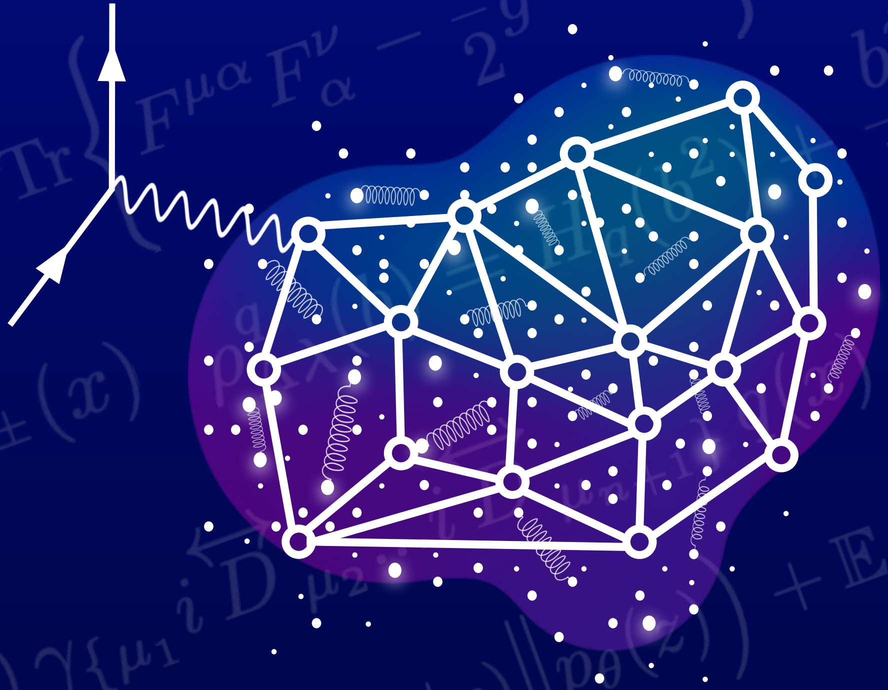


Learning PDFs through Interpretable Representations in Mellin Space

Brandon Kriesten • 4 June 2024 • CTEQ Spring Meeting

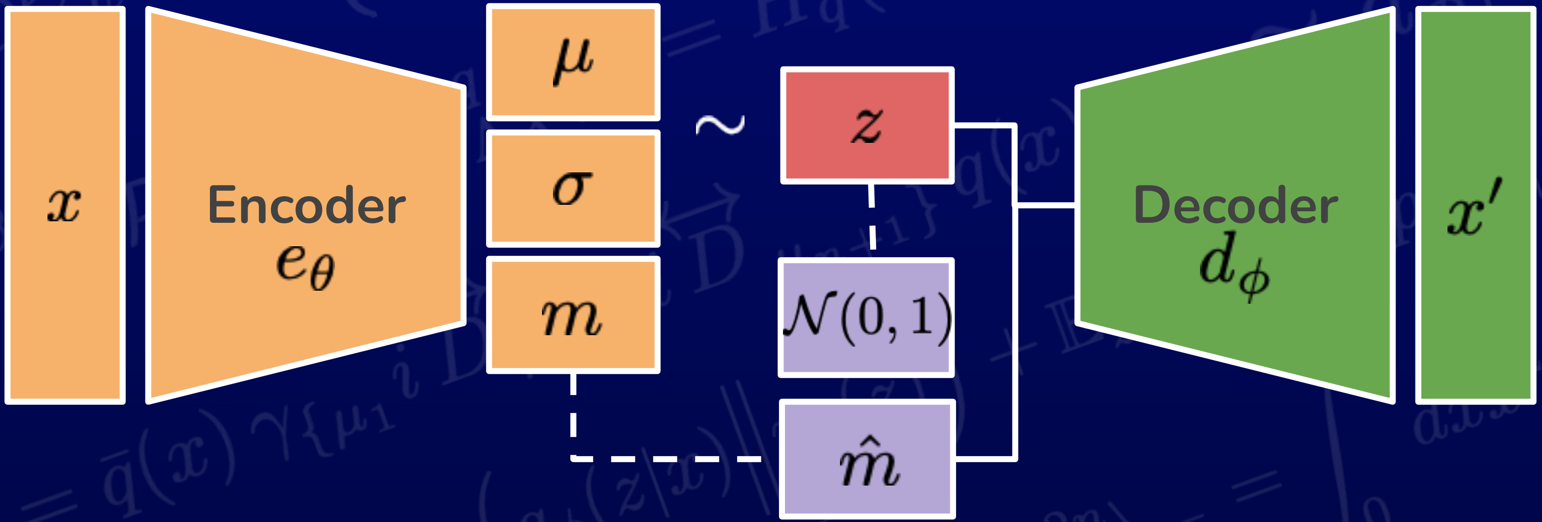
Motivation

- Neural network representations of PDFs?
  - reformatting a phenomenological PDF fit as an *inverse problem*
  - physics constraints (lattice QCD inputs)
  - possible many solutions
- A jumble of questions with neural networks
  - How do we quantify uncertainties? Aleatoric / epistemic (/distributional OOD) separation?
  - Can we interpret the 'black-box'? etc …
- **Outline**:

  BK, T.J. Hobbs **arXiv: 2312.02278 (submitted to PRD)**
  - reconstruct PDFs from their Mellin moments
  - explore interpretability techniques
  - uncertainty quantification studies

  } in progress works

Variational Autoencoder
Inverse Mapper

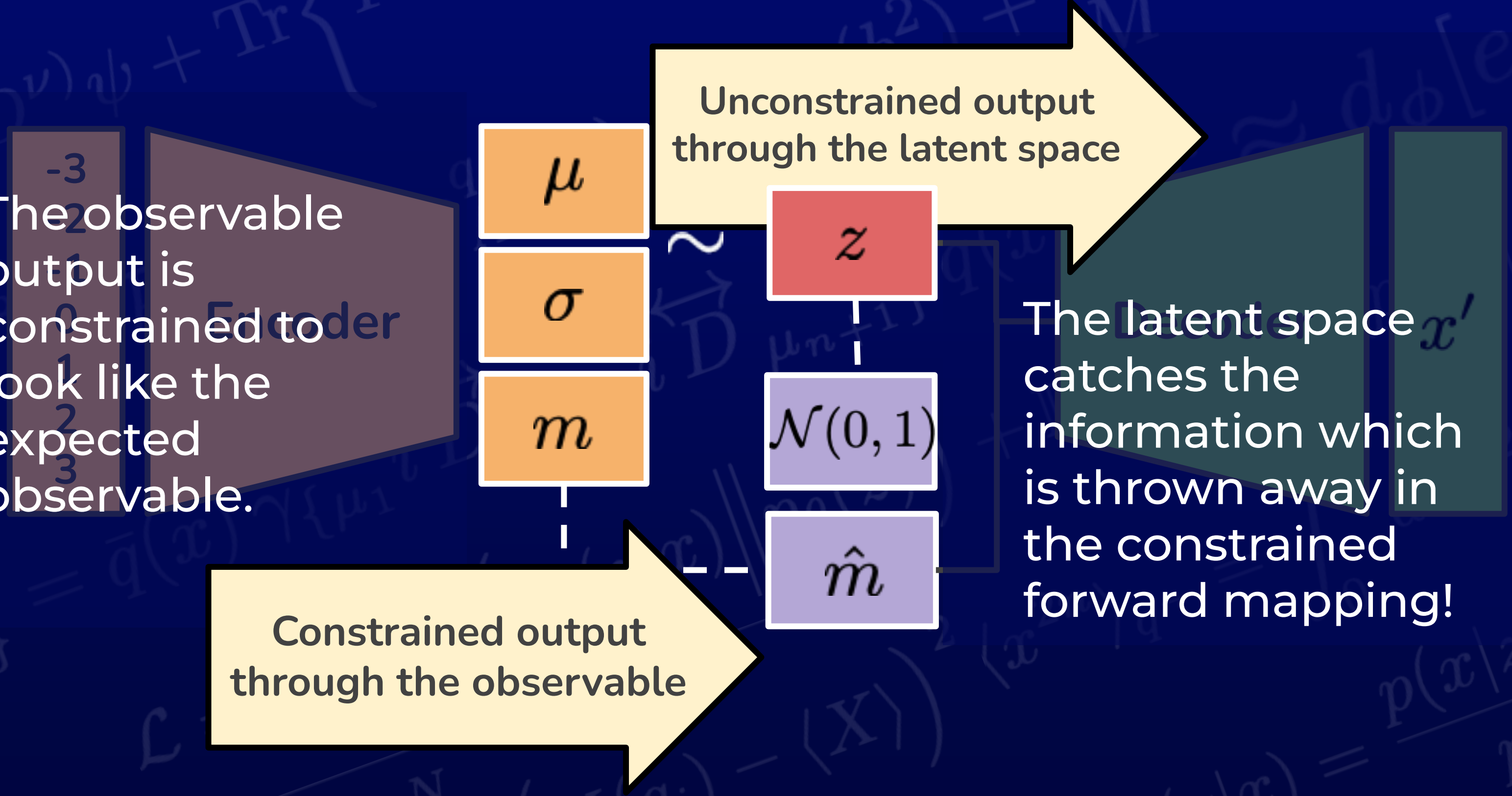


We utilize variational autoencoders as a powerful tool to dissect inverse problems!

M. Almaeen, Y. Alanazi, N. Sato, W. Melnitchouk, M.P. Kuchera, Y. Li **IJCNN (2021)**

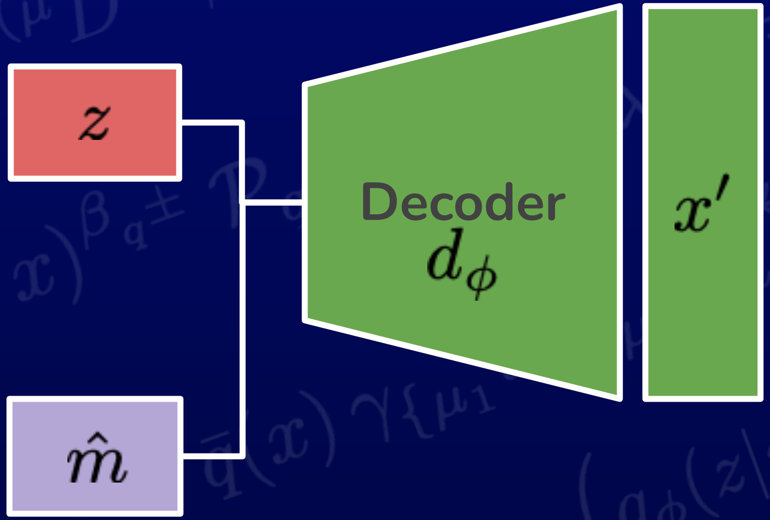
Information Flow and Reconstruction

The observable output is constrained to look like the expected observable.



The latent space catches the information which is thrown away in the constrained forward mapping!

Variational Autoencoder
Inverse Mapper



The goal is to train a decoder model to accept latent information and an observable to generate a never before seen input!

Generative algorithms provide access to new technologies beyond neural network interpolation.

M. Almaeen, Y. Alanazi, N. Sato, W. Melnitchouk, M.P. Kuchera, Y. Li **IJCNN (2021)**

Parton Distribution Functions
from latent space Mellin moments

If we have an infinite amount of information, we can construct the PDF exactly from classical methods

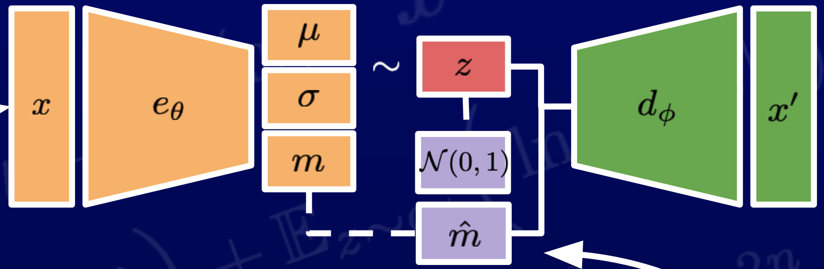
$$q(x) + (-1)^{n+1}\bar{q}(x) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} dn x^{-n-1}\langle x^n\rangle_q$$

We don't have an infinite amount of information, typically we have just a few of these moments from the lattice.

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

Parton Distribution Functions from latent space Mellin moments

Question: Can we determine the full x-dependence of PDFs from a set number of Mellin moments?



Model Inputs

Randomly generated PDFs with 5 parameters.

$$q(x) \pm \bar{q}(x) = \mathcal{N}_{q^\pm} x^{\alpha_{q^\pm}} (1-x)^{\beta_{q^\pm}} \mathcal{P}_{q^\pm}(x)$$
$$\mathcal{P}_{q^\pm}(x) = 1 + \gamma_{q^\pm}\sqrt{x} + \delta_{q^\pm} x$$

Latent Observable

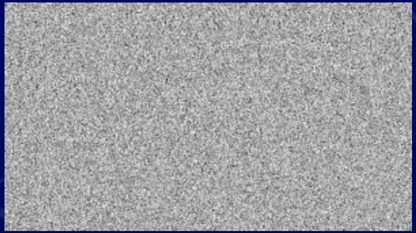
Organized and interpretable as a series of moments

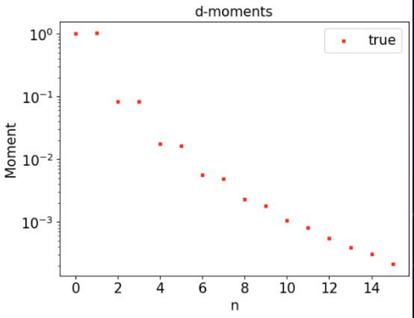
$$\langle 1 \rangle_{q^-}, \langle x \rangle_{q^+}, \langle x^2 \rangle_{q^-}, \langle x^3 \rangle_{q^+}, \ldots$$

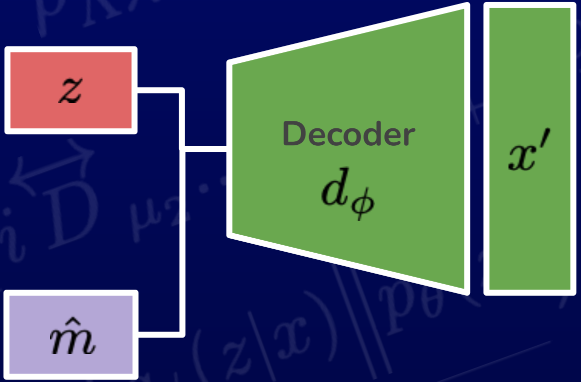
BK, T.J. Hobbs **arXiv: 2312.02278 (submitted to PRD)**

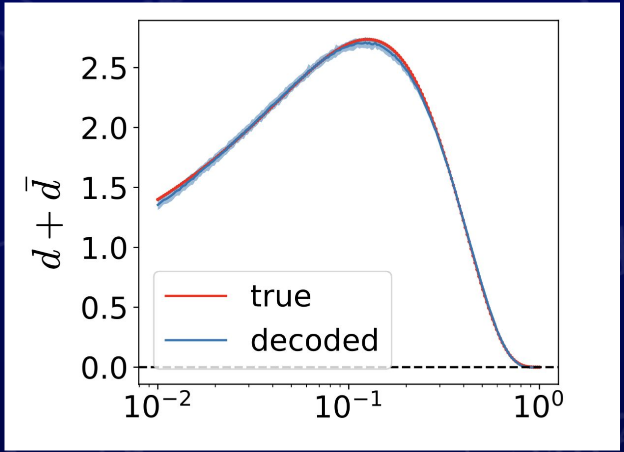
Parton Distribution Functions
from latent space Mellin moments

Making predictions from a trained decoder model.



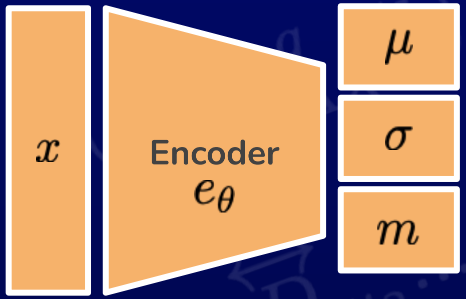


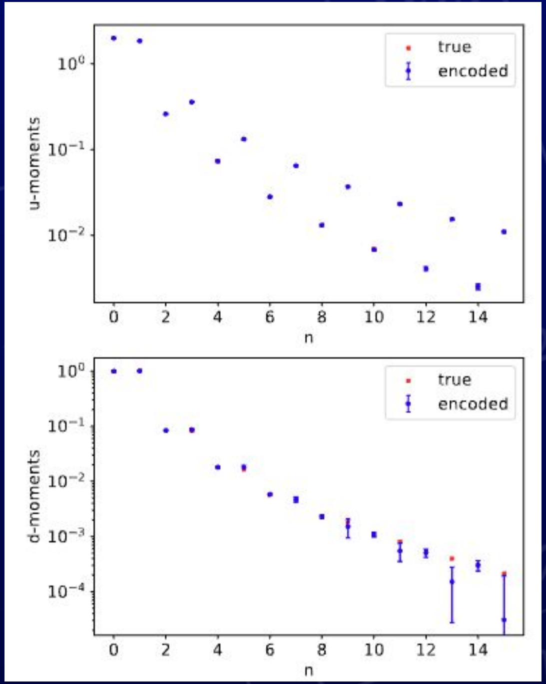




BK, T.J. Hobbs **arXiv: 2312.02278 (submitted to PRD)**

Parton Distribution Functions
from latent space Mellin moments





We can also look at the trained encoder model to see how well we reconstruct the observable.

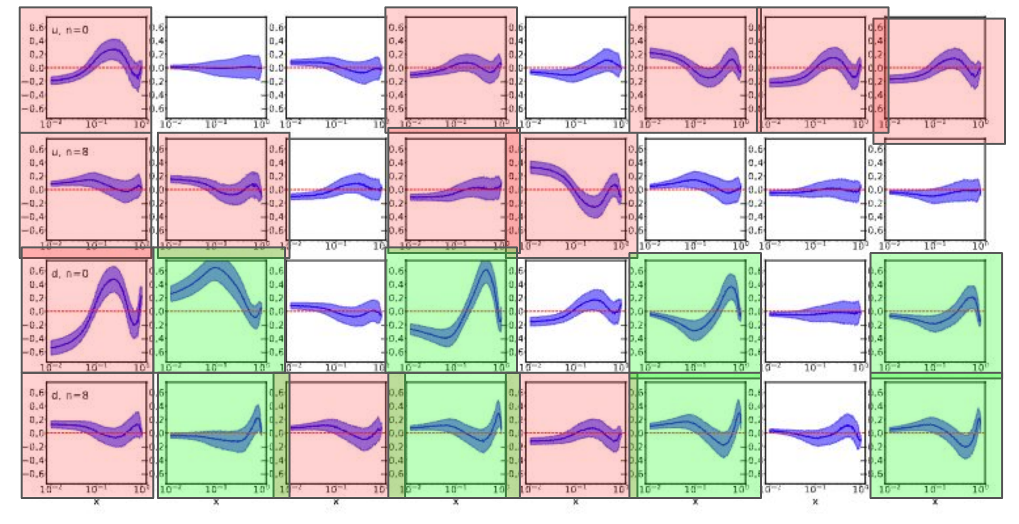
BK, T.J. Hobbs **arXiv: 2312.02278 (submitted to PRD)**

Parton Distribution Functions from latent space Mellin moments

We utilize the Pearson correlation between the learned moments from the encoder and the decoded PDF (d^+) as an explainability technique. One can see spurious correlated effects as well as consistent correlations.

$$\mathrm{Corr}(X,Y) = \frac{\langle XY\rangle - \langle X\rangle\langle Y\rangle}{\Delta X \Delta Y}$$

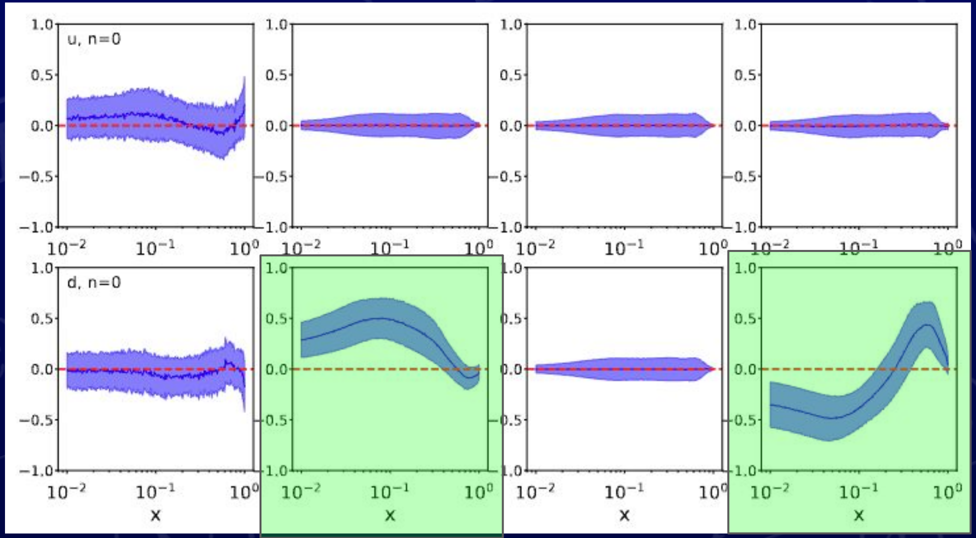
$$\mathrm{Corr}[d^+(x), \langle x^n\rangle_{u^\pm, d^\pm}]$$



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

Parton Distribution Functions from latent space Mellin moments

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

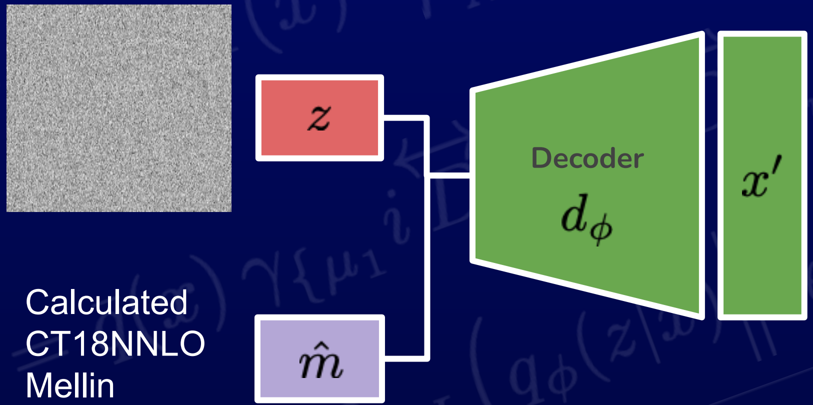


With a more dramatically undercomplete autoencoder architecture, the correlations are statistically consistent with 0 everywhere except for the d^+ moments. Obvious spurious correlations seem to disappear.

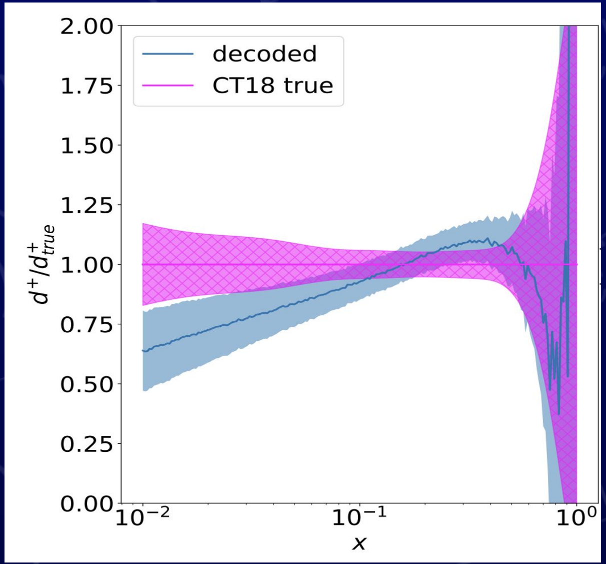
BK, T.J. Hobbs **arXiv: 2312.02278 (submitted to PRD)**

Parton Distribution Functions from latent space Mellin moments

Question: With a fully trained decoder network from the toy problem, can we construct completely unknown PDF from a full phenomenological fit from its moments?

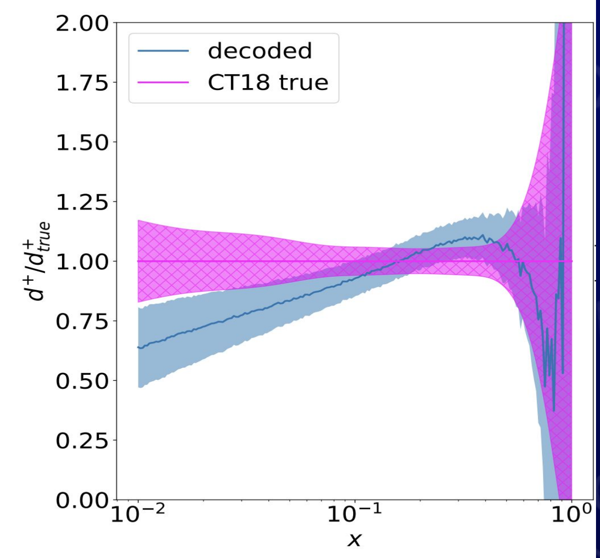


Calculated CT18NNLO Mellin moments.



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

Parton Distribution Functions
from latent space Mellin moments



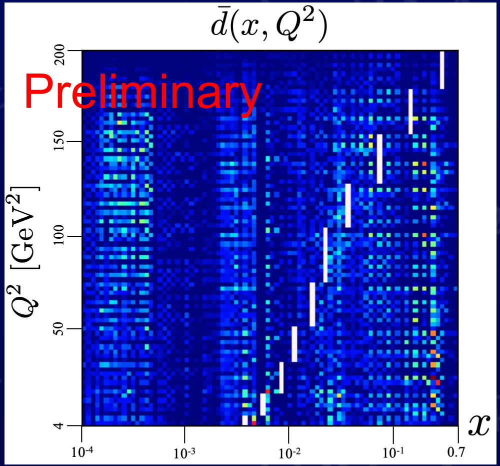
Is CT18NNLO within the parameter space of this toy model somewhere? Or is this just coincidence? Opens a lot of questions regarding parameterization dependence.

Highlights the need for UQ and explainability techniques to fully understand.

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

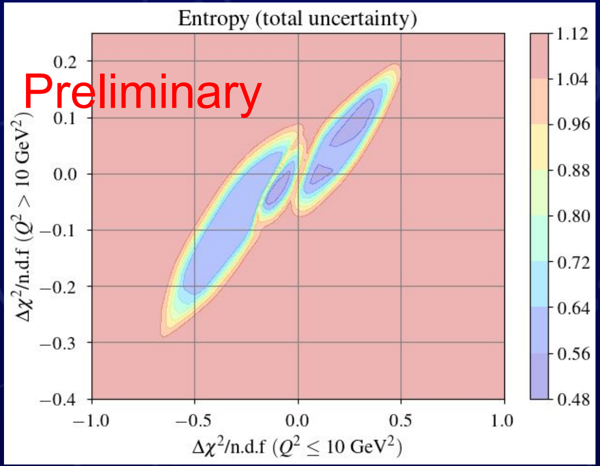
Exploring Next Gen ML techniques

XAI4PDF



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

Evidential Deep Learning for UQ



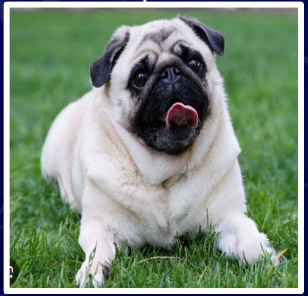
BK, T.J. Hobbs **(in progress)**

XAI: Guided Backpropagation

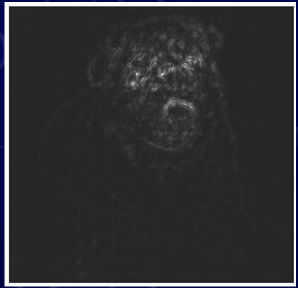
$$\frac{\partial f^{\text{out}}}{\partial f_i^{\ell}} = (f_i^{\ell} > 0) \cdot \left(\frac{\partial f^{\text{out}}}{\partial f_i^{\ell+1}} > 0 \right) \cdot \frac{\partial f^{\text{out}}}{\partial f_i^{\ell+1}}$$

Guided backprop is a technique in which the gradients of a neural network layer are masked during backpropagation holding the weights fixed to determine which input features positively affect the classification outcome the most.

Input

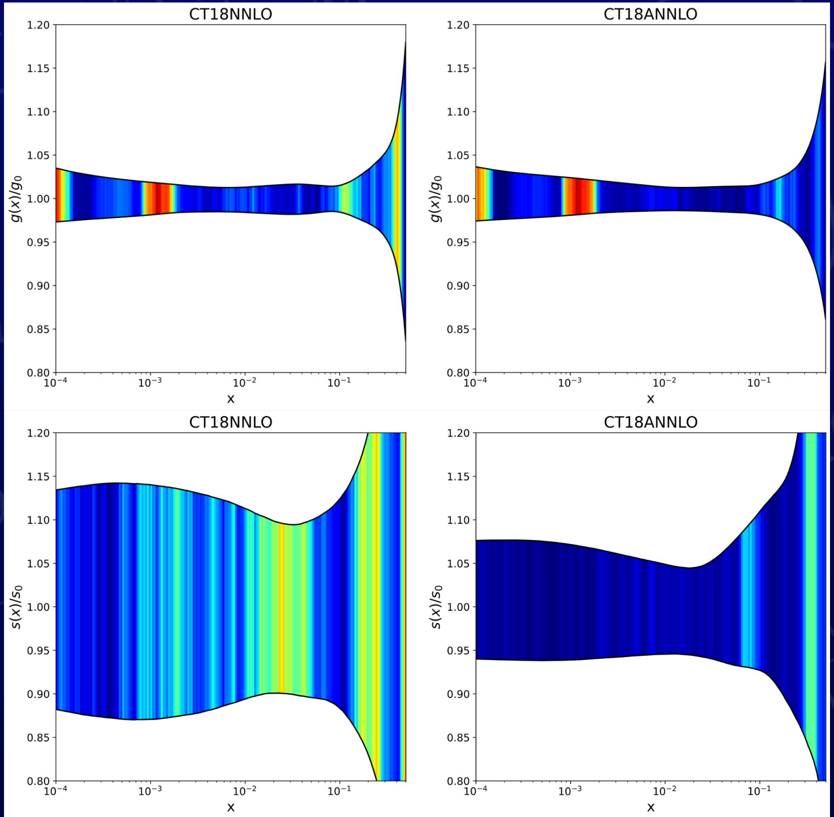


Guided Backprop



XAI4PDF: Explainability across 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	$\mu^2_{F,DIS} = 0.8^2\left(Q^2 + \frac{0.3\ \text{GeV}^2}{x_B^{0.3}}\right)$	No	Yes	1.3
CT18Z	$\mu^2_{F,DIS} = 0.8^2\left(Q^2 + \frac{0.3\ \text{GeV}^2}{x_B^{0.3}}\right)$	Yes	No	1.4

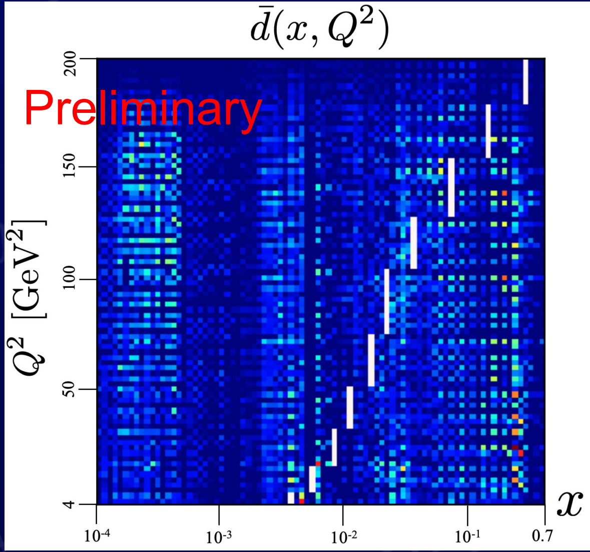


We classify different PDF fits with various theory assumptions.

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

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

XAI4PDF: Explainable AI for PDFs



Explainability within a fitted phenomenological framework to investigate how the χ^2 on the CDHSW data (neutrino-lead scattering) traces back to regions in the phase space of the fitted PDF. The aim is to identify regions in x-Q^2 space that most impact the χ^2.

This is a novel approach to dissect ML-based PDF analyses and understand their internal behavior.

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

Uncertainty in High Energy Theory

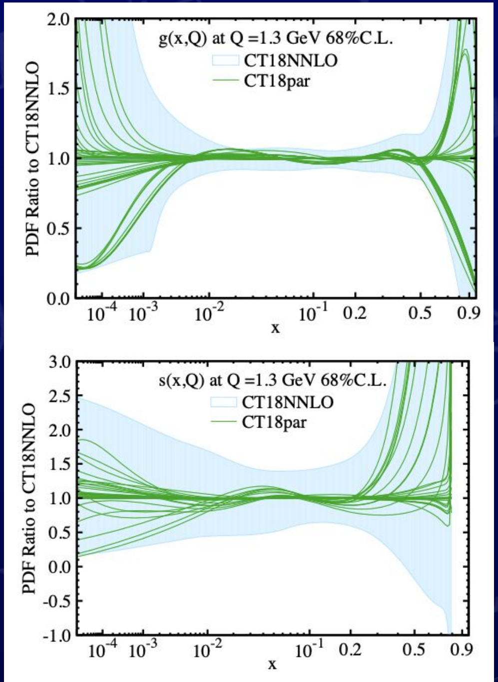


Image credit: PhysRevD.103.014013

Obtaining stable PDF uncertainties is a significant area of research with a substantial history.

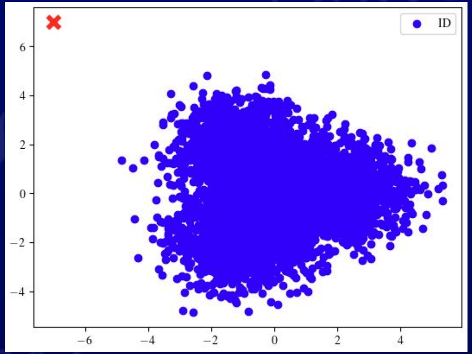
<u>Perhaps ML can offer additional perspective.</u>

<u>**Problem**</u>: ML models often are overconfident about predictions in regions it's never trained on.

<u>**Question**</u>: How do we teach an ML model what it does and doesn't know?

Evidence-based learning!

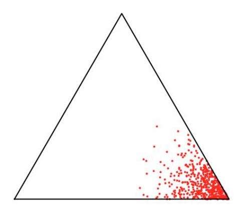
Evidential Deep Learning



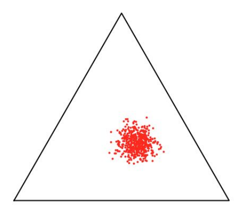
Prior Networks - making the implicit explicit

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

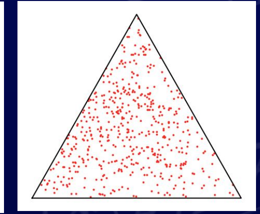
By factorizing the predictive posterior distribution, we can explicitly model the dependence on the prior of the ensemble. In a single forward pass, we can describe aleatoric, epistemic, and distributional uncertainties.



ID
low error



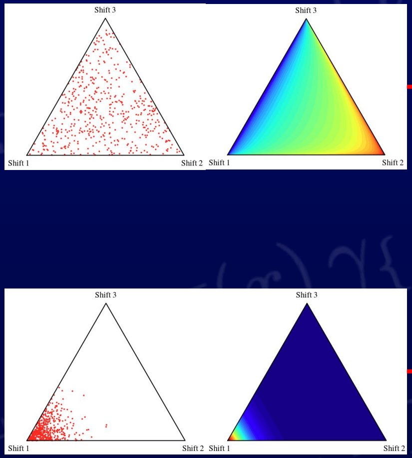
ID
high error

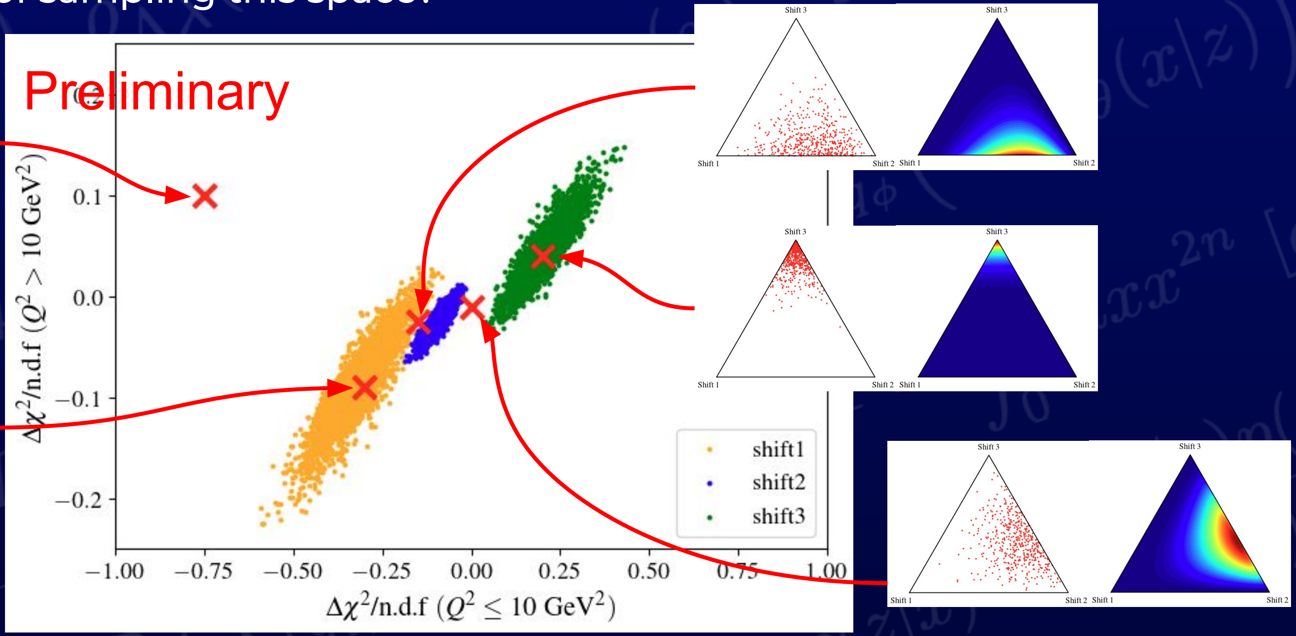


OOD

Evidential Deep Learning for GNIs

Consider a BSM scenario in which, due to the presence of non-standard interactions in a weak effective field theory Lagrangian, the values of the CKM matrix elements can appear shifted from their fitted standard values resulting in a shift of the χ^2 against a dataset. How to quantify uncertainty of sampling this space?





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

Conclusions

ML in HEP and NP is nascent, not to mention theory applications, but is expected to grow substantially.

The research I have discussed here - solving large scale inverse problems for robust insights into hadron structure - is the nucleus of a wide-reaching program.

The activity fills an urgent need for more direct collaboration between theorists, data scientists, and computational experts.

Thank you for your attention!