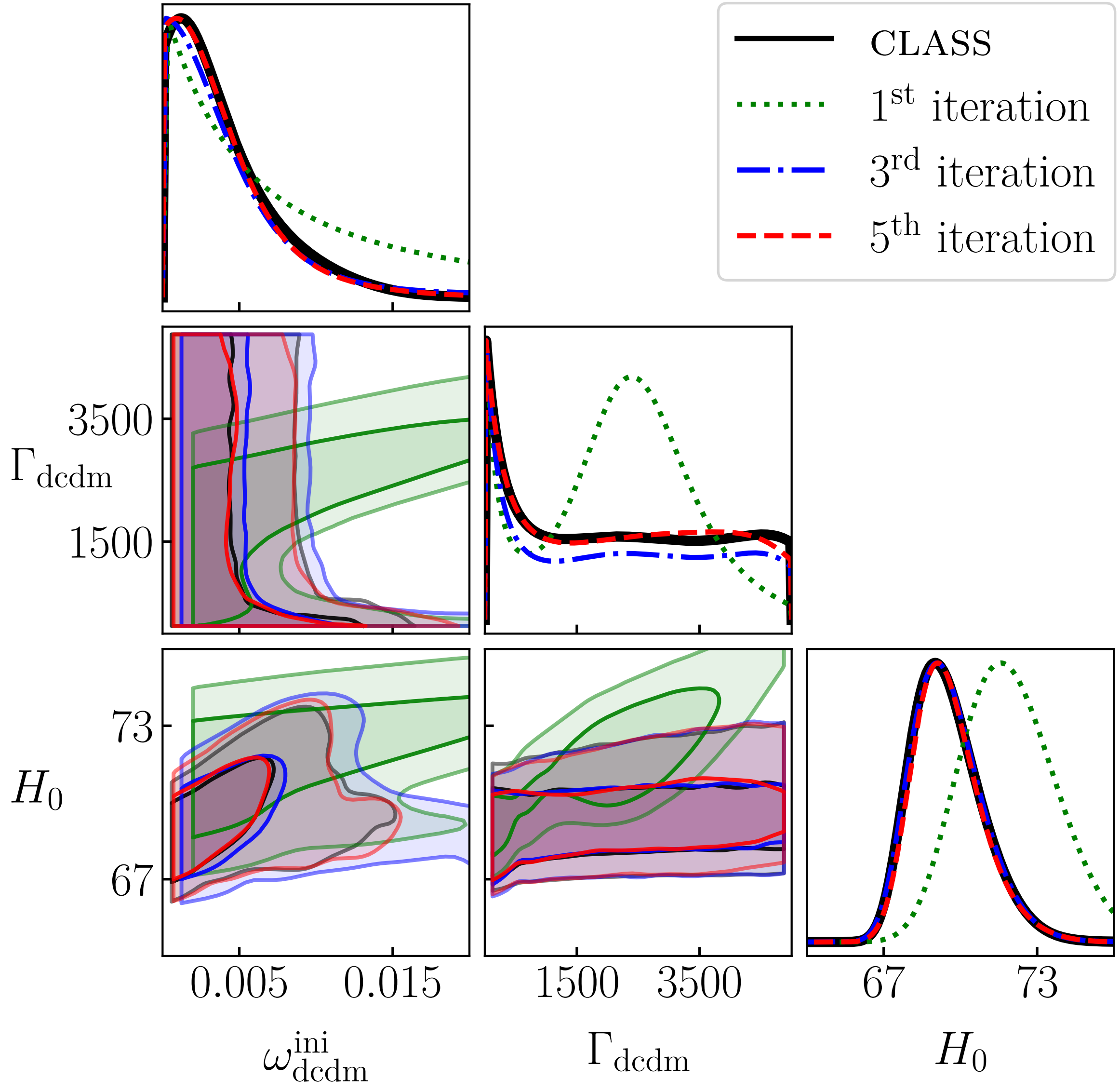


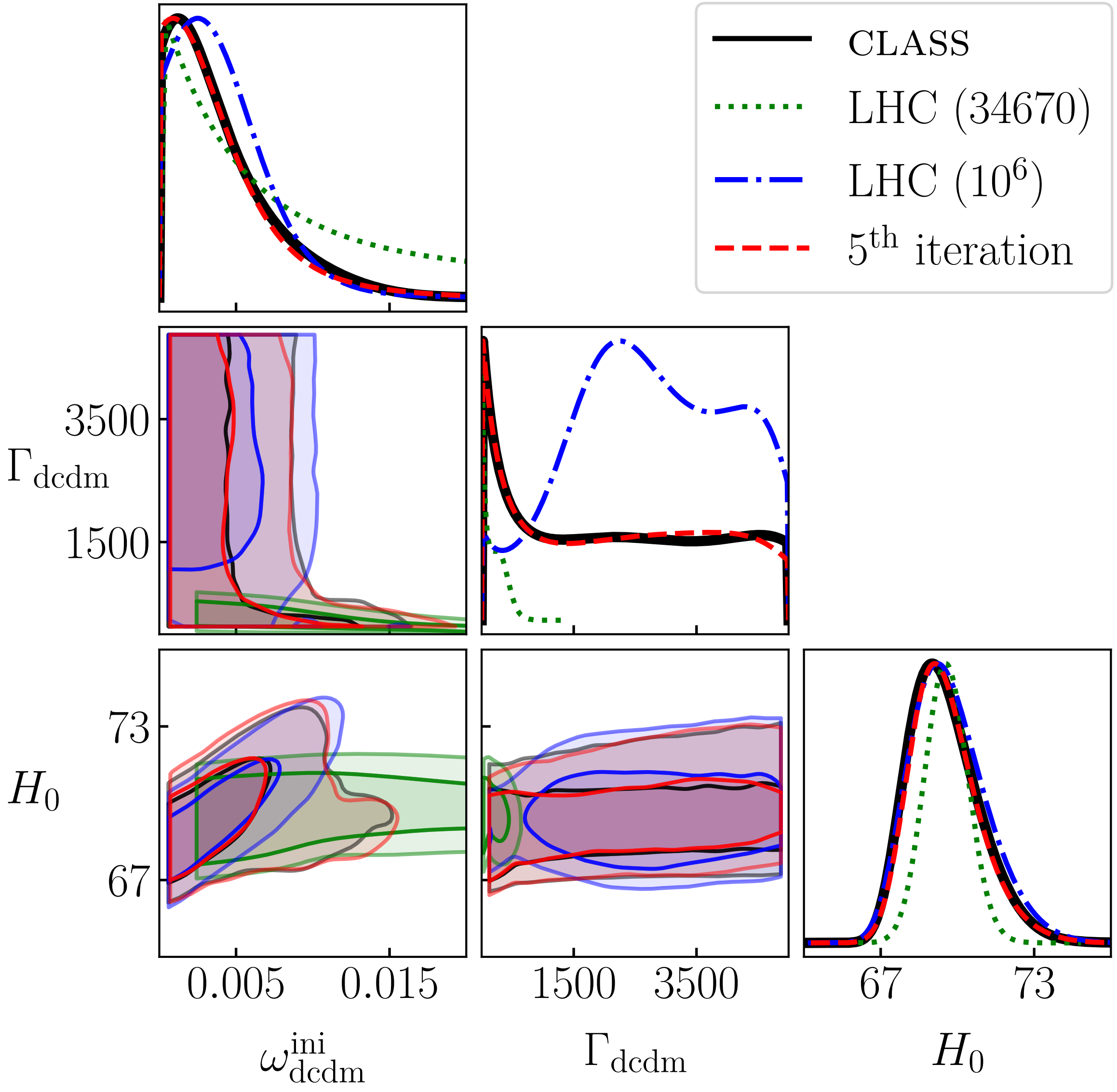
Results of iterative sampling

- DCDM model
  - Near perfect overlap!
- Massive neutrinos
  - Same excellence!



Results of iterative sampling





Current and future developments

Current and future developments

- Profile likelihoods
- GPU sampling
- Rewriting likelihoods to TensorFlow
- Emulating likelihoods

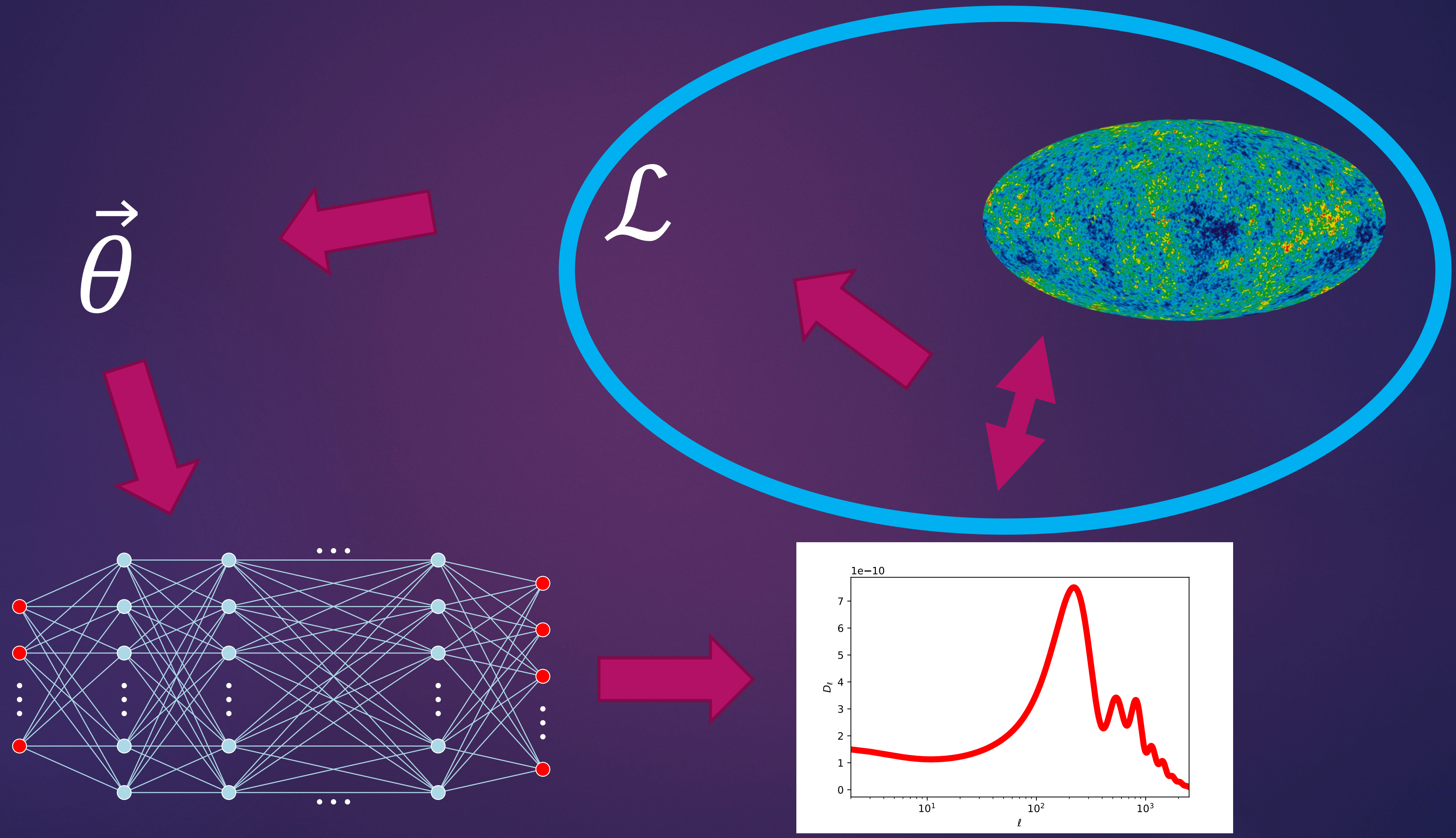
Current and future developments

- Profile likelihoods

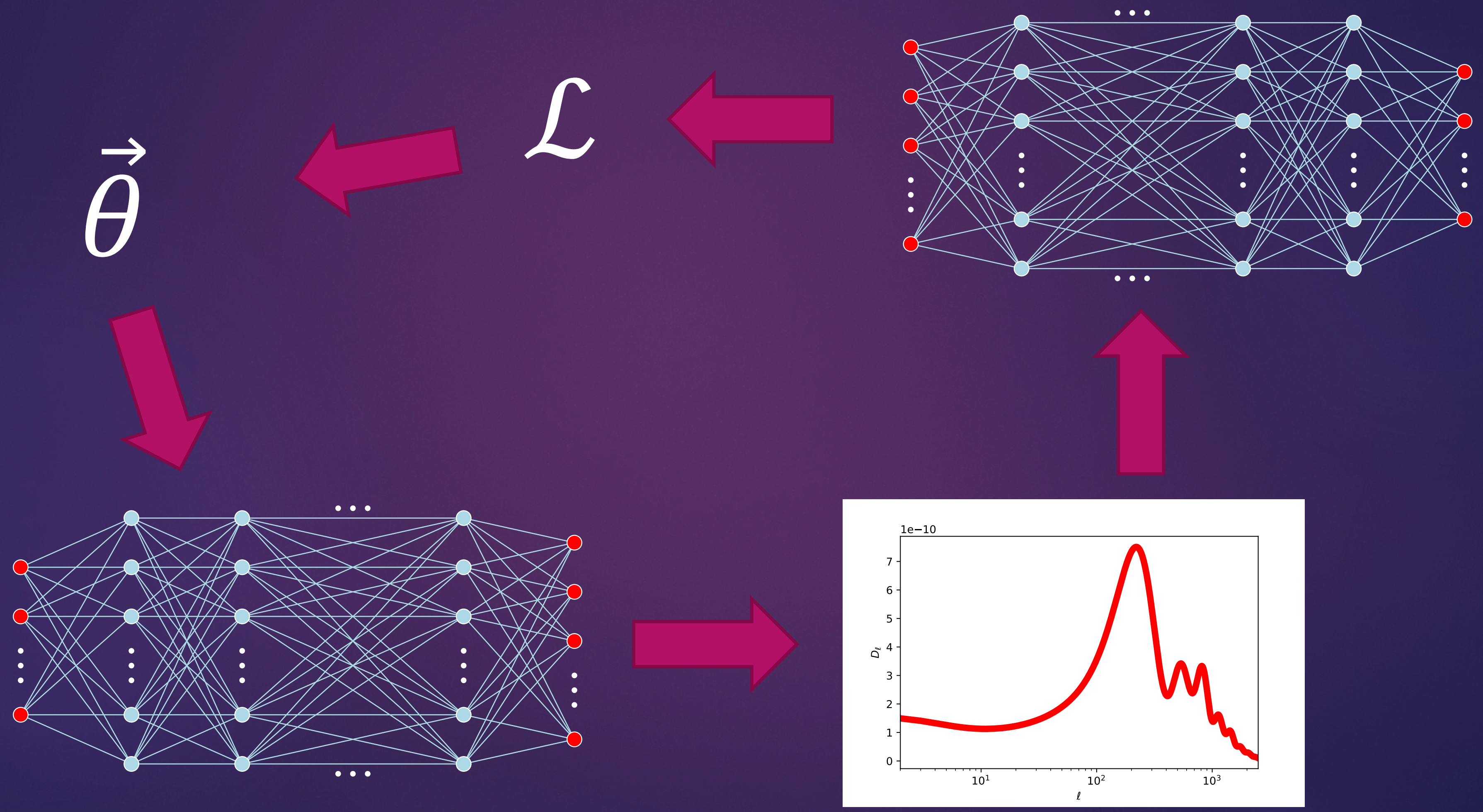
Emil’s talk today 17:30 (149)

- GPU sampling
- Rewriting likelihoods to TensorFlow
- Emulating likelihoods

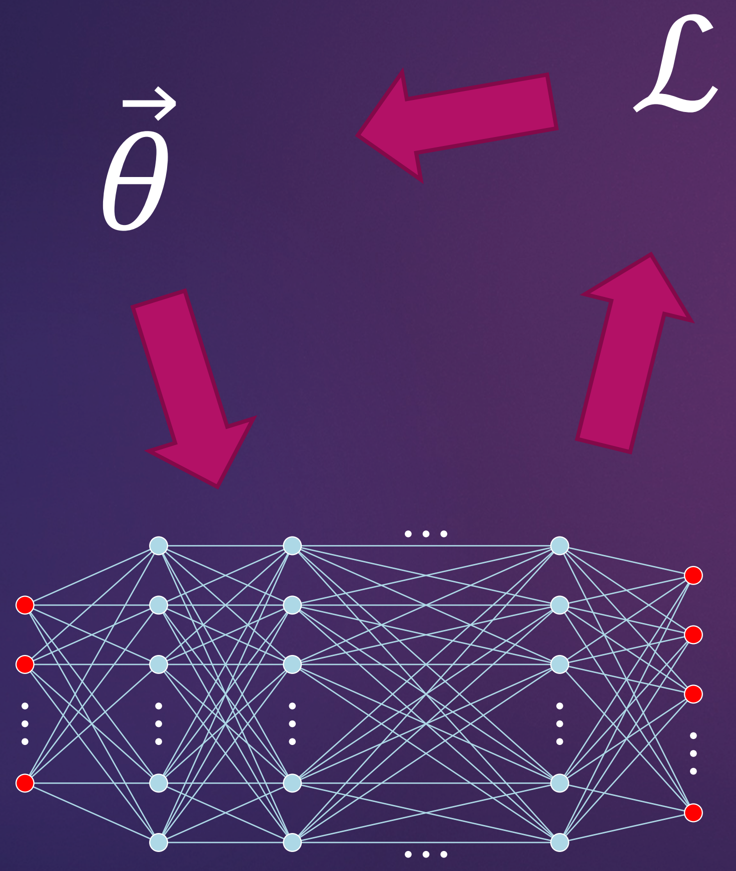
Current and future developments



Current and future developments



Current and future developments



Current and future developments

- Cobaya (Lewis and Torrado) plugin
- CAMB as alternative to CLASS
- Emulation of matter power spectra

Thank you for your time!