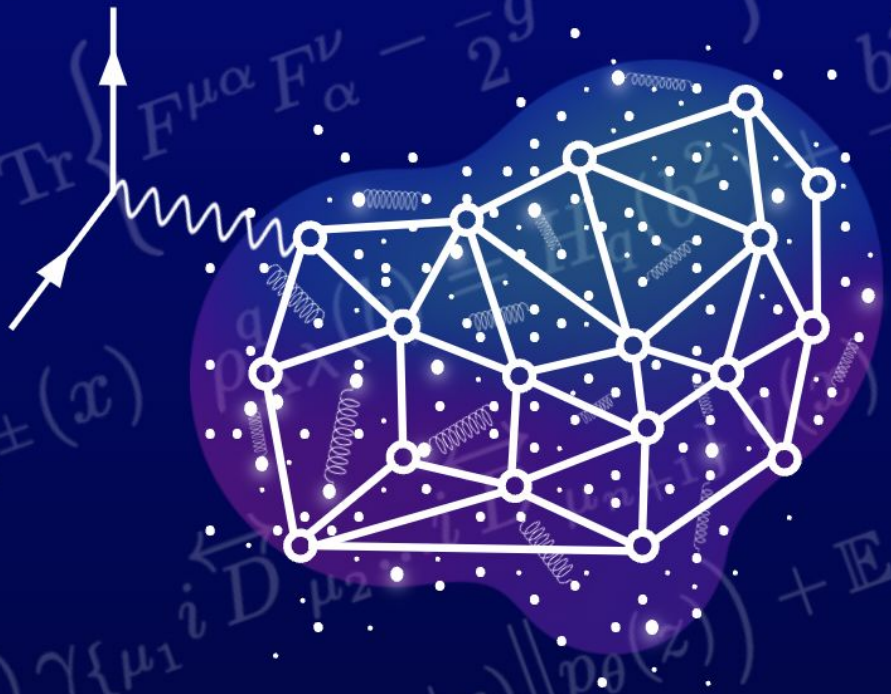


From Uncertainty to Discovery: Machine Learning at the Frontier of Phenomenology

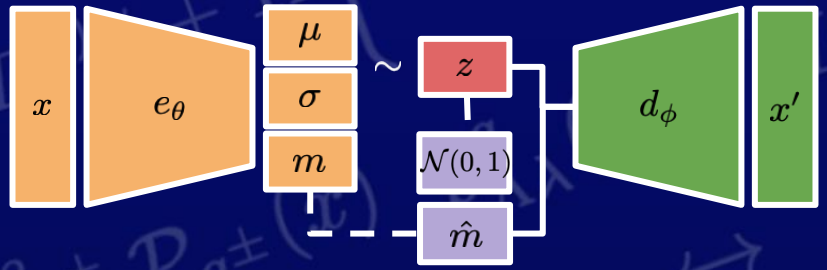
Brandon Kriesten • 19 November 2024 • PDFLattice

Argonne NATIONAL LABORATORY

Motivation / Outline

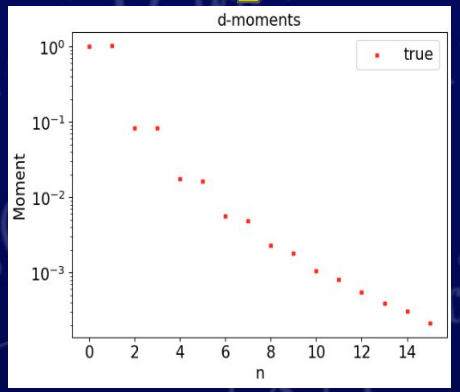
- Challenges in PDF determination and precision theory
  - reformatting a phenomenological PDF fit as an *inverse problem*
  - physics constraints (lattice QCD inputs, theory constraints)
  - Uncertainty quantification - major limitation in physics searches
- A jumble of questions with machine learning
  - How do we quantify uncertainties?
    - Aleatoric / epistemic (/ distributional OOD) separation?
  - Can we dissect and explain the 'black-box'?
  - Repurpose standard ML tools for physics discovery ...
- **Works**:
  - reconstruct PDFs from their Mellin moments BK, T.J. Hobbs **arXiv: 2312.02278**
  - explore explainability techniques BK, J. Gomprecht, T.J. Hobbs **arXiv: 2407.03411**
  - uncertainty quantification studies BK, T.J. Hobbs arXiv: 2412.XXXXX

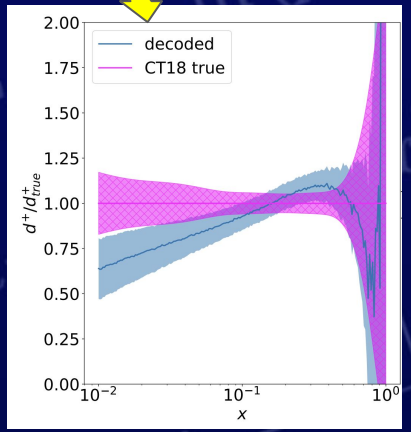
Generative AI for Inverse Problems



Using variational autoencoder as powerful generative model to generate solutions to inverse problems.

The latent variables are organized into interpretable physics constraints such as Mellin moments calculated on the lattice.



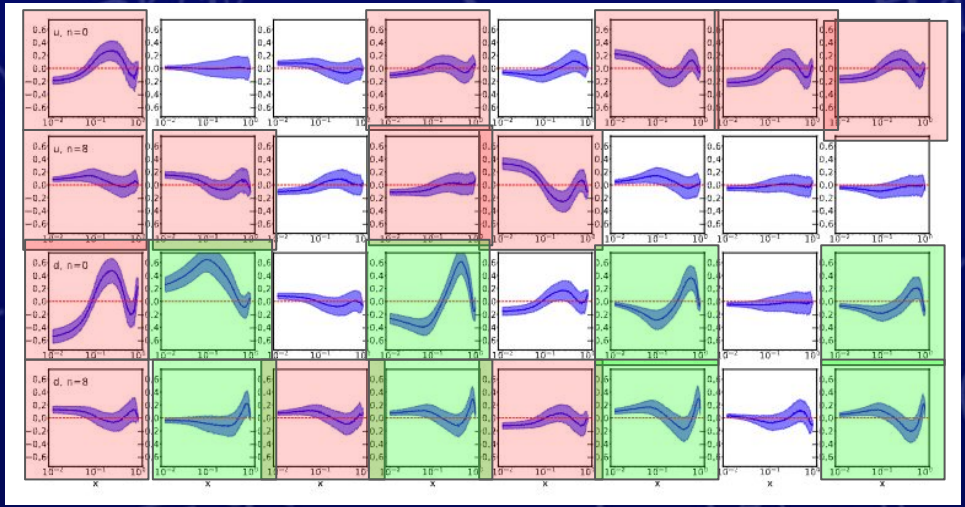


BK, T.J. Hobbs arXiv: 2312.02278 (accepted to PRD)

Argonne NATIONAL LABORATORY

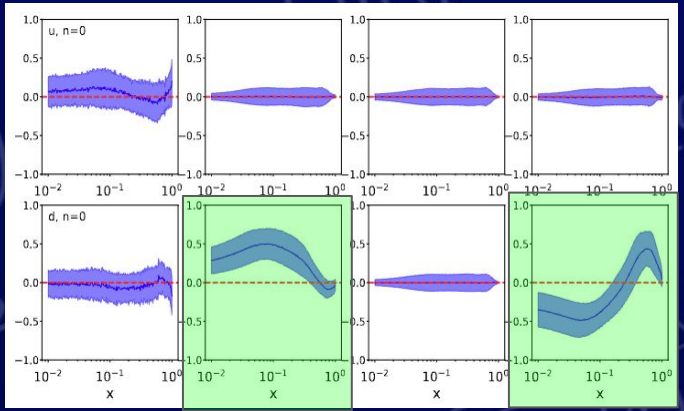
Generative AI for Inverse Problems

$$\mathrm{Corr}[d^+(x), \langle x^n \rangle_{u^\pm, d^\pm}]$$



With a large latent space, information is too free to create spurious correlations between moments and PDFs – not physics.

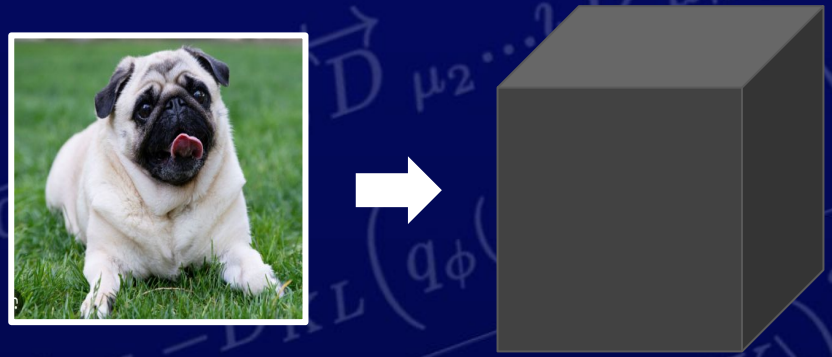
By constraining the latent dimensions, squeezing the bottleneck, one can force the AI to generate physics-like properties.



BK, T.J. Hobbs **arXiv: 2312.02278 (accepted to PRD)**

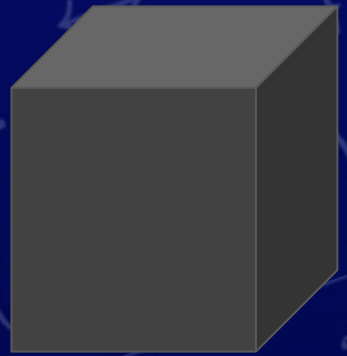
Probabilistic AI / ML

Probabilistic machine learning is an interpretation of AI / ML in which one thinks of the outputs of a specific machine model as learning the parameters of some probability distribution which describes your data. Ex. **classification models** learn the parameters of a **categorical distribution**.



ML can be confidently wrong

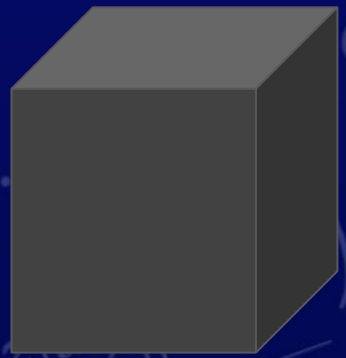
The problem with AI / ML is it can often be really **confidently wrong!**



Dog: 23 %
Cat: 65 %
Bird: 12 %

Uncertainty Quantification

What we want is more like this.



Much better!

Dog: 0.3 %
Cat: 0.2 %
Bird: 0.5 %
Idk?: 99%

This is why techniques like explainability (XAI) and uncertainty quantification (UQ) are important. Probabilistic AI / ML offers a mathematical language for these techniques. How to introduce this 4th category?

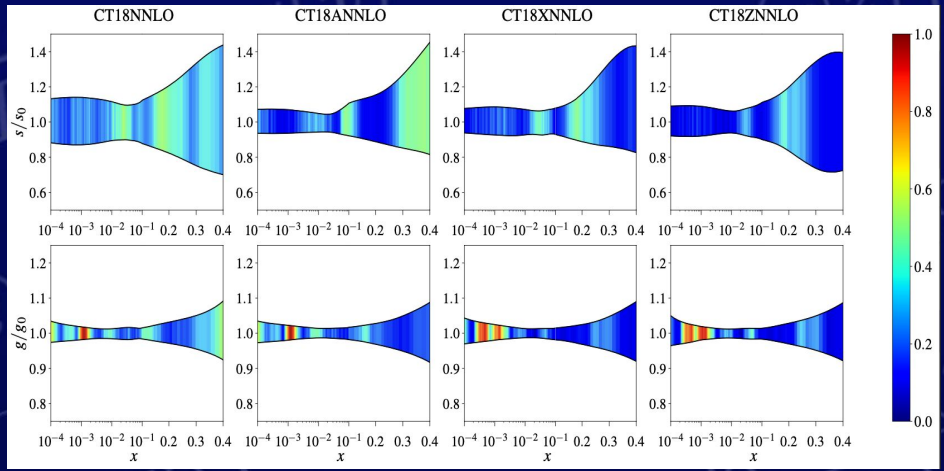
Argonne NATIONAL LABORATORY

Uncertainty Quantification

There are many open questions in phenomenological fitting of PDFs, many of which boil down to the open question of parameterization dependence: “How to effectively capture the associated effects of underlying theory assumptions on the fitted shape of the PDFs?”

Model discrimination among classes of parton densities - a classification problem. We can therefore trace-back the classification score to the x-dependence of the PDF (XAI).

How do we map parameterizations to some known space and quantify overlaps?



BK, J. Gomprecht, T.J. Hobbs **JHEP 11 (2024) 007**

Argonne NATIONAL LABORATORY

Uncertainty quantification for classification

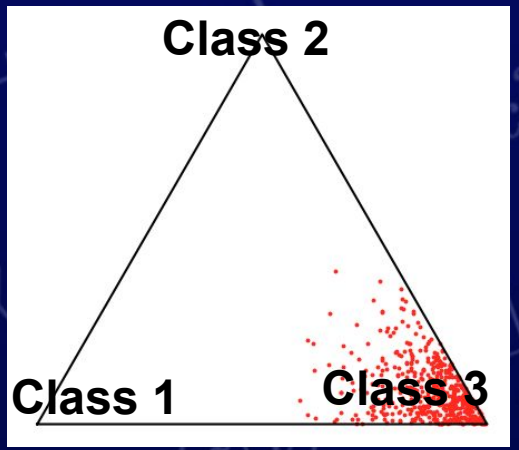
Bayesian Neural Networks

$$p(w_c|x^*, \mathcal{D}) = \int \underbrace{p(w_c|x^*, \theta)}_{\text{Aleatoric}} \underbrace{p(\theta|\mathcal{D})}_{\text{Epistemic}} d\theta$$

where through Monte Carlo sampling of the model parameters, we can create an ensemble of models which induces a distribution over the output.

$$\left\{ p(w_c|x^*, \theta^{(i)}) \right\}_{i=1}^{M}$$

There is an implicit prior distribution which is generating this ensemble of categorical distributions - an implicit Dirichlet.



Aleatoric / Epistemic Uncertainties

Through Maximum Likelihood Estimation (MLE) training we can approximately factorize the total uncertainty into pieces coming from the underlying data distribution (aleatoric) and the model's capabilities to recreate the data distribution (epistemic).

$$\mathbb{E}_{p_{true}(\mathbf{x},y)}\left[\mathcal{L}^{NLL}(y,\mathbf{x},\theta)\right] = \mathbb{E}_{p_{true}(\mathbf{x})}\left[D_{KL}\Big(p_{true}(y|\mathbf{x})\Big\|p(y|\mathbf{x},\theta)\Big) - \mathbb{H}\Big(p_{true}(y|\mathbf{x})\Big)\right]$$

Epistemic: reducible in theory, trickier to define in practice.

Aleatoric: irreducible, directly from underlying data distribution.

Argonne NATIONAL LABORATORY

Evidential Deep Learning

(Dirichlet) Prior Networks - making the implicit explicit

$$p(w_c|x^*, \mathcal{D}) = \int\int \underbrace{p(w_c|\mu)}_{\text{Aleatoric}} \underbrace{p(\mu|x^*, \theta)}_{\text{Distributional}} \underbrace{p(\theta|\mathcal{D})}_{\text{Epistemic}} d\theta d\mu$$

We can explicitly model the dependence on the prior of the ensemble. In a single forward pass, we can describe aleatoric, epistemic, and distributional uncertainties.

Malinin and Gales arXiv:1802.10501

Argonne NATIONAL LABORATORY

Separation of Uncertainties

Aleatoric

The uncertainty from the inherent noise or variability in the true underlying data distribution. **Cannot be reduced** by adding more training data nor by improving the training procedure.

Epistemic

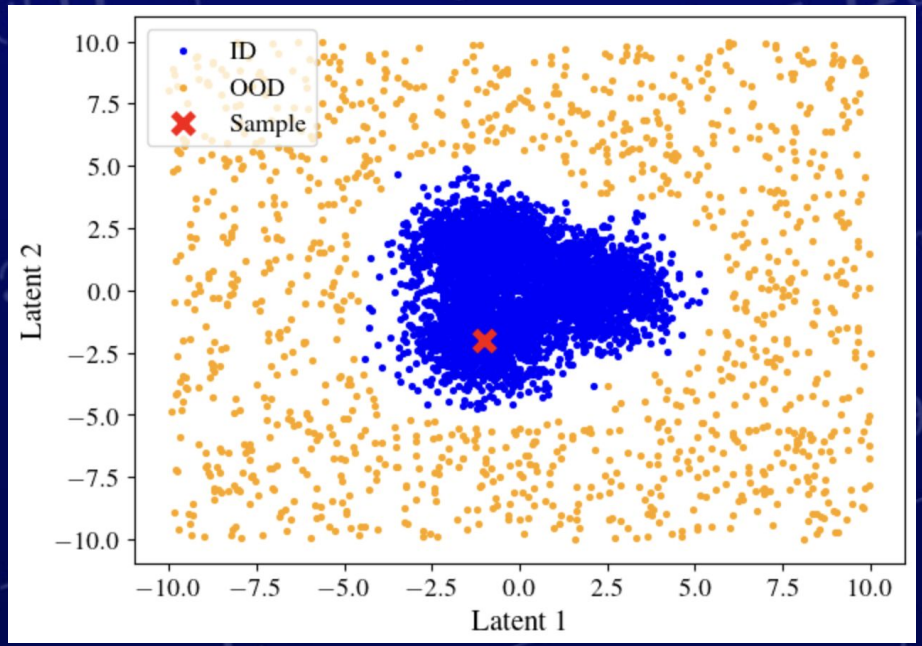
The uncertainty of how well the machine learning model has learned the underlying distribution of the data. **Can be reduced** by adding training data and improving training procedures.

Distributional

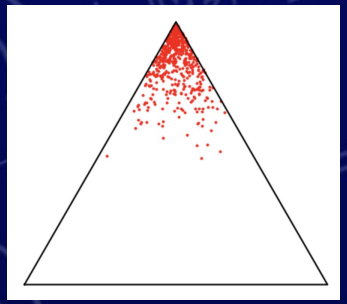
A form of epistemic uncertainty, represents the uncertainty of your choice of probability distribution to represent your data. Often reflected through uncertainty when encountering an example not represented in the training set.

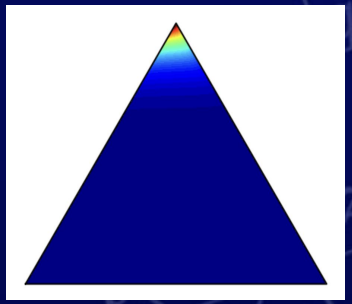
Malinin and Gales arXiv:1802.10501

Dirichlet Prior Networks - an example

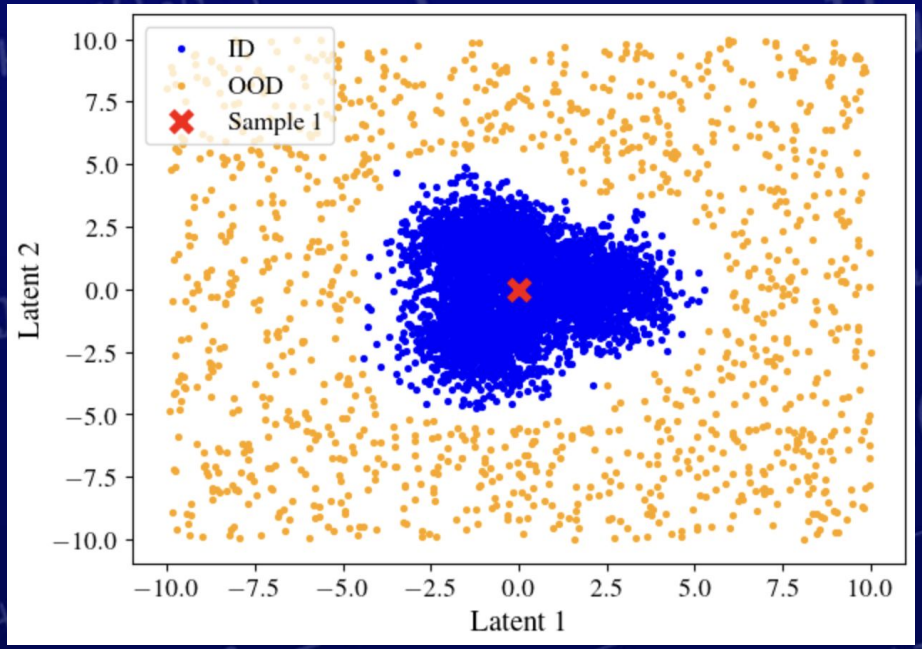


In-distribution sampling with low data uncertainty (samples are located on a corner of the simplex) and low knowledge uncertainty (high sample density).

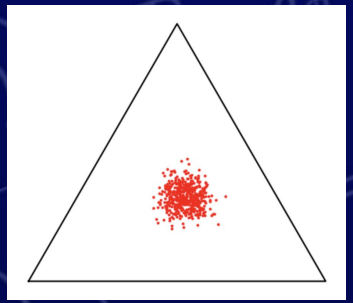


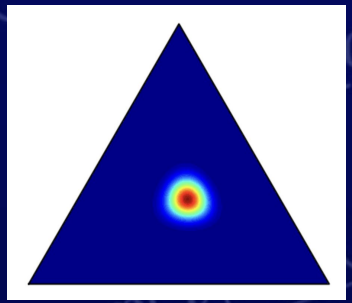


Dirichlet Prior Networks - an example

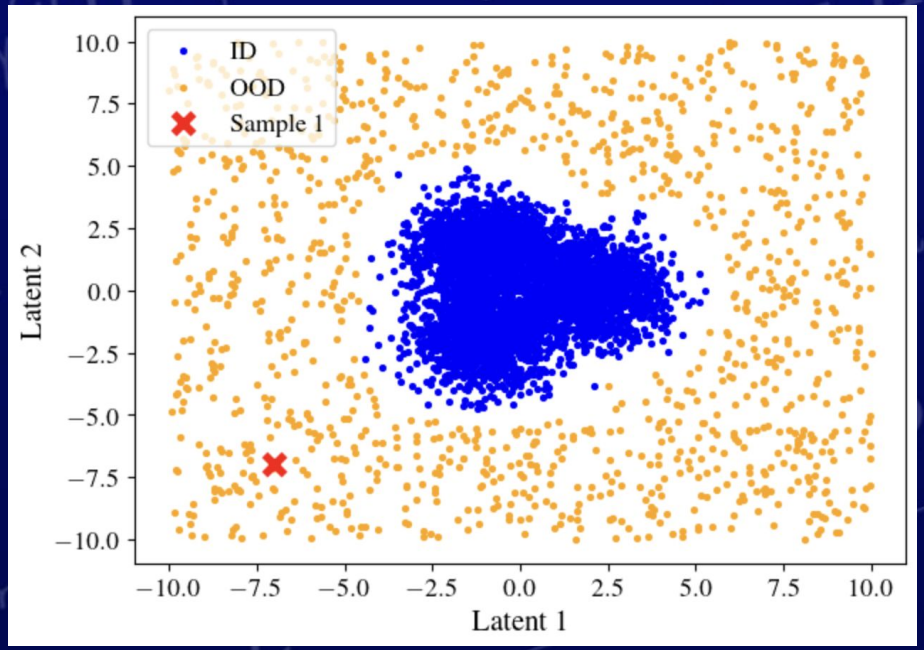


In-distribution sampling with high data uncertainty (samples are squeezed to the center of the simplex) with low knowledge uncertainty (high sample density).

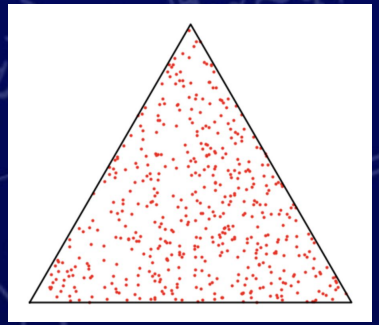


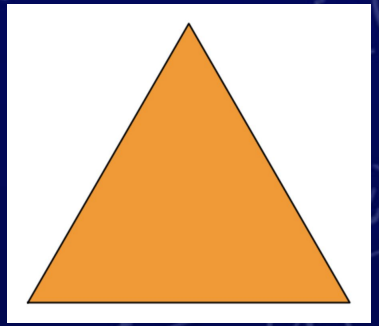


Dirichlet Prior Networks - an example



Out-of-distribution sampling with high data uncertainty (samples are diffuse) and high knowledge uncertainty (low sample density).

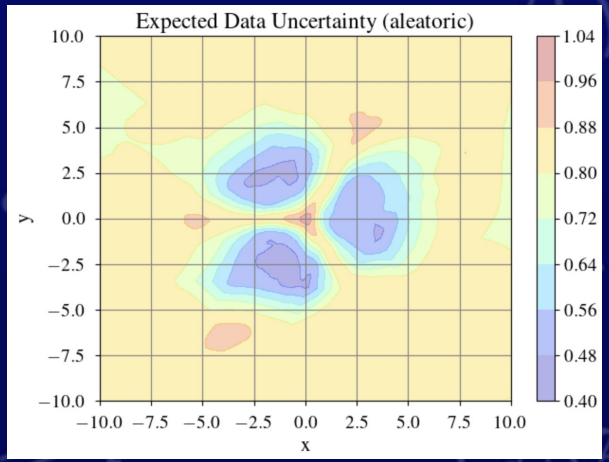


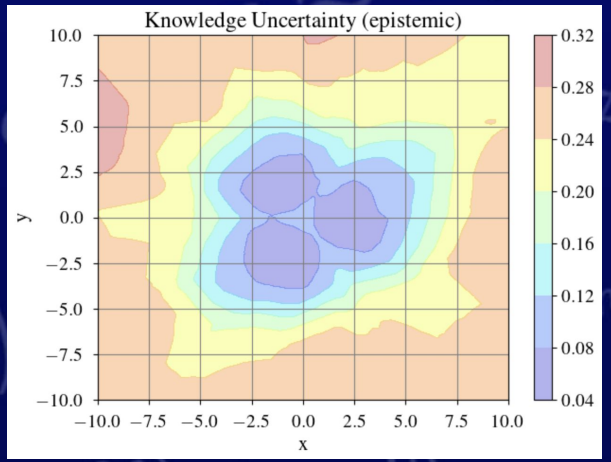


Dirichlet Prior Networks - an example

$$\mathbb{E}_{p(\mu|\mathbf{x}^*,\hat{\theta})}[\mathbb{H}[p(y|\mu)]] = \sum_{c=1}^{C} -\frac{\alpha_c}{\alpha_0}(\psi(\alpha_c+1) - \psi(\alpha_0+1))$$

$$\mathcal{I}[y,\mu|x^*,\mathcal{D}] = \mathbb{H}[\mathbb{E}_{p(\mu|x^*,\mathcal{D})}[p(y|\mu)]] - \mathbb{E}_{p(\mu|\mathbf{x}^*,\mathcal{D})}[\mathbb{H}[p(y|\mu)]]$$





We can separate the classification uncertainty into aleatoric (high in regions of high class overlap) and epistemic (high in regions where there is no data). Information theory description through the entropy and mutual information.

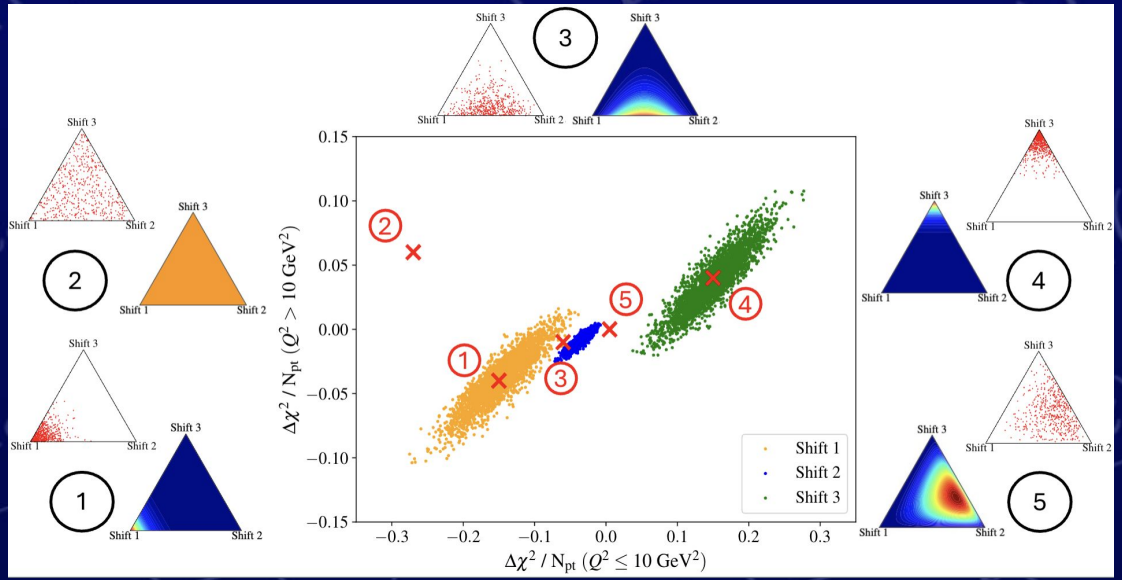
Uncertainty Quantification with prior networks

Mapping parametric theory space of BSM models

We construct mock BSM configurations EW standard model parameters in CC ν-DIS cross section.

Construct latent space and use UQ metrics to study how these models overlap.

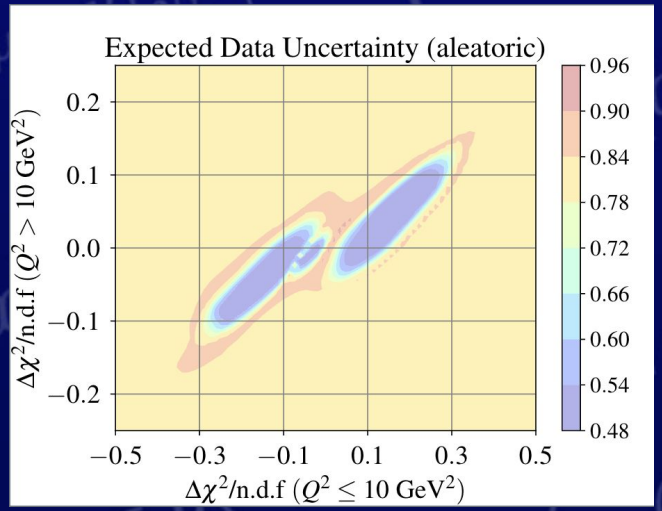
Dirichlet prior networks and classification for model discrimination.

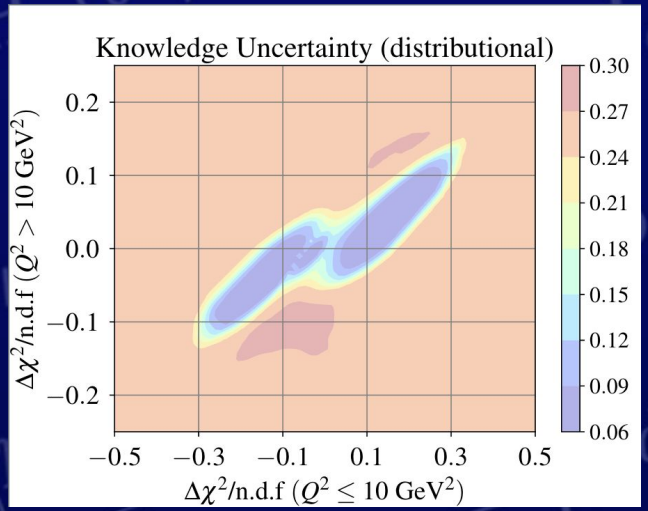


BK, T.J. Hobbs arXiv: 2412.XXXXX

Argonne NATIONAL LABORATORY

Uncertainty Measures for model overlap





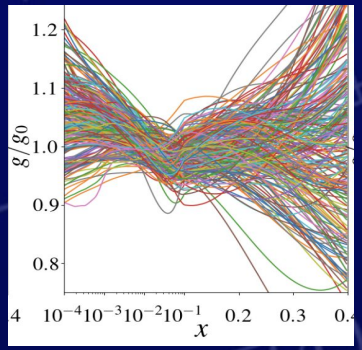
Quantitative information theory measures of where models overlap and are confused by the classification scheme …

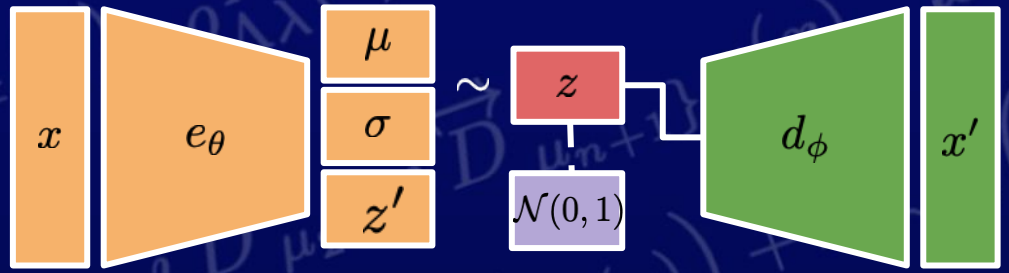
… and where there is no data. Seems trivial in 2D but in higher dimensions it becomes even more important.

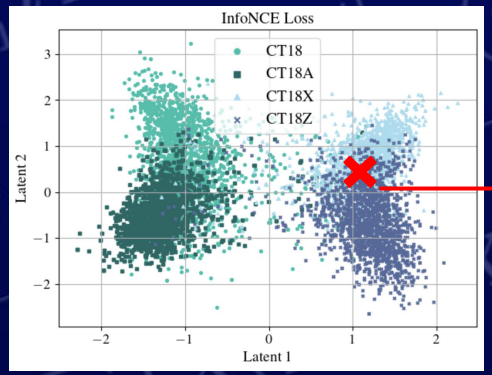
BK, T.J. Hobbs **arXiv: 2412.XXXXX**

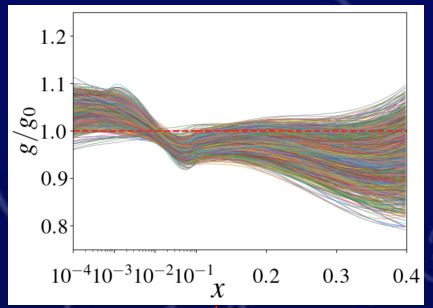
Future work

Mapping PDF parameterizations to constrained embedding space based on similarity metrics (contrastive learning). Model discrimination and generation in these spaces.









BK, T.J. Hobbs **(in progress)**

Argonne NATIONAL LABORATORY

UQ takeaways

Explainability and uncertainty quantification techniques are essential tools to understand learned behaviors from machine learning algorithms. Are ML algorithms learning physics and do we care?

Aleatoric uncertainty - *irreducible* - associated with inherent noise in training data. Tensions between data representation and class label.

Epistemic uncertainty - *reducible* - how well is your model learning the underlying training data distribution.

Distributional uncertainty - *reducible* - how well does your assumed prior match the true learned behavior of the data. Out of distribution sampling.

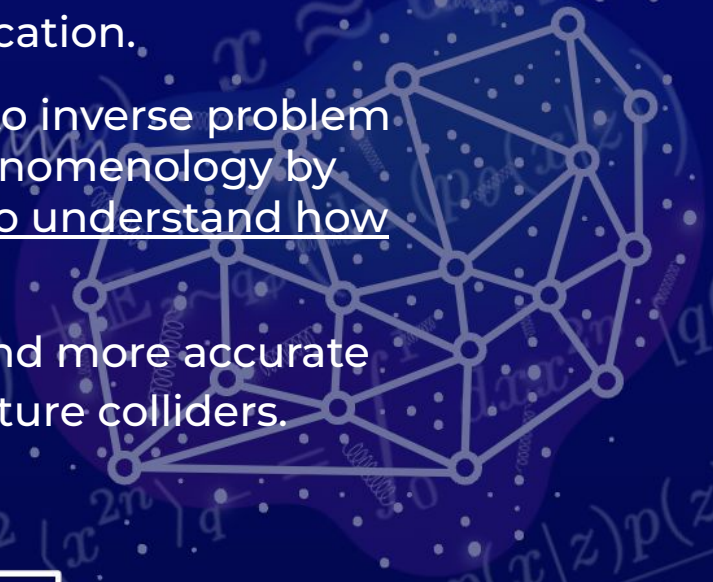
Extrapolation in ML models must be treated with extreme caution due to overconfidence in regions where there is no training data. (not negative!)

Conclusions

The research I have discussed here is the nucleus of a wide-reaching program culminating in comprehensive phenomenological fits with uncertainty quantification.

Generative AI offers a transformative approach to inverse problem solutions, driving the next wave of precision phenomenology by pushing beyond traditional methods. We need to understand how these models work with XAI and UQ methods.

Goal: Frontier discoveries for precision physics and more accurate predictions for high impact measurements at future colliders.



This work at Argonne National Laboratory was supported by the U.S. Department of Energy under contract DE-AC02-06CH11357.

Argonne NATIONAL LABORATORY

Thank you for your attention!

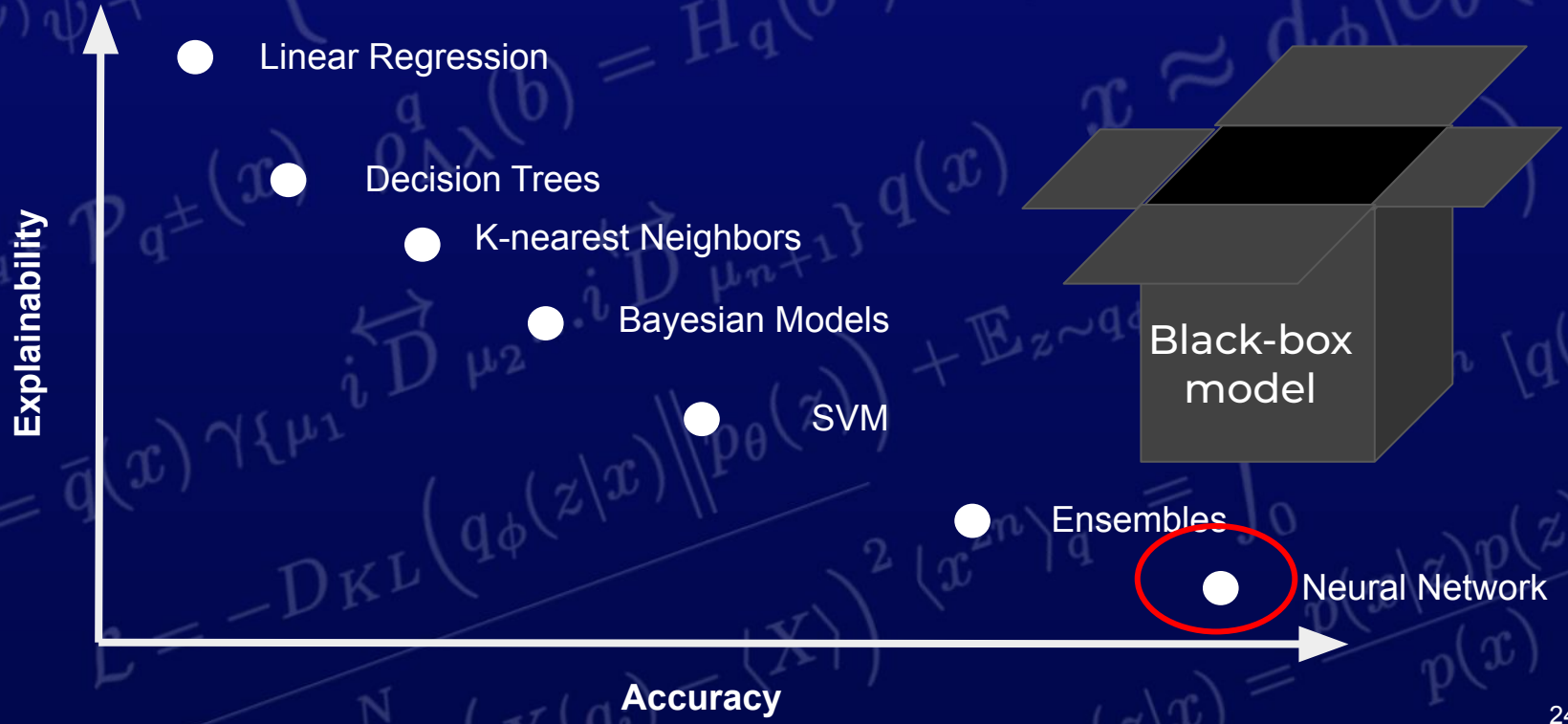
This work at Argonne National Laboratory was supported by the U.S. Department of Energy under contract DE-AC02-06CH11357.

Backup Slides

This work at Argonne National Laboratory was supported by the U.S. Department of Energy under contract DE-AC02-06CH11357.

Argonne NATIONAL LABORATORY

Explainability vs. accuracy



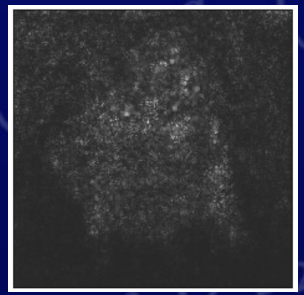
Argonne NATIONAL LABORATORY

Explainability ... a fun example!

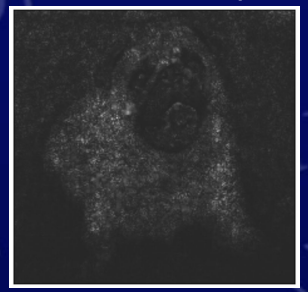
A survey of techniques

Input

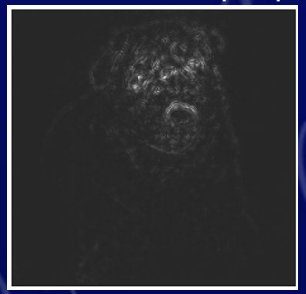
Gradients



Gradients ⊙ Input

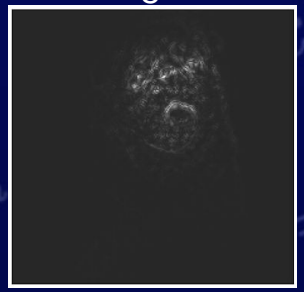


Guided Backprop



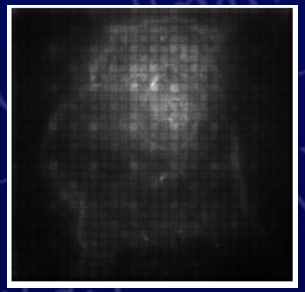
gradCAM

Guided gradCAM



Integrated Gradients

smoothGrad



Occlusion

Edge Detection

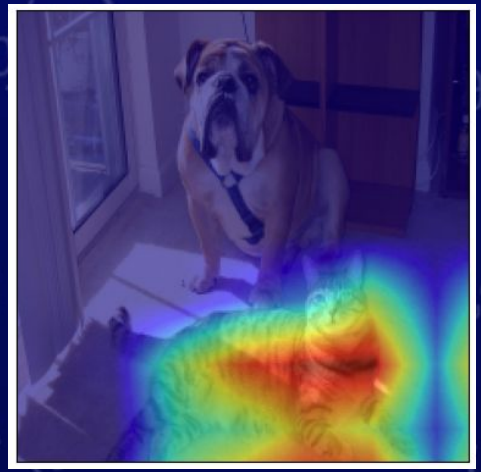
Argonne NATIONAL LABORATORY

Picking out features from image with multiple possible labels

Original

Dog

Cat



Argonne NATIONAL LABORATORY

Guided backpropagation

$$\frac{\partial f^{\text{out}}}{\partial f_i^{\ell}} = (f_i^{\ell} > 0) \cdot \left(\frac{\partial f^{\text{out}}}{\partial f_i^{\ell+1}} > 0 \right) \cdot \frac{\partial f^{\text{out}}}{\partial f_i^{\ell+1}}$$

Guided backprop is a “re-purposing” of the auto differentiation process in ML in which the gradients of a neural network layer are masked during a single backpropagation pass holding the weights fixed post-learning to determine which input features positively affect the classification outcome the most.

Simonyan et. al. arXiv: 1312.6034

Argonne NATIONAL LABORATORY

Guided backpropagation

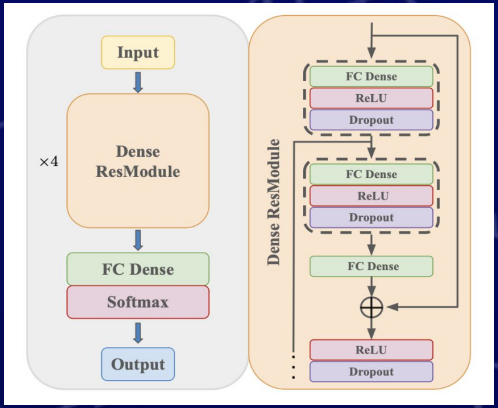
$$\frac{\partial f^{\text{out}}}{\partial f_i^{\ell}} = (f_i^{\ell} > 0) \cdot \left(\frac{\partial f^{\text{out}}}{\partial f_i^{\ell+1}} > 0 \right) \cdot \frac{\partial f^{\text{out}}}{\partial f_i^{\ell+1}}$$

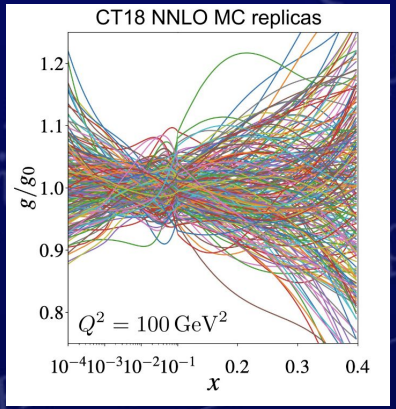
$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} \otimes \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} \otimes \begin{bmatrix} -1 & 2 & -3 \\ 6 & -1 & 4 \\ 1 & -7 & 5 \end{bmatrix} = \begin{bmatrix} 0 & 2 & 0 \\ 0 & 0 & 4 \\ 1 & 0 & 5 \end{bmatrix}$$

The double-masking procedure during backpropagation generates highly detailed saliency maps, effectively highlighting fine-grained input features that most influence the network's output.

Classifying PDFs from salient features

Train a ResNet-like model on PDF MC replicas to identify salient features in x-dependence for classification tasks.





Fitting methodology: can we trace effects from the underlying theory back to the x-dependence of the PDF?

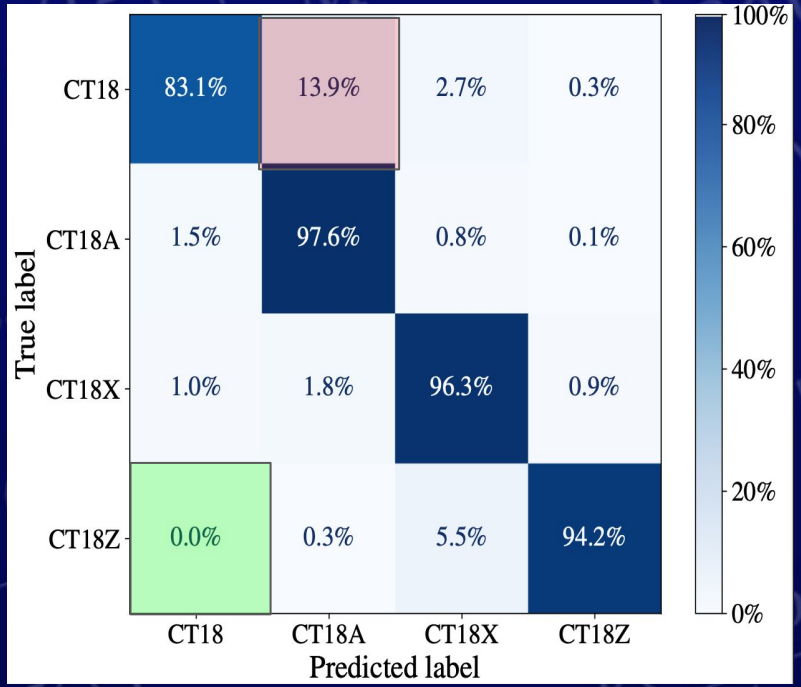
BK, J. Gomprecht, T.J. Hobbs **JHEP 11 (2024) 007**

Argonne NATIONAL LABORATORY

XAI4PDF: Explainability for fitted PDFs

PDF fits	Factorization scale in DIS	ATLAS 7 TeV W/Z data included?	CDHSW $F_2^{p,d}$ data included?	Pole charm mass, GeV
CT18	$\mu^2_{F,DIS} = Q^2$	No	Yes	1.3
CT18A	$\mu^2_{F,DIS} = Q^2$	Yes	Yes	1.3
CT18X	$\mu^2_{F,DIS} = 0.8^2\left(Q^2 + \frac{0.3\ \mathrm{GeV}^2}{x_B^{0.3}}\right)$	No	Yes	1.3
CT18Z	$\mu^2_{F,DIS} = 0.8^2\left(Q^2 + \frac{0.3\ \mathrm{GeV}^2}{x_B^{0.3}}\right)$	Yes	No	1.4

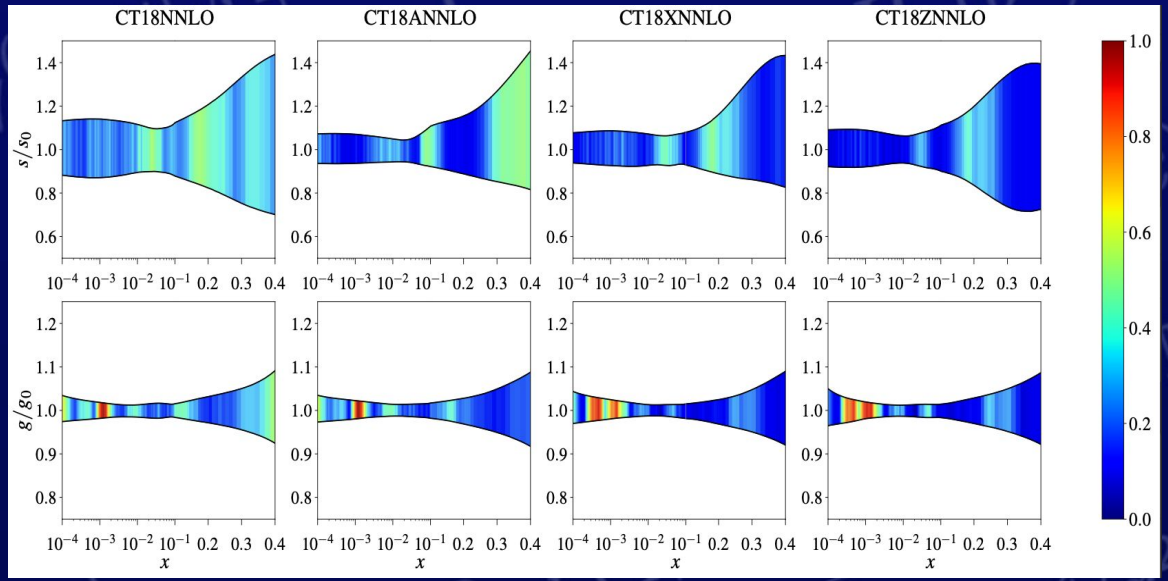
The two analyses which are “furthest” from each other (CT18 and CT18Z) are also the least confused, confirming that the shift in theory assumptions drives the statistical distinguishability as inferred by the XAI calculation.



BK, J. Gomprecht, T.J. Hobbs **JHEP 11 (2024) 007**

Argonne NATIONAL LABORATORY

XAI4PDF: Explainability for fitted PDFs



The strange and gluon PDFs stand out while discerning between different theory fits!

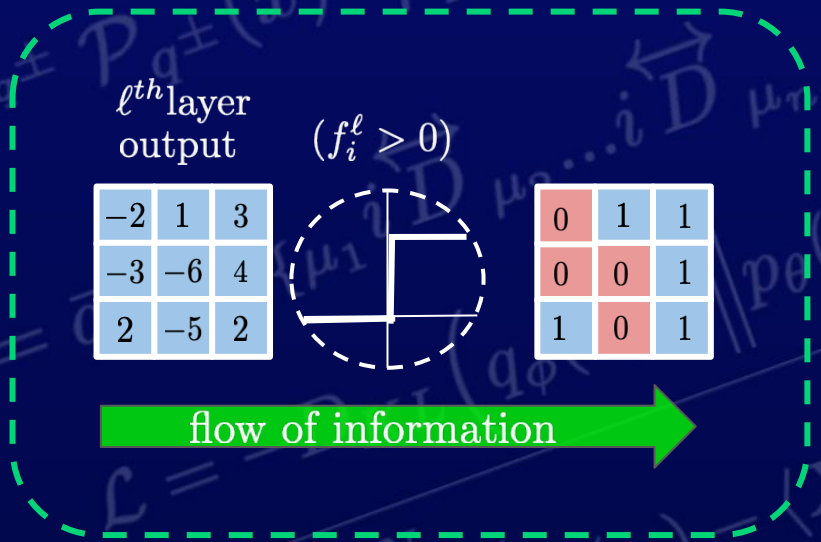
The gluon replicas have a dominant role in the classification among the CT18 series with highly localized gradients.

The strange replicas have smoother gradients indicating a weaker role.

BK, J. Gomprecht, T.J. Hobbs **JHEP 11 (2024) 007**

Guided backpropagation

$$\frac{\partial f^{\text{out}}}{\partial f_i^{\ell}} = (f_i^{\ell} > 0)\left(\frac{\partial f^{\text{out}}}{\partial f_i^{\ell+1}} > 0\right) \cdot \frac{\partial f^{\text{out}}}{\partial f_i^{\ell+1}}$$

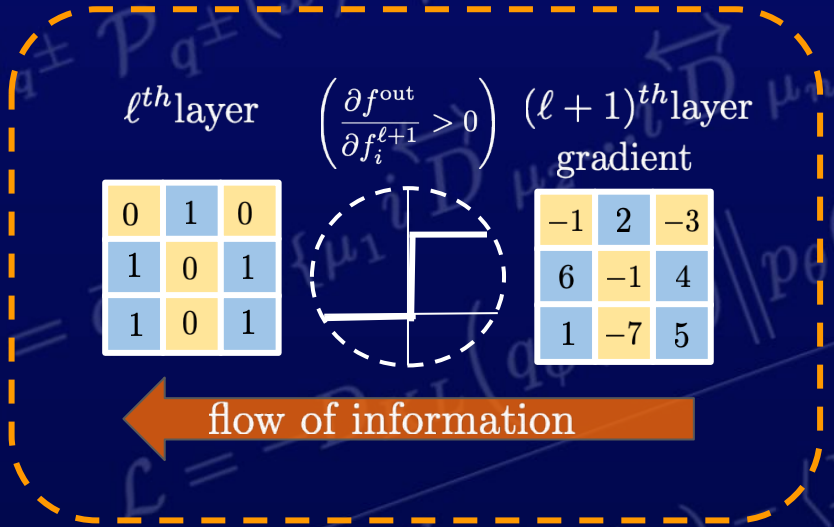


Forward Pass Logic Gate

Ensures only positive activations in the ℓ^{th} layer are considered when backpropagating the gradient. Prevents input from negative activations in the forward flow of information.

Guided backpropagation

$$\frac{\partial f^{\text{out}}}{\partial f_i^{\ell}} = (f_i^{\ell} > 0) \cdot \left(\frac{\partial f^{\text{out}}}{\partial f_i^{\ell+1}} > 0 \right) \cdot \frac{\partial f^{\text{out}}}{\partial f_i^{\ell+1}}$$

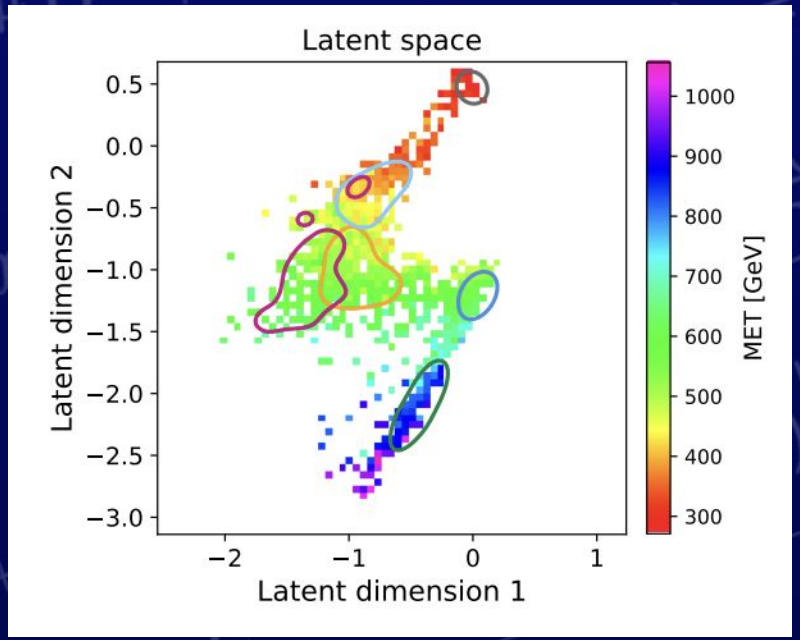


Backward Gradient Logic Gate

Ensures only positive gradients from the $(\ell+1)^{\text{th}}$ layer are considered when backpropagating the gradient. Prevents negative gradients in the backward flow of information.

Argonne NATIONAL LABORATORY

Mapping Beyond Standard Model theory landscape



There is significant effort to map the theoretical landscape of BSM configurations.

This will help us understand where theoretical models are lacking.

But how do we define where models overlap or are just simply missing? A model discrimination task … classification with uncertainty quantification.

Image: Hallin et. al. 2407.20315