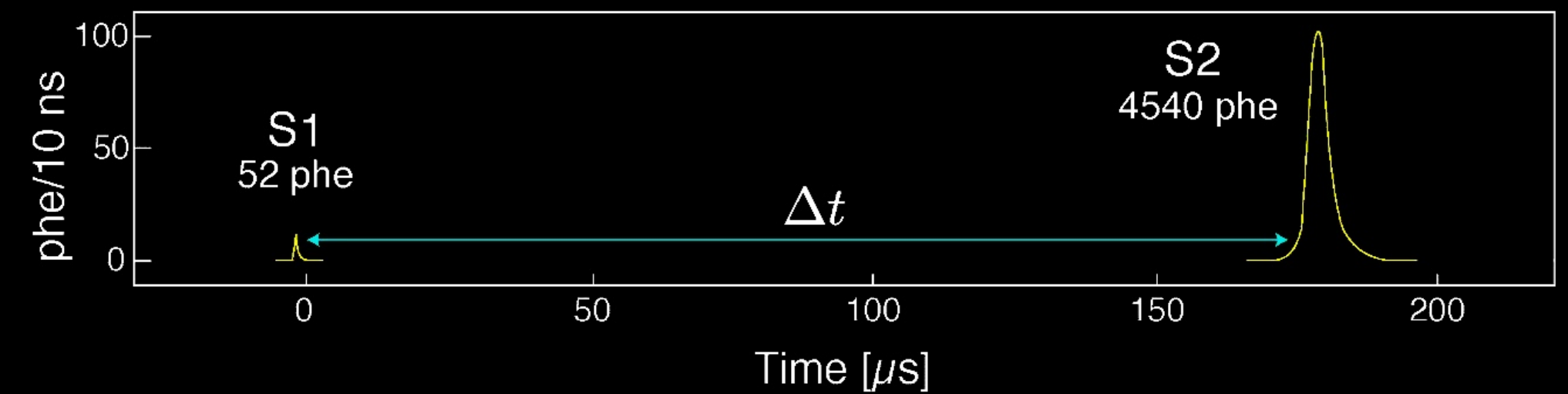
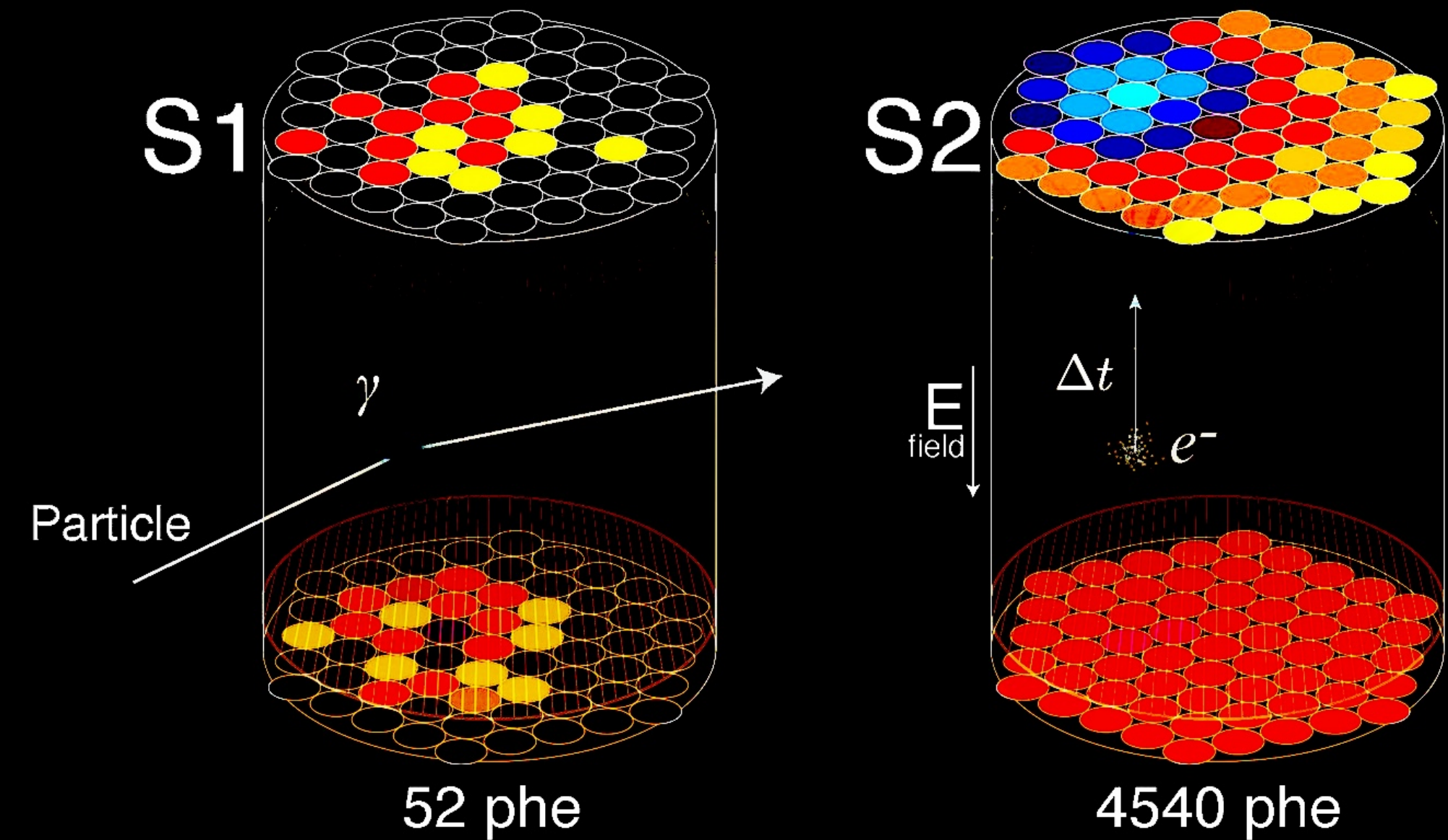
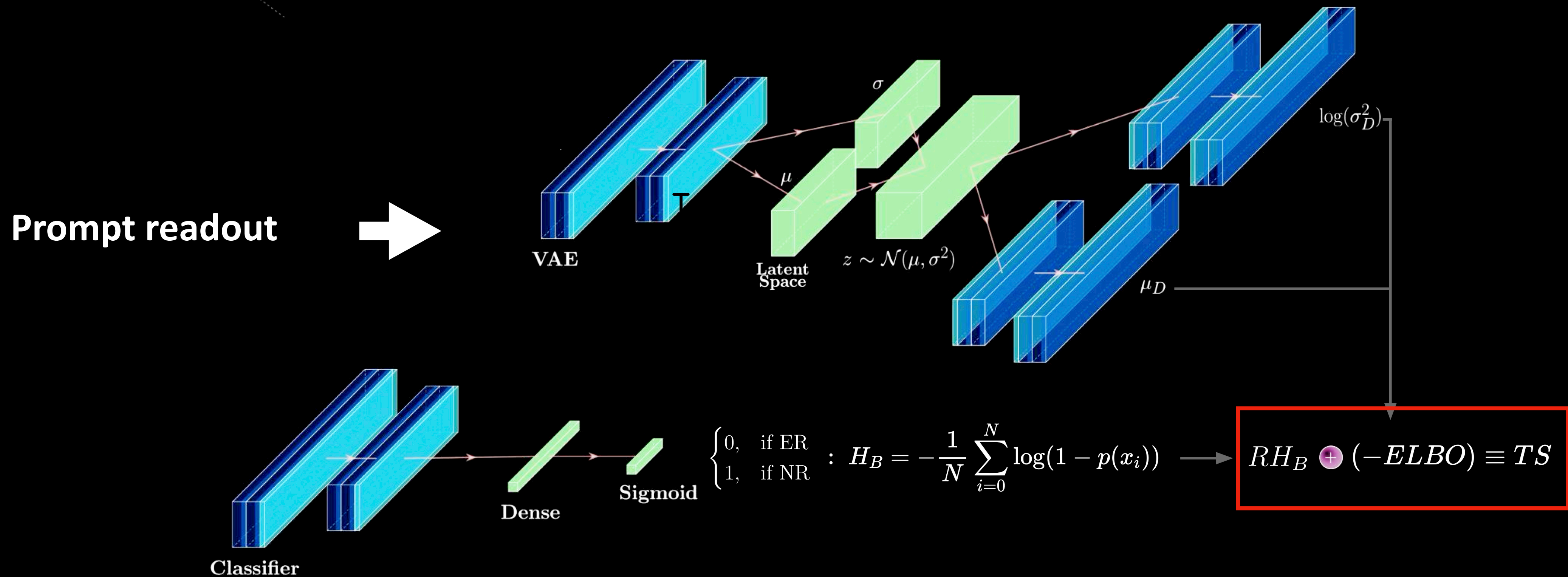


# Anomaly detection for dark matter direct detection

Aim: Use event-by-event prompt readout to train anomaly detector



$$ELBO = \underbrace{\frac{1}{2}\beta \sum_{j=1}^{L-\text{Space}} \left(1 + \log((\sigma_j)^2) - (\mu_j)^2 - (\sigma_j)^2\right)}_{\text{VAE Loss}} + \frac{1}{L} \sum_{l=1}^{\text{Batch}} \log \mathcal{N}(\mathbf{x}^{(l)} | \boldsymbol{\mu}_D^{(l)}, \log \sigma_D^{2(l)}).$$



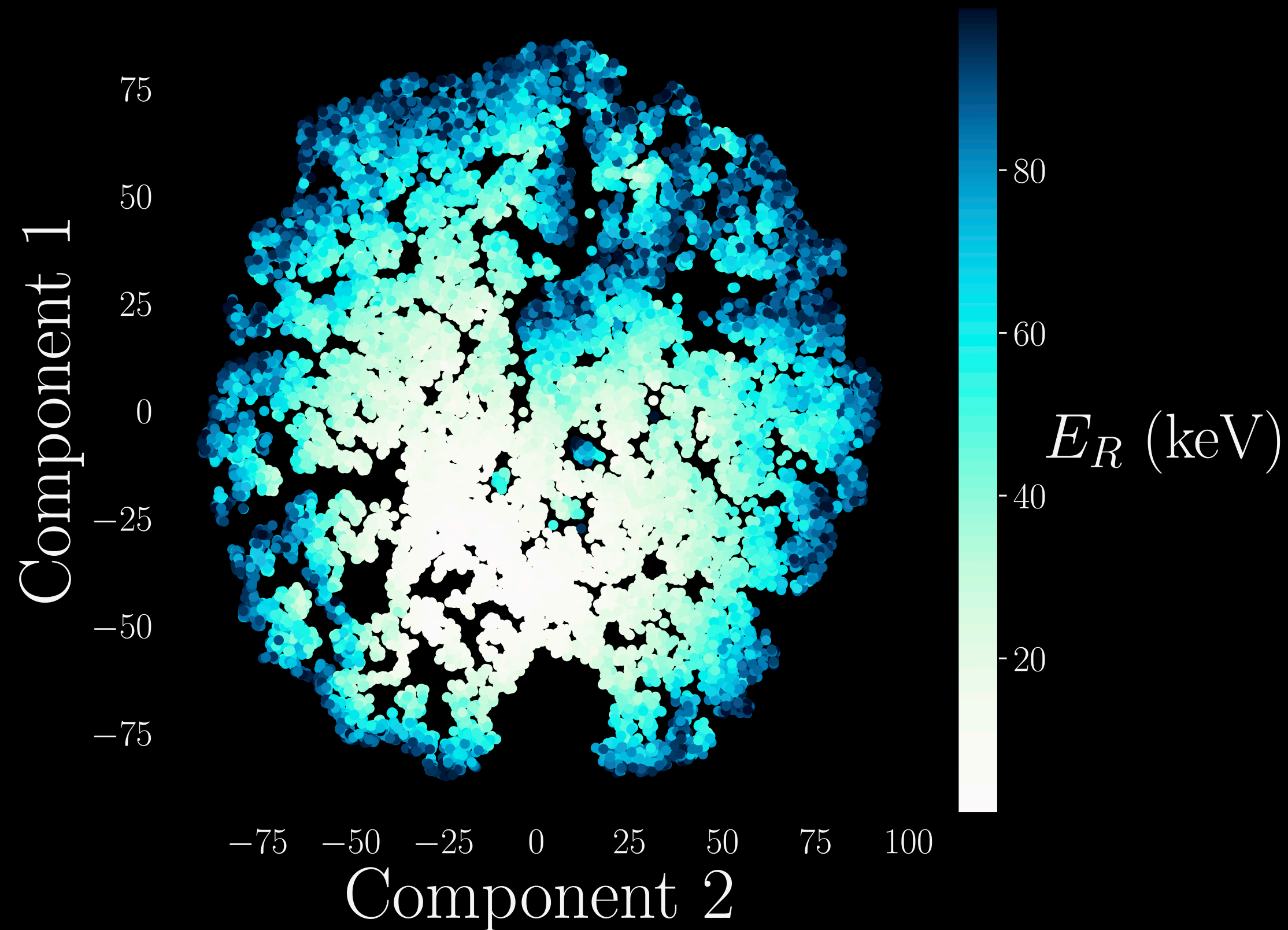
**Semi-supervised pipeline:** Combined anomaly function leverages discrete nature of detector events

Non-background events → Higher TS



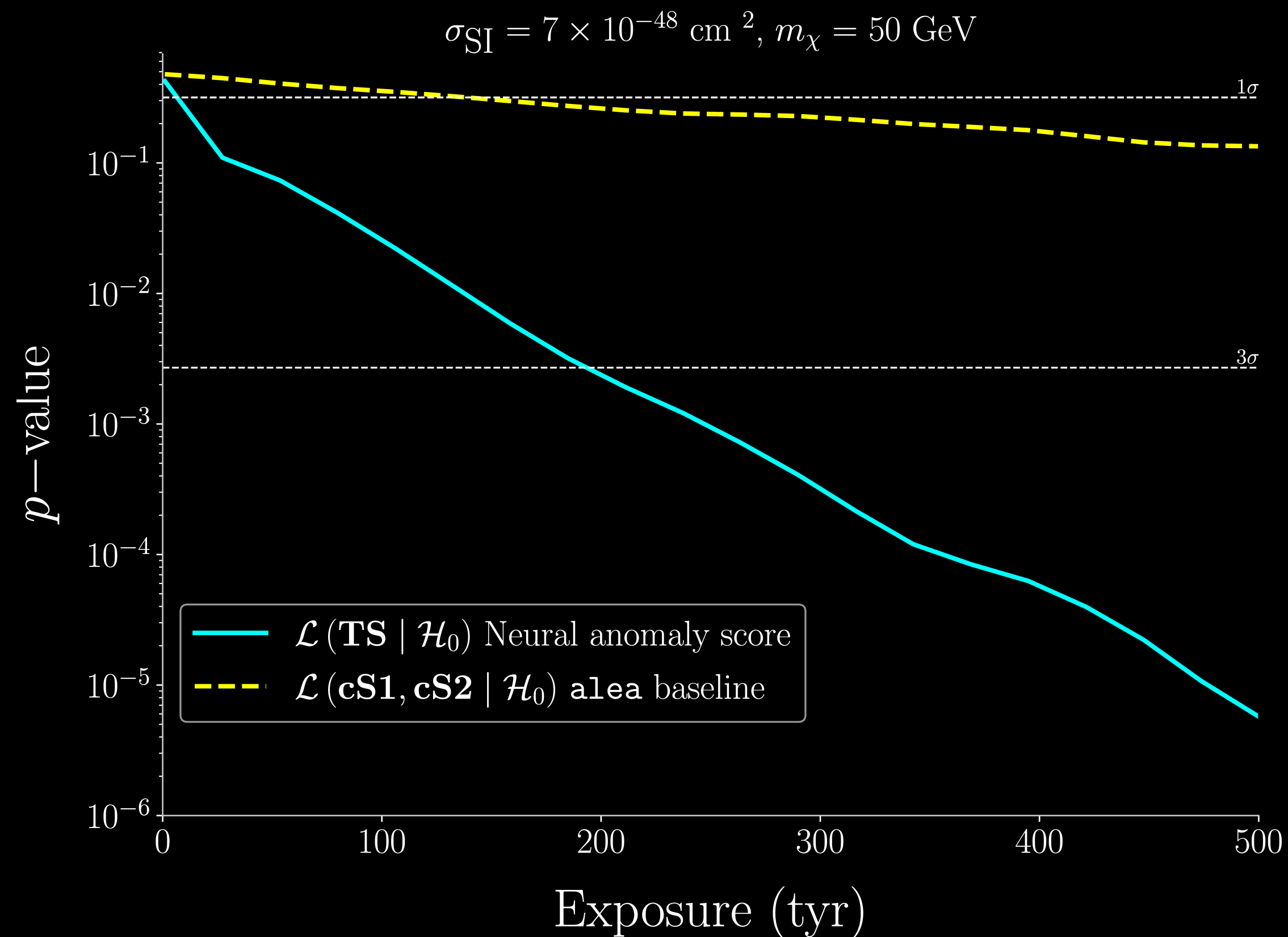
# What we get...

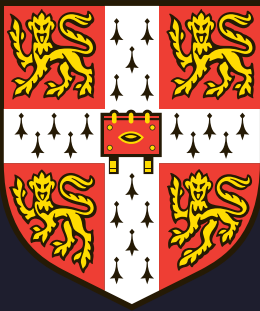
VAE learns event energy...



## Establish 1D two-sample test

- No nuisance parameters
- Spectral info encoded
- No dependence on high level summary statistics (cS1/cS2)
- Trivially implement calibration data directly
- Likelihood free and model independent





# Clustering Considerations for Nested Sampling

Adam Ormondroyd

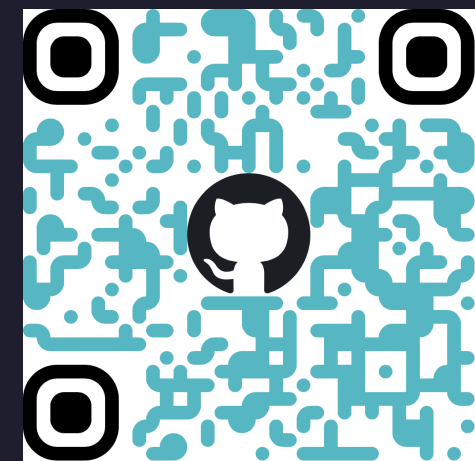
University of Cambridge

Kavli Institute for Cosmology

5.3.1 Cluster recognition

Any cluster recognition algorithm can be substituted at this point. One must take care that this is not run too often, or one runs the risk of adding a large overhead to the calculation. In practice, checking for clustering every  $\sim \mathcal{O}(n_{\text{live}})$  iterations is sufficient, since the prior will have only compressed by a factor  $e$ . We encourage users of POLYCHORD to experiment with their own preferred cluster recognition, in addition to that provided and described below.

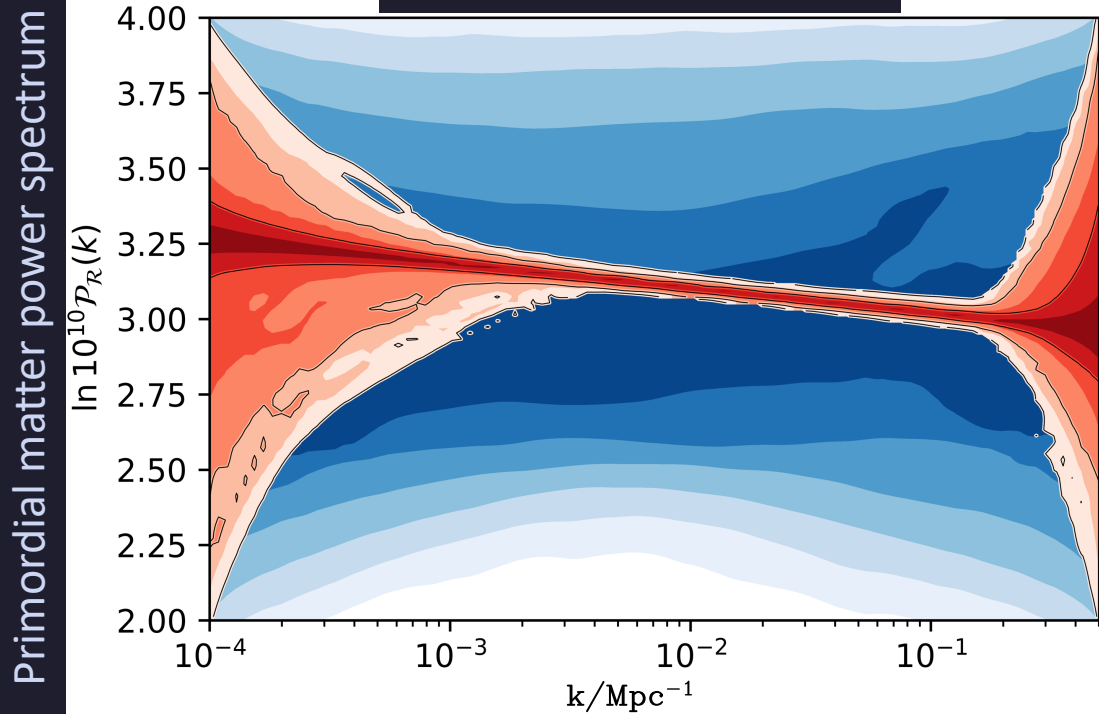
“PolyChord: next-generation nested sampling”  
 arXiv:1506.00171





# Motive: primordial matter power spectrum flex-knot reconstruction

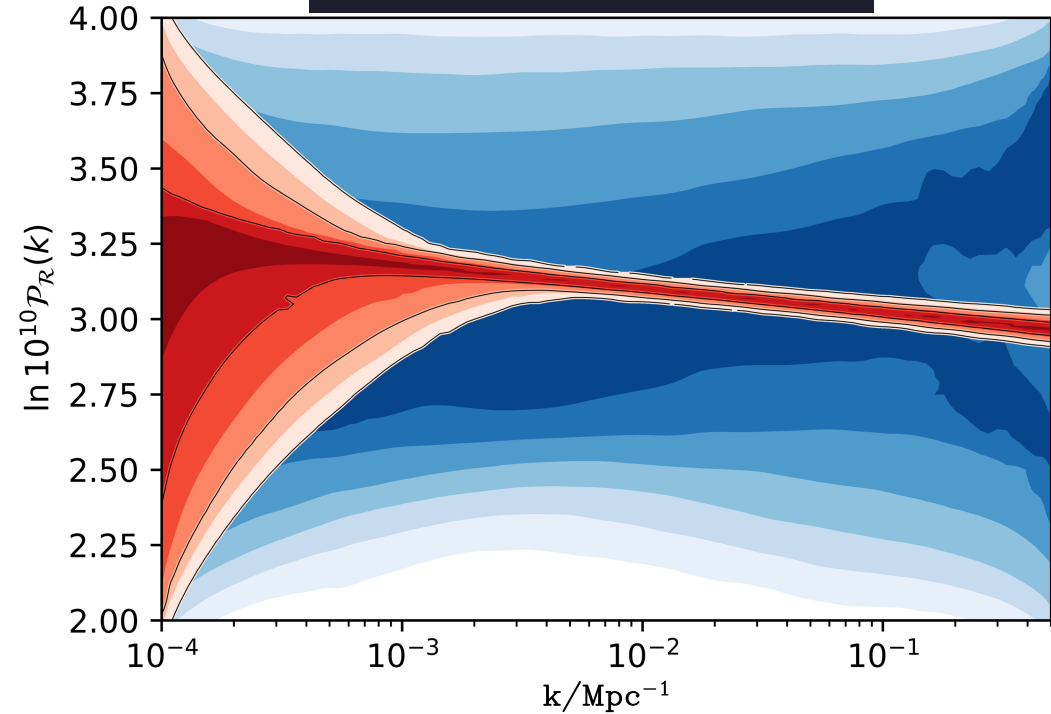
100 nested sampling live points (low resolution)



large scales  
(cosmic variance)

small scales  
(detector resolution)

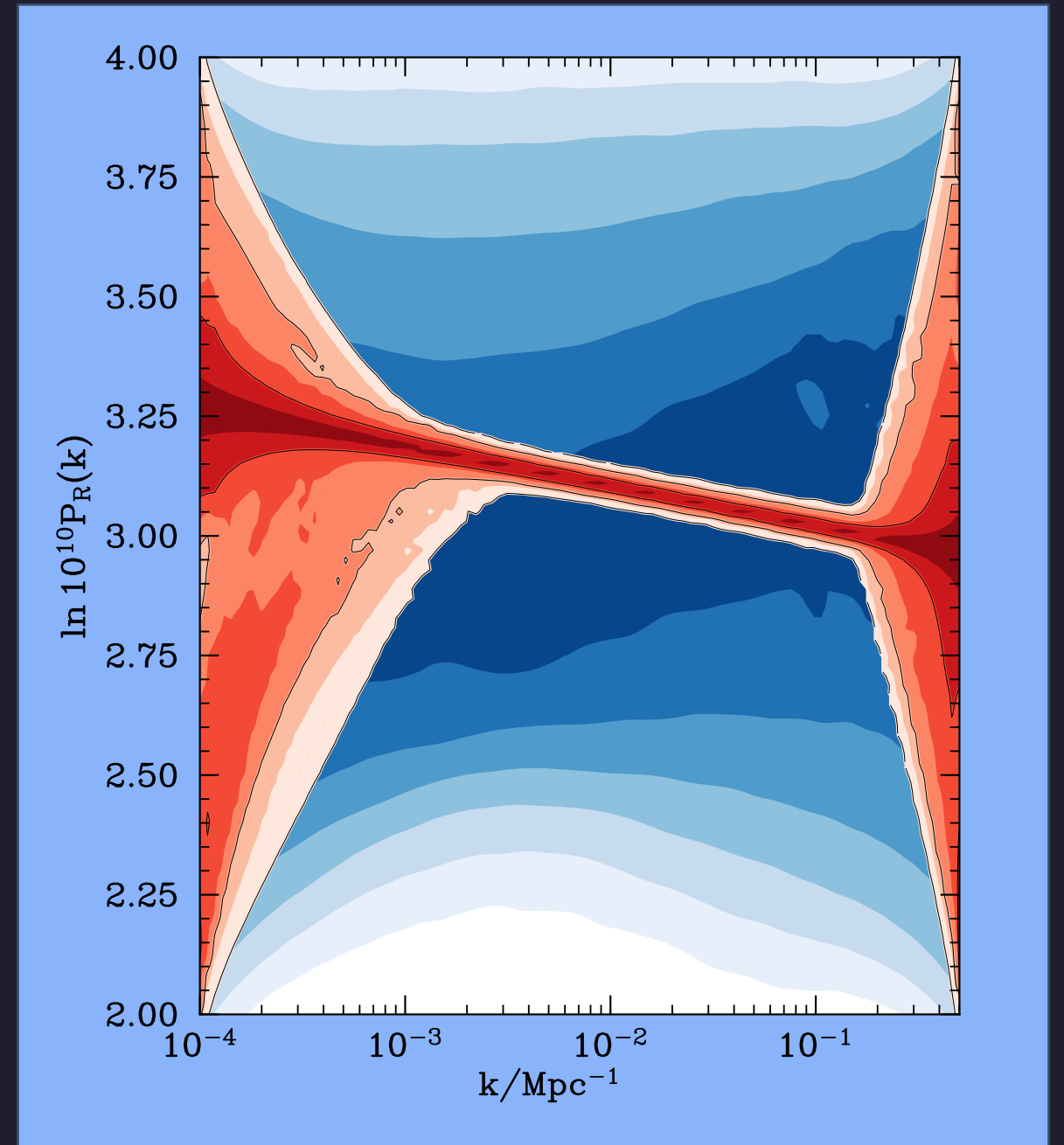
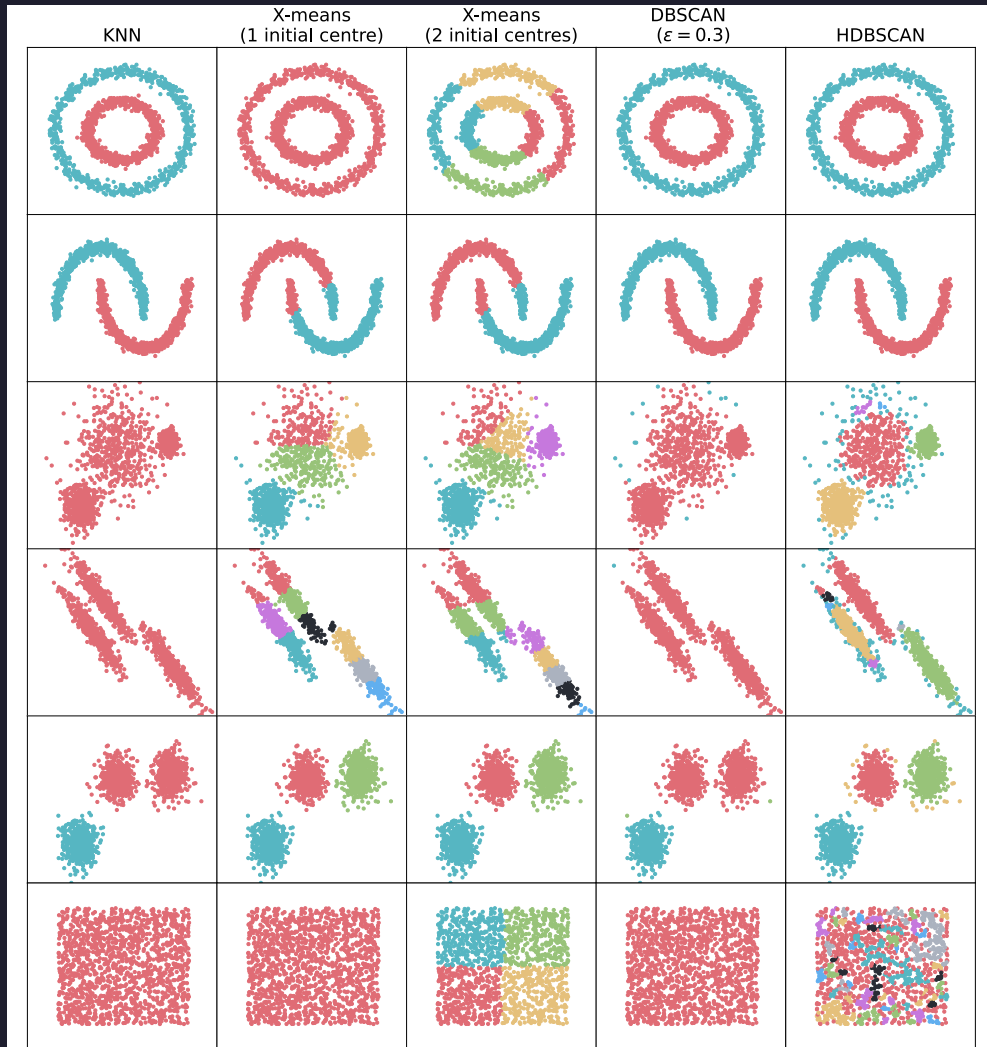
725 nested sampling live points (normal resolution)



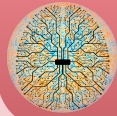
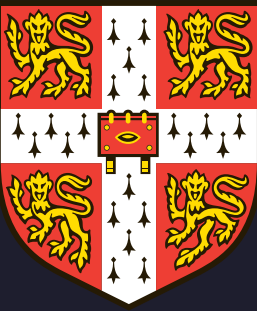
large scales  
(cosmic variance)

small scales  
(detector resolution)

# Clustering choices are abundant!







# Clustering Considerations for Nested Sampling

Adam Ormondroyd ano23@cam.ac.uk

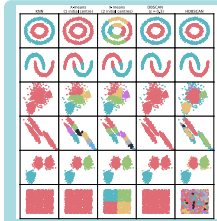
Kavli Institute for Cosmology · Cavendish Laboratory · University of Cambridge



Clustering algorithms are integral to multi-modal nested sampling, for both region-based samplers such as MultiNest, and chain-based samplers such as PolyChord. Robust identification clusters of live points is crucial for effective spawning of new live points, prior volume estimation and therefore the total evidence calculation. Reliable cluster detection also allows the calculation of the sub-evidences of each cluster, which may correspond to different physical phenomena. We have explored extensions to the clustering approach within PolyChord, and found that including correlation between the volume estimates of clusters increases the accuracy of evidence calculations. We show how different clustering methods affect a reconstruction of the cosmological primordial matter power spectrum  $\mathcal{P}_{\mathcal{R}}(k)$ .



## Clustering choice

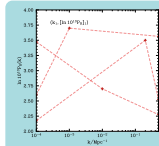


2D demonstration of PolyChord's K-Nearest Neighbours, pyclustering's X-means and scikit-learn's (H)DBSCAN.

The nested sampling [1, 2] algorithm PolyChord was originally advertised suggesting that the user should experiment with their favourite cluster-identifying methods, but provides no guidance on how to do so [3, 4]. We have added an interface which allows the user to substitute any clustering strategy at the Python level, allowing for easier experimentation with alternatives such as the selection provided by scikit-learn and pyclustering [5, 6]. Some algorithms are better suited than others to identifying posterior modes of nested sampling live points, for example K-means and spectral clustering need to be told the number of clusters to look for, others may not assign every point to a cluster. Some approaches find clusters where there are none!

## Application to cosmology $\mathcal{P}_{\mathcal{R}}(k)$

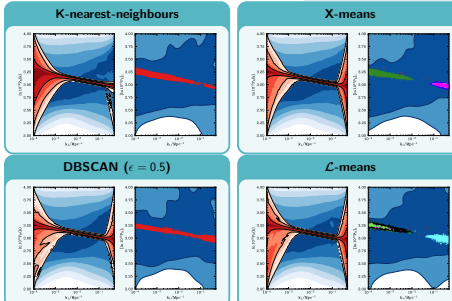
This investigation was initially motivated by a reconstruction of the primordial matter power spectrum  $\mathcal{P}_{\mathcal{R}}(k)$  using flexknots [7, 8], and the Planck 2018 likelihoods [9, 10]. Planck measured the  $C_l$  multipole range  $1 \leq l \leq 7000$ , corresponding to  $10^{-4} \leq k/\text{Mpc}^{-1} \leq 10^{-0.3} [11]$ . Flexknots are parameterisations of 1D functions, consisting of a series of splines (in this case linear) joined at knots. The number and positions of knots are determined by the data, which can be performed by either combining several runs with fixed number of knots, or the number of knots being itself a parameter. In the former case, we noticed that with three knots only the mode with the central knot towards the left was being fully explored. PolyChord's native K-nearest-neighbours clustering was unable to separate the positions of the central knot into two distinct clusters, so we explored both off-the-shelf clustering algorithms and an approach which includes likelihood information,  $\mathcal{L}$ -means.



Examples of  $\mathcal{P}_{\mathcal{R}}(k)$  flexknots with three knots each.

### Prototype $\mathcal{L}$ -means algorithm

Partition  $X$  into  $n$  sets and calculate the likelihood at the centres of the two clusters. Find the midpoint of the two centres. Calculate the likelihood at the midpoint. If the midpoint has greater likelihood than either cluster, all points are within the same cluster and iteratively apply  $\mathcal{L}$ -means to each cluster.



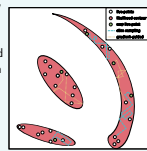
K-nearest-neighbours and DBSCAN fail to separate the central knot into two clusters, while X-means does so reliably.  $\mathcal{L}$ -means is able to separate the central knot into multiple clusters, but also tends to over-cluster. Functional posterior plots were created using fgviznx [12], and scatterplots showing clusters were created using a development branch of anaesthetic [13].

## Classes of Nested Sampler

Nested sampling algorithms have two main strategies for sampling new points from the prior:

### Chain-based

A sufficiently long random walk within the likelihood contour from an existing live point will generate a new live point.

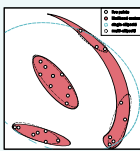


PolyChord first uses the covariance matrix of the live points to whiten the space, then performs Neal's slice sampling along orthogonal directions in that space [3, 4]. GGNIS (gradient-guided nested sampling) also implements both Neal's and Hamiltonian slice sampling, along with uniform sampling and random walks [14].

Multiple-modal problems render contour whitening ineffective, and a strategy is required to decide from which mode to sample since a random walk cannot pass through the likelihood contour.

### Region-based

Region-based samplers construct regions around the live points to approximate the likelihood contour.



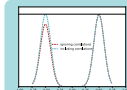
MultiNest constructs a series of ellipsoids [15–17]. These regions are usually expanded by a numerical factor to improve their chances of fully enclosing the likelihood contour, then a point is rejection-sampled from them. The curse of dimensionality means that these techniques are only effective up to  $\mathcal{O}(10)$  dimensions, as either rejection sampling becomes inefficient, or the expansion factor would have to be so low that significant regions of the likelihood contour would be missed. Clustering can be used to separate the live points into discrete regions, rather than a single sparse region.

Hybrid methods combine the two approaches in an attempt to alleviate the dimensionality scaling of region-based methods, while reducing the number of likelihood evaluations made outside the contour [14, 18–20].

## Correlated cluster volumes (in progress!)

When a cluster  $p$  is divided, the remaining prior volume  $X_p$  is divided among its subclusters  $X_i$ . However, in nested sampling, we do not know the precise prior volumes, only expectation values and errors. This is divided according to the proportion of live points in each cluster  $n_i$ :

$$\bar{X}_i = \frac{n_i}{n_p} \bar{X}_p, \quad \bar{X}_i^2 = \frac{n_i(n_i+1)}{n_p(n_p+1)} \bar{X}_p^2, \\ \overline{X_i X_j} = \frac{n_i n_j}{n_p(n_p+1)} \bar{X}_p^2.$$



Accounting for prior volume correlations may provide more accurate posteriors.

Since  $\overline{X_i X_j} \neq \bar{X}_i \bar{X}_j$ , the error on the prior volumes estimates of each cluster are correlated. This is important when deciding from which cluster to sample; currently PolyChord chooses a cluster proportionally to its prior volume  $\bar{X}_i$ , but neglects their correlation. We are experimenting with drawing a set of  $X_i$  from their joint distribution before each live point is generated, which has shown promise with symmetric multi-modal likelihoods.

## References

[1] John Skilling, 'Nested Sampling', in Robert Fisher, Roland Pfeifer, and Udo von Toussaint, editors, Bayesian Inference and Maximum Entropy Methods in Science and Engineering, the International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, volume 518 of *Frontiers in Physics*, Springer, Berlin, 2004. MR2060009

[2] John Skilling, 'Bayesian Adaptive Nested Sampling', in Robert Fisher, Roland Pfeifer, and Udo von Toussaint, editors, Bayesian Inference and Maximum Entropy Methods in Science and Engineering, volume 518 of *Frontiers in Physics*, Springer, Berlin, 2004. MR2060010

[3] John Skilling, 'Bayesian Adaptive Nested Sampling', in Robert Fisher, Roland Pfeifer, and Udo von Toussaint, editors, Bayesian Inference and Maximum Entropy Methods in Science and Engineering, volume 518 of *Frontiers in Physics*, Springer, Berlin, 2004. MR2060011

[4] John Skilling, 'Bayesian Adaptive Nested Sampling', in Robert Fisher, Roland Pfeifer, and Udo von Toussaint, editors, Bayesian Inference and Maximum Entropy Methods in Science and Engineering, volume 518 of *Frontiers in Physics*, Springer, Berlin, 2004. MR2060012

[5] John Skilling, 'Bayesian Adaptive Nested Sampling', in Robert Fisher, Roland Pfeifer, and Udo von Toussaint, editors, Bayesian Inference and Maximum Entropy Methods in Science and Engineering, volume 518 of *Frontiers in Physics*, Springer, Berlin, 2004. MR2060013

[6] John Skilling, 'Bayesian Adaptive Nested Sampling', in Robert Fisher, Roland Pfeifer, and Udo von Toussaint, editors, Bayesian Inference and Maximum Entropy Methods in Science and Engineering, volume 518 of *Frontiers in Physics*, Springer, Berlin, 2004. MR2060014

[7] Adam Ormondroyd, 'Flexknots: A Flexible, Efficient, and Accurate Method for Modelling and Predicting the Primordial Matter Power Spectrum', in *Journal of Cosmology and Astroparticle Physics*, volume 2018, page 018, 2018. MR3840000

[8] Adam Ormondroyd, 'Flexknots: A Flexible, Efficient, and Accurate Method for Modelling and Predicting the Primordial Matter Power Spectrum', in *Journal of Cosmology and Astroparticle Physics*, volume 2018, page 018, 2018. MR3840001

[9] Planck Collaboration, 'Planck 2018 Results. I. Overview and the cosmological framework', in *Journal of Cosmology and Astroparticle Physics*, volume 2018, page 029, 2018. MR3840002

[10] Planck Collaboration, 'Planck 2018 Results. II. Constraints on the cosmological model', in *Journal of Cosmology and Astroparticle Physics*, volume 2018, page 020, 2018. MR3840003

[11] Planck Collaboration, 'Planck 2018 Results. III. Cosmological parameters', in *Journal of Cosmology and Astroparticle Physics*, volume 2018, page 029, 2018. MR3840004

[12] fgviznx, 'fgviznx: A Python package for functional posterior plotting', 2022. MR3840005

[13] anaesthetic, 'anaesthetic: A Python package for nested sampling', 2022. MR3840006

[14] Adam Ormondroyd, 'A Hybrid Method for Nested Sampling', in *Journal of Cosmology and Astroparticle Physics*, volume 2022, page 018, 2022. MR3840007

[15] John Skilling, 'Bayesian Adaptive Nested Sampling', in Robert Fisher, Roland Pfeifer, and Udo von Toussaint, editors, Bayesian Inference and Maximum Entropy Methods in Science and Engineering, volume 518 of *Frontiers in Physics*, Springer, Berlin, 2004. MR2060015

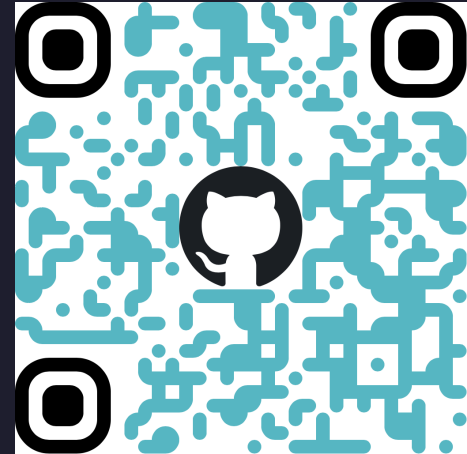
[16] John Skilling, 'Bayesian Adaptive Nested Sampling', in Robert Fisher, Roland Pfeifer, and Udo von Toussaint, editors, Bayesian Inference and Maximum Entropy Methods in Science and Engineering, volume 518 of *Frontiers in Physics*, Springer, Berlin, 2004. MR2060016

[17] John Skilling, 'Bayesian Adaptive Nested Sampling', in Robert Fisher, Roland Pfeifer, and Udo von Toussaint, editors, Bayesian Inference and Maximum Entropy Methods in Science and Engineering, volume 518 of *Frontiers in Physics*, Springer, Berlin, 2004. MR2060017

[18] Adam Ormondroyd, 'A Hybrid Method for Nested Sampling', in *Journal of Cosmology and Astroparticle Physics*, volume 2022, page 018, 2022. MR3840008

[19] Adam Ormondroyd, 'A Hybrid Method for Nested Sampling', in *Journal of Cosmology and Astroparticle Physics*, volume 2022, page 018, 2022. MR3840009

[20] Adam Ormondroyd, 'A Hybrid Method for Nested Sampling', in *Journal of Cosmology and Astroparticle Physics*, volume 2022, page 018, 2022. MR3840010



# Enhancing Robustness

BSM Parameter Inference with n1D-CNN and Novel Data Augmentation

Koay Yong Sheng

Stefano Moretti, Prashant Singh, Harri Waltari

## Keywords:

- Parameter Inference
- Improving CNNs
- Data Augmentation
- Signal Region Evaluation

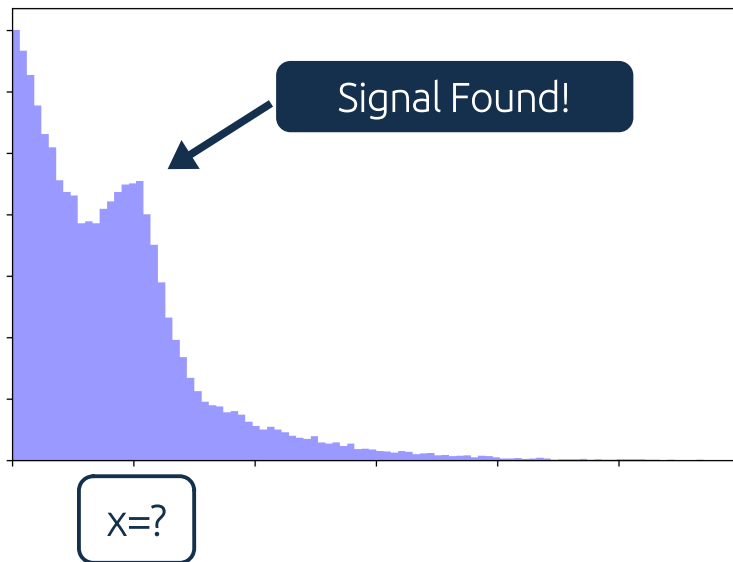


UPPSALA  
UNIVERSITET

Poster  
**A:14**



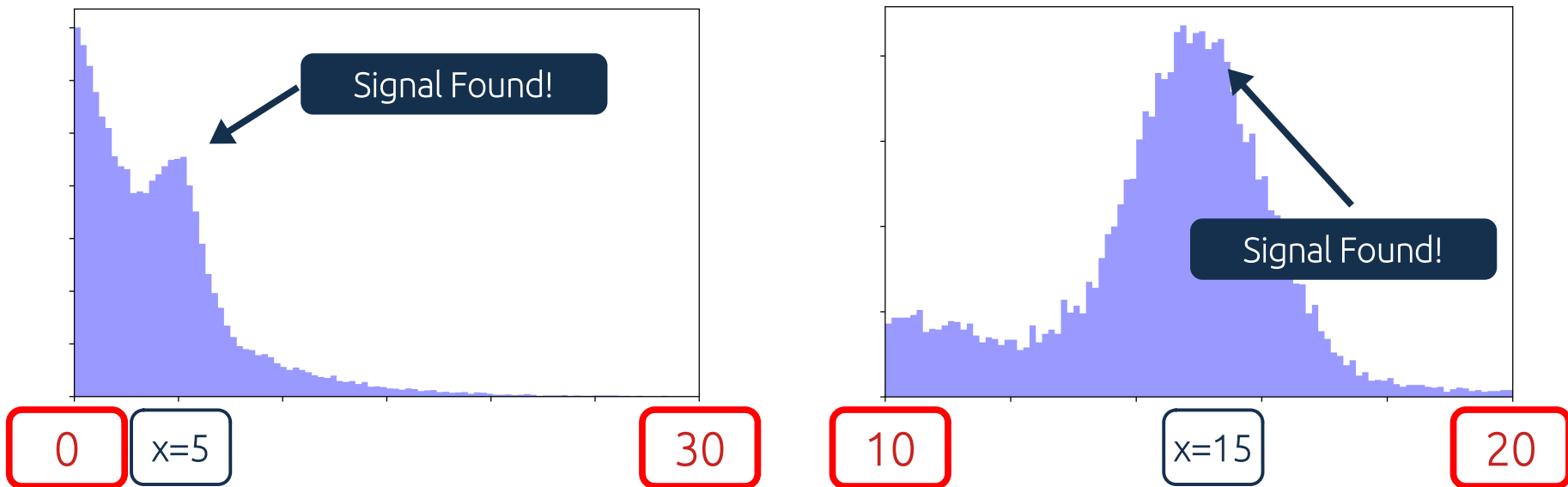
Given the following images, what is  $x_{\text{signal}}$ ?



Impossible! No scale of reference!  
Recall CNN can only see images/pixels

(Convolutional Neural Network)

# Auxiliary Information Important !



CNNs need **Aux Info** to infer from Different Signal Regions





# n-1DCNN-wA

(Extending 1DCNN with Auxiliary Inputs)

## BSM Inference

(Beyond Standard Model)

from

ANY SIGNAL REGION

ANY FIXED NO. OF OBSERVABLES

## Poster A:14

# Enhancing Robustness

BSM Parameter Inference with n1D-CNN and Novel Data Augmentation

Koay Yong Sheng

## BUMP HUNTS

1 Observable

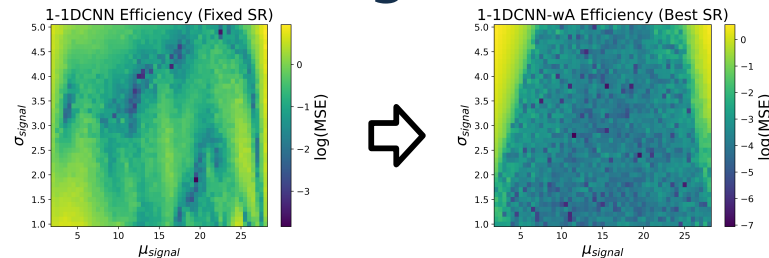
## BUMP-LESS MONO-X SIGNALS

(Dark Matter Search)

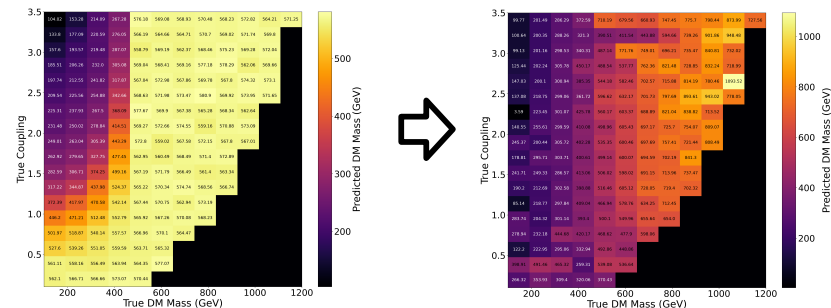
3 Observables  
(and more with Mono-Z)

## BONUS ADVANTAGES

## Orders of Magnitude Better



## Expands Infer-able Parameters Space



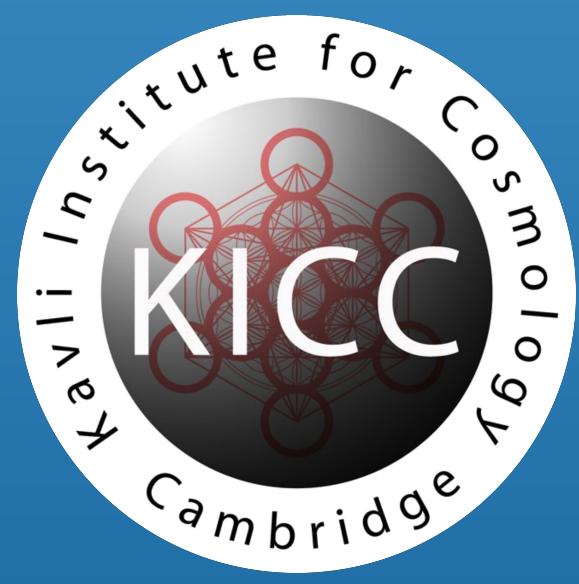
## Data Augmentation

(that grows with  $N_{\text{Signal Region}}$  and  $N_{\text{Observables}}$ )

+

## Signal Region Evaluator

(Which is the Best Signal Region?)



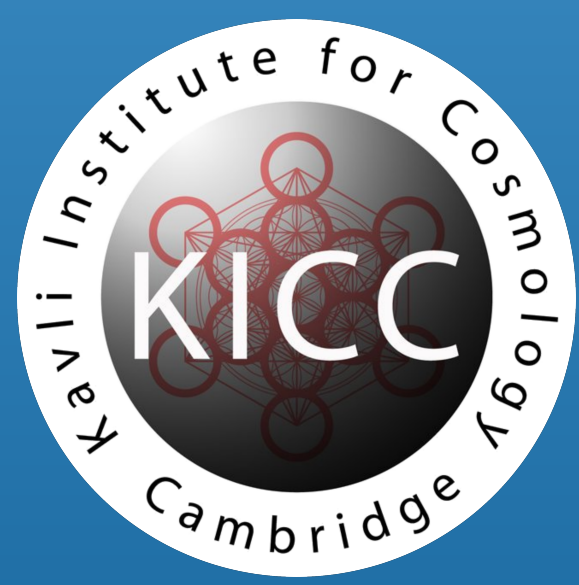
# Fully Bayesian Forecasts with Neural Bayes Ratio Estimation (Poster #9, arXiv:2309.06942)



Thomas Gessey-Jones (tg400@cam.ac.uk) and Will Handley (wh260@cam.ac.uk)

Problem - Data Marginalized Forecasts for Bayesian Model Comparison are Typically Infeasible

**Nested Sampling:** Generally applicable but too slow to explore data space  
**Savage–Dickey forecasts:** Restrictive assumptions and require nested models  
Both require explicit likelihoods.



# Fully Bayesian Forecasts with Neural Bayes Ratio Estimation (Poster #9, arXiv:2309.06942)

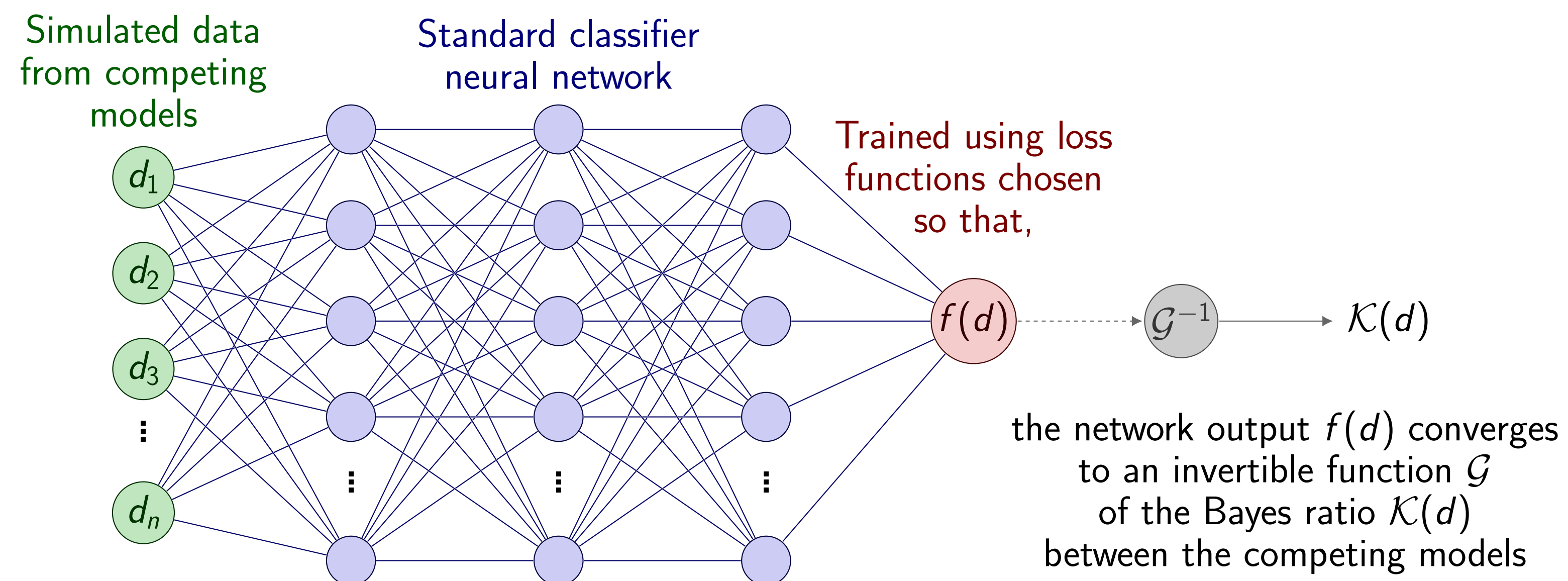


Thomas Gessey-Jones (tg400@cam.ac.uk) and Will Handley (wh260@cam.ac.uk)

## Problem - Data Marginalized Forecasts for Bayesian Model Comparison are Typically Infeasible

**Nested Sampling:** Generally applicable but too slow to explore data space  
**Savage–Dickey forecasts:** Restrictive assumptions and require nested models  
Both require explicit likelihoods.

## Solution - Neural Bayes Ratio Estimation (Evidence Network)

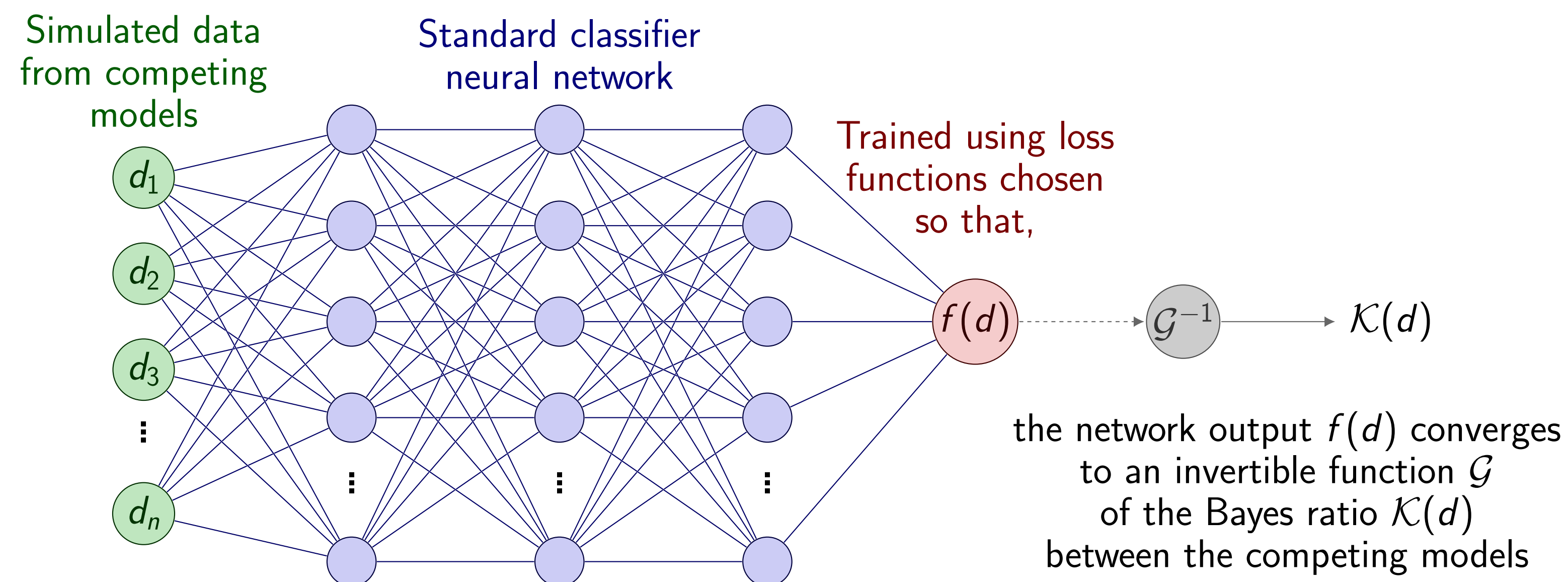




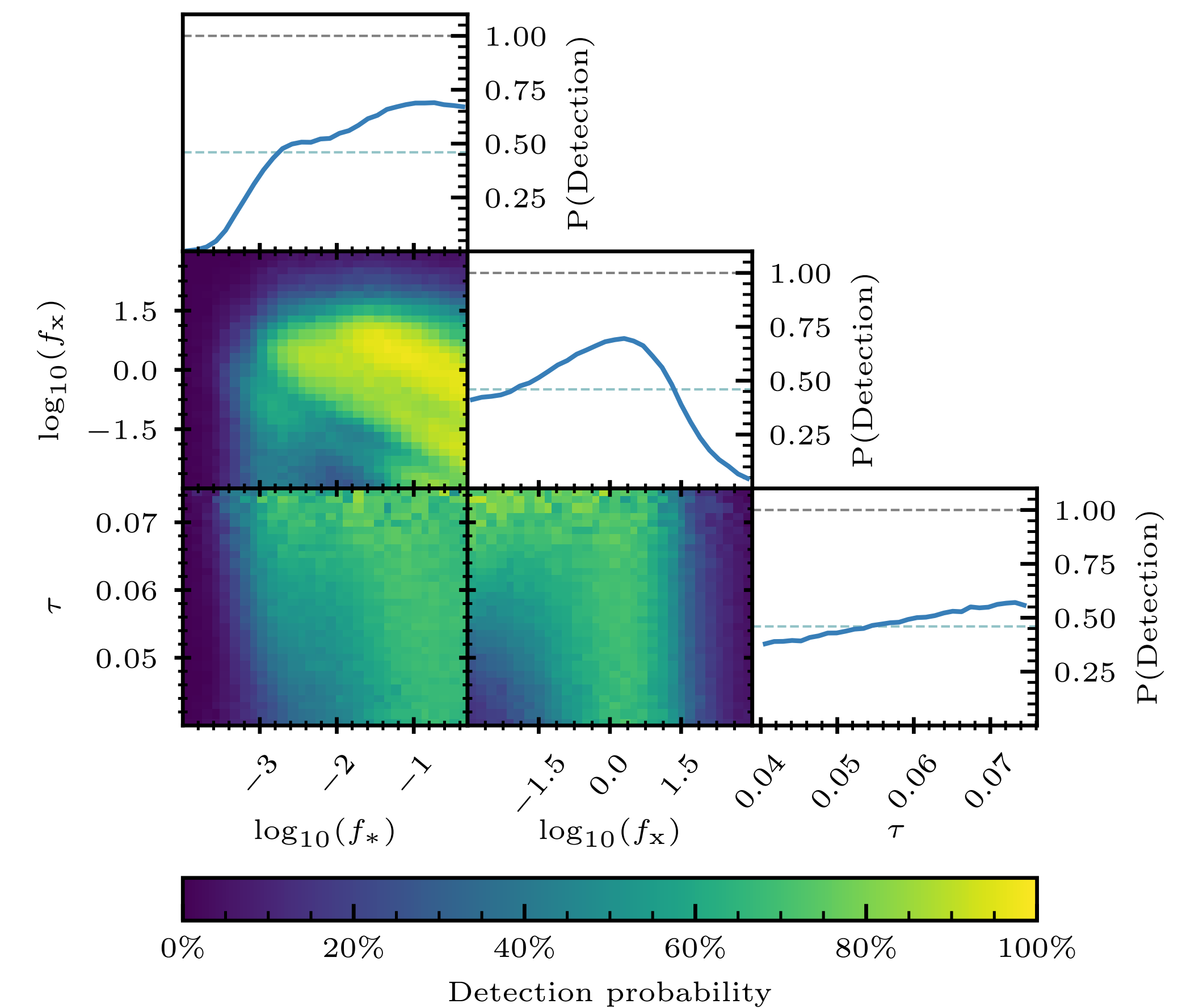
## Problem - Data Marginalized Forecasts for Bayesian Model Comparison are Typically Infeasible

**Nested Sampling:** Generally applicable but too slow to explore data space  
**Savage–Dickey forecasts:** Restrictive assumptions and require nested models  
 Both require explicit likelihoods.

## Solution - Neural Bayes Ratio Estimation (Evidence Network)



## Example - Detecting the Global 21-cm Signal



Just 5.54 GPU hours to perform instead of the 45,000,000 CPU hours nested-sampling would have taken.