Quantifying uncertainty in ML-based global analyses

Manal Almaeen

Old Dominion University

CENTER *for* NUCLEAR FEMTOGRAPHY

Generalized Parton Distributions and Global Analysis workshop
June 13, 2023

Motivation: uncertainty in learning

Self driving cars

Healthcare

Face detection

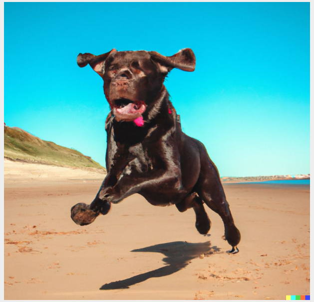
Security

Manufacturing

Finance

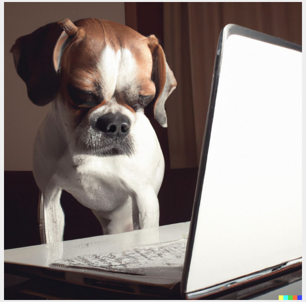
Neural Networks: training vs testing

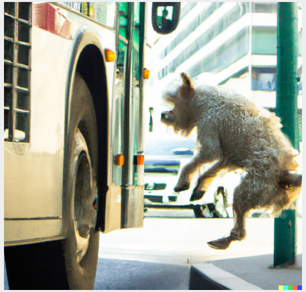
Training on your dataset

Expectation

Testing in reality



Reality

Types of The Uncertainty

Uncertainty in Machine learning:

- Aleatoric
- Epistemic

Uncertainty in physics

- Systematic
- Statistical

Types of The Uncertainty

Uncertainty in Machine learning:

- **Aleatoric (data uncertainty):**

→ Uncertainty coming from corrupted data (detector noise).

- **Epistemic (model uncertainty):**

→ Uncertainty coming from imperfect models.

Uncertainty in physics:

- **Statistical:**

→ Uncertainty that can be statistically determined from input data (std of measurements).

- **Systematic:**

→ Uncertainty coming from theory or other measurements (not statistical).

Types of The Uncertainty

Uncertainty in Machine learning:	Uncertainty in physics:
● **Aleatoric (data uncertainty):** ➔ Uncertainty coming from corrupted data (detector noise).	● **Statistical:** ➔ Uncertainty that can be statistically determined from input data (std of measurements).
● **Epistemic (model uncertainty):** ➔ Uncertainty coming from imperfect models.	● **Systematic:** ➔ Uncertainty coming from theory or other measurements (not statistical).

Types of The Uncertainty

Uncertainty in Machine learning:

- **Aleatoric (data uncertainty):**

➔ Uncertainty coming from corrupted data (detector noise).

- **Epistemic (model uncertainty):**

➔ Uncertainty coming from imperfect models.

Uncertainty in physics:

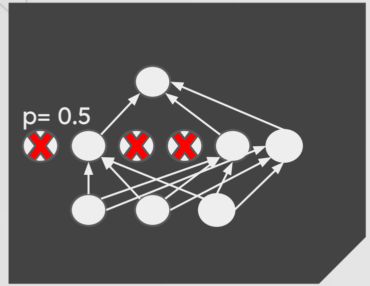
- **Statistical:**

➔ Uncertainty that can be statistically determined from input data (std of measurements).

- **Systematic:**

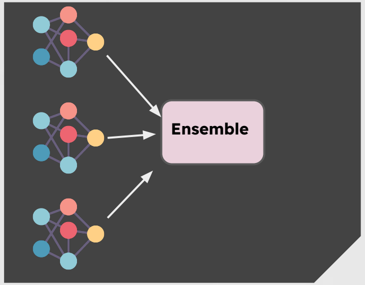
➔ Uncertainty coming from theory or other measurements (not statistical).

Uncertainty Quantification Methods for Machine Learning



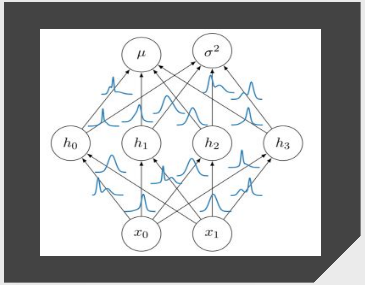
Dropout

- Introduced for regularization.
- Adopted for UQ – MC Dropout.
- Enabling Dropout during inference



Deep Ensembles

- Training the same architecture several times with different initializations.



Bayesian Neural Network

- NN with the weights of the layers form a probability distribution.

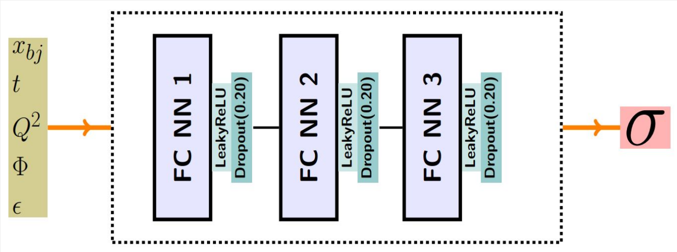
Figure source: Cabiscol, J.A. (2019). Understanding Uncertainty in Bayesian Neural Networks.

Quantifying the Uncertainty of DVCS Cross Sections

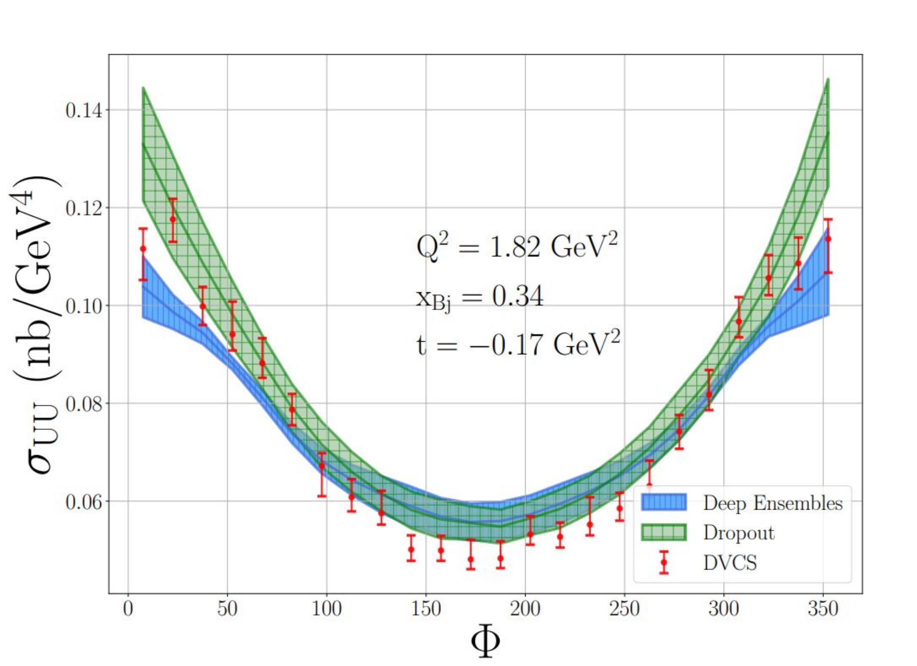
UQ on predicting the cross sections

from the kinematics using:

- ➔ Deep Ensembles
- ➔ Dropout



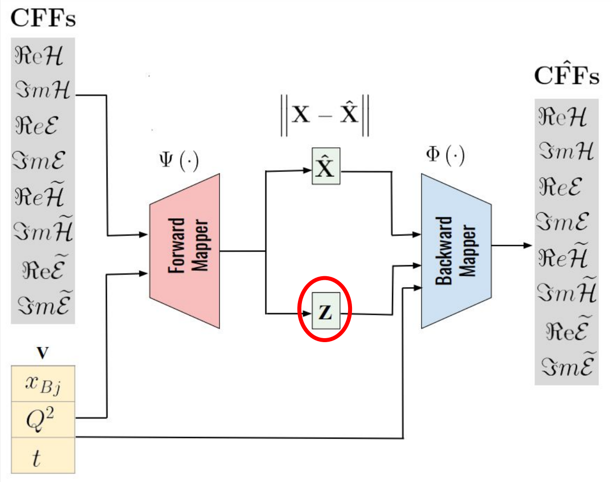
NNs architecture for DVCS cross sections.



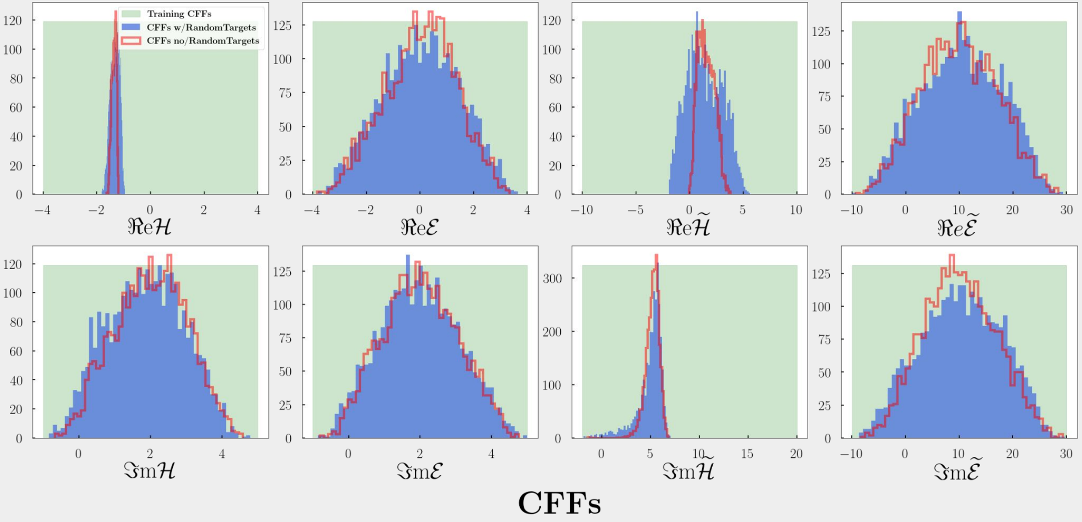
M. Almaeen, J. Grigsby, J. Hoskins, B. Kriesten, Y. Li, H-W. Lin, S. Liuti arXiv:2207.10766

Quantifying the Uncertainty of the CFFs

$x_{Bj} = 0.343$, $t = -0.172$, $Q^2 = 1.82$, $E_b = 5.75$



C-VAIM architecture for extracting CFFs.



Random target → latent space z + incorporating the DVCS error bar.

No random target → latent space z

M. Almaeen, J. Hoskins, B. Kriesten, Y. Li, H-W. Lin, S. Liuti (in progress)

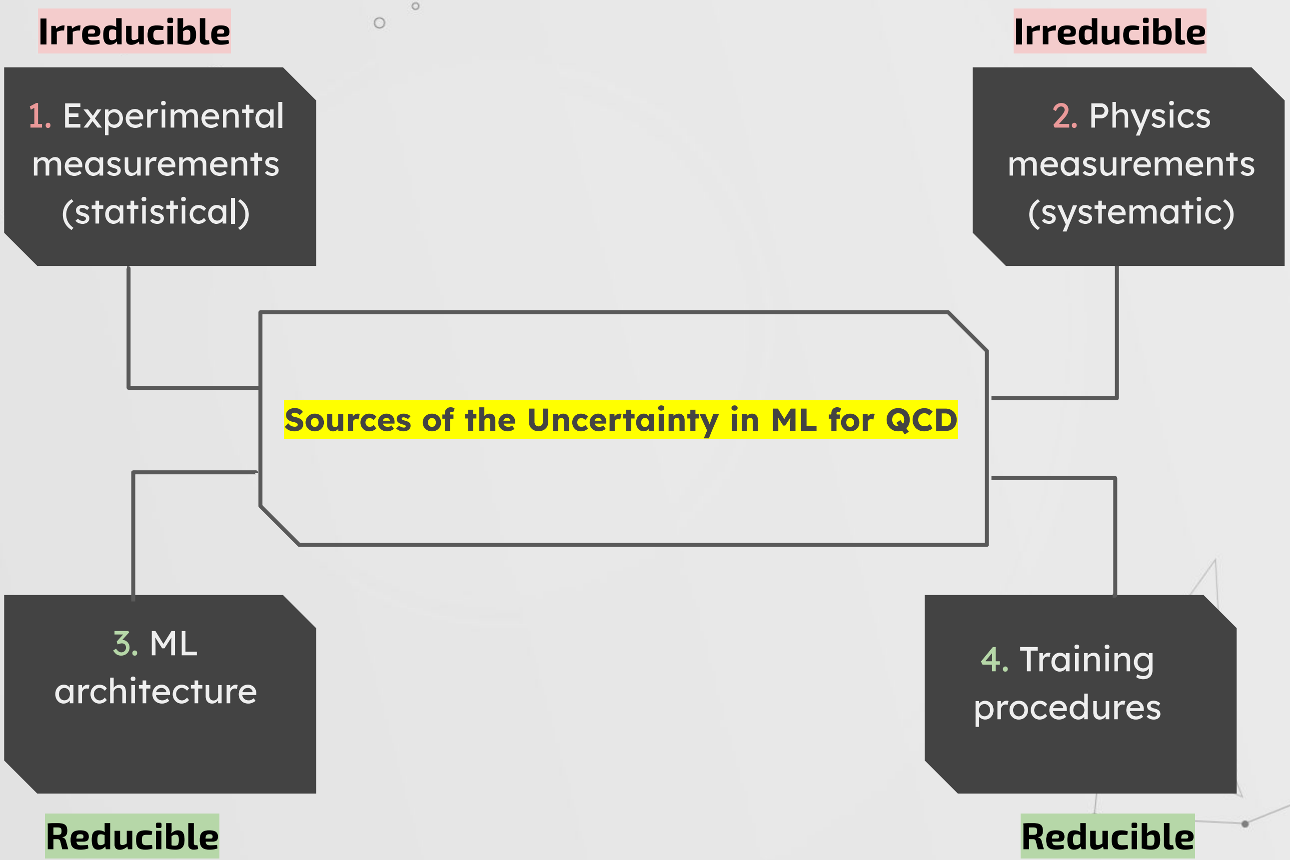
Sources of the Uncertainty in ML for QCD

1. Experimental measurements (statistical)

2. Physics measurements (systematic)

3. ML architecture

4. Training procedures


Irreducible
1. Experimental measurements (statistical)
Irreducible
2. Physics measurements (systematic)
Sources of the Uncertainty in ML for QCD
3. ML architecture
Reducible
4. Training procedures
Reducible

How can we reduce the model uncertainty?

- **Active Learning**
    - Collecting sufficient data for modeling is challenging.
    - AL can reduce the number of training samples.
    - AL is basically categorized into two types:
        - Population based AL.
        - Pool based AL.

- # Active Learning

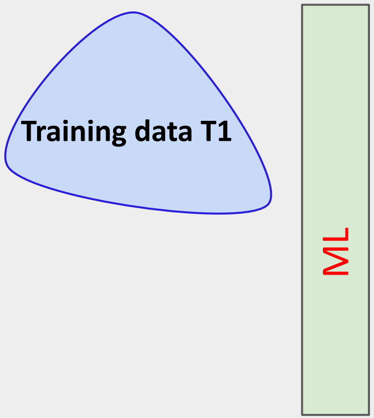
G1 $=\{\alpha_u = 0.5, \beta_u = 2.5, \alpha_d = 0.1, \beta_d = 3.0\}$

Training data T1

1st level

• Active Learning

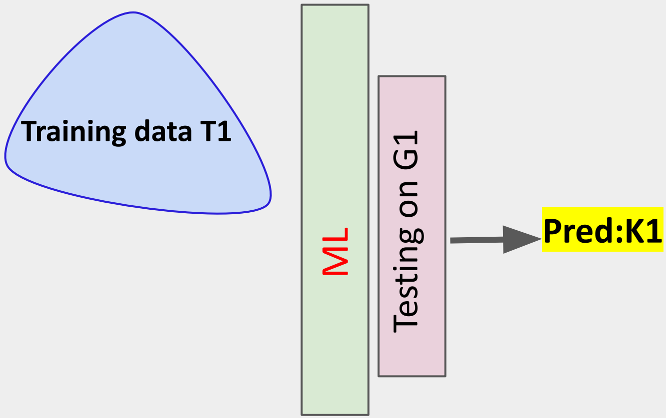
G1 $=\{\alpha_u = 0.5, \beta_u = 2.5, \alpha_d = 0.1, \beta_d = 3.0\}$



1st level

- # Active Learning

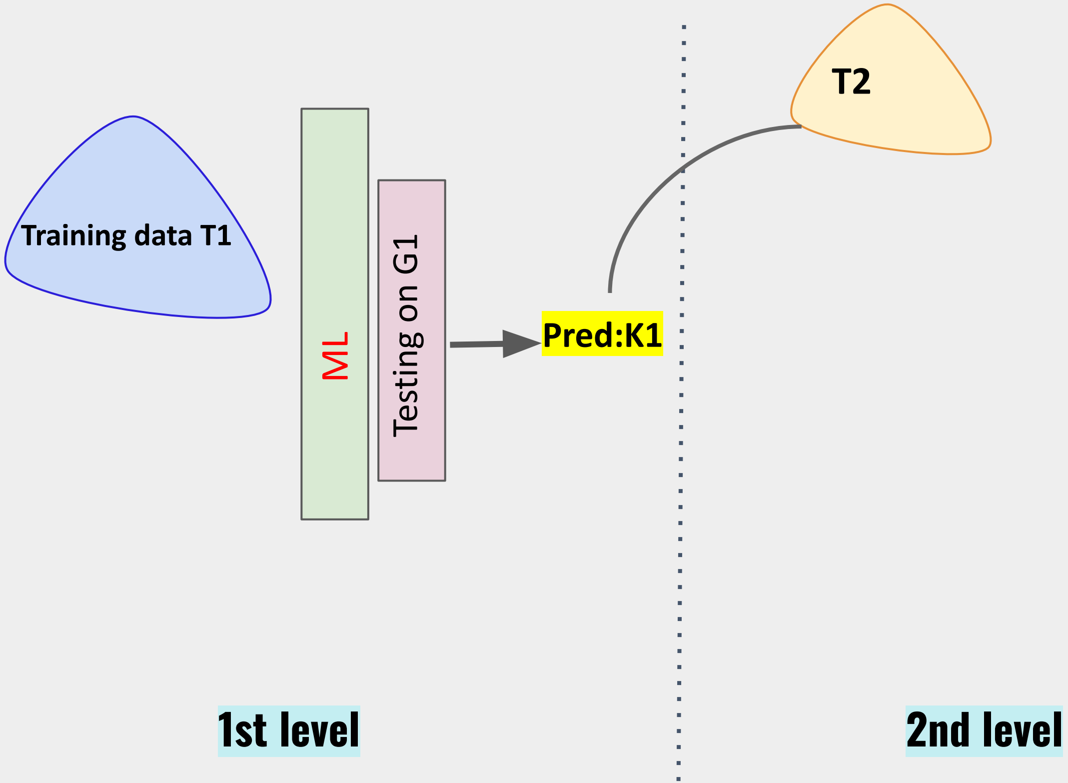
$\mathbf{G1} = \{\alpha_u = 0.5, \beta_u = 2.5, \alpha_d = 0.1, \beta_d = 3.0\}$



1st level

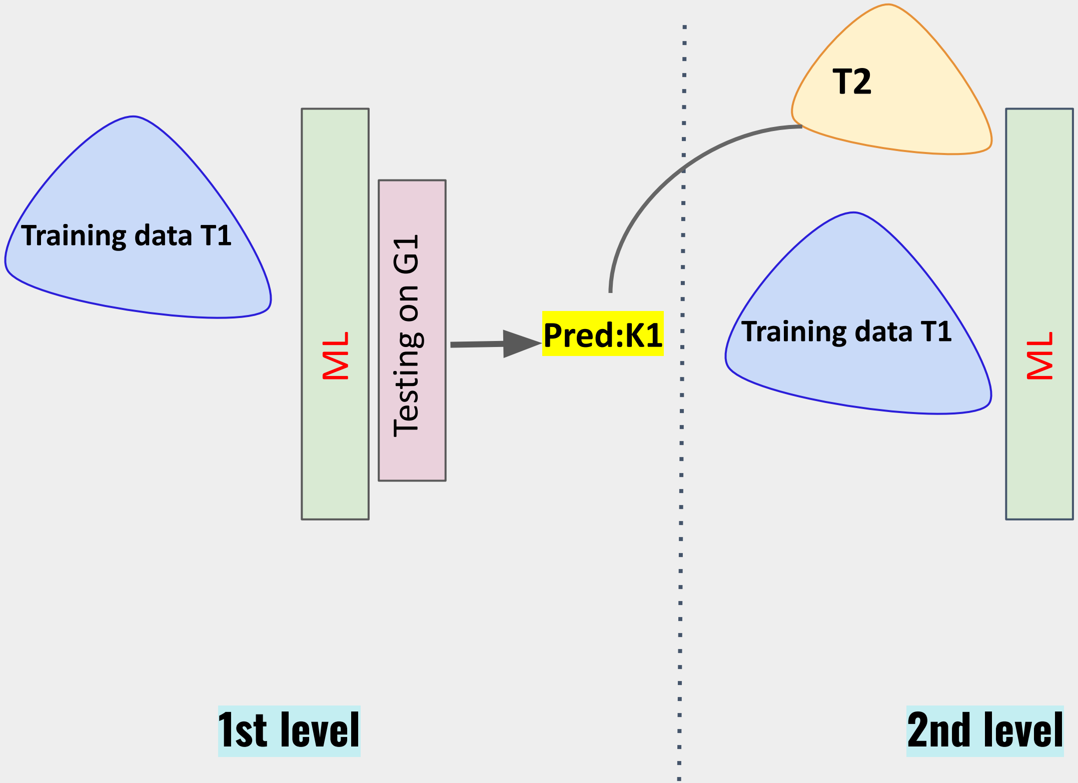
- # Active Learning

G1 =$\{\alpha_u = 0.5, \beta_u = 2.5, \alpha_d = 0.1, \beta_d = 3.0\}$



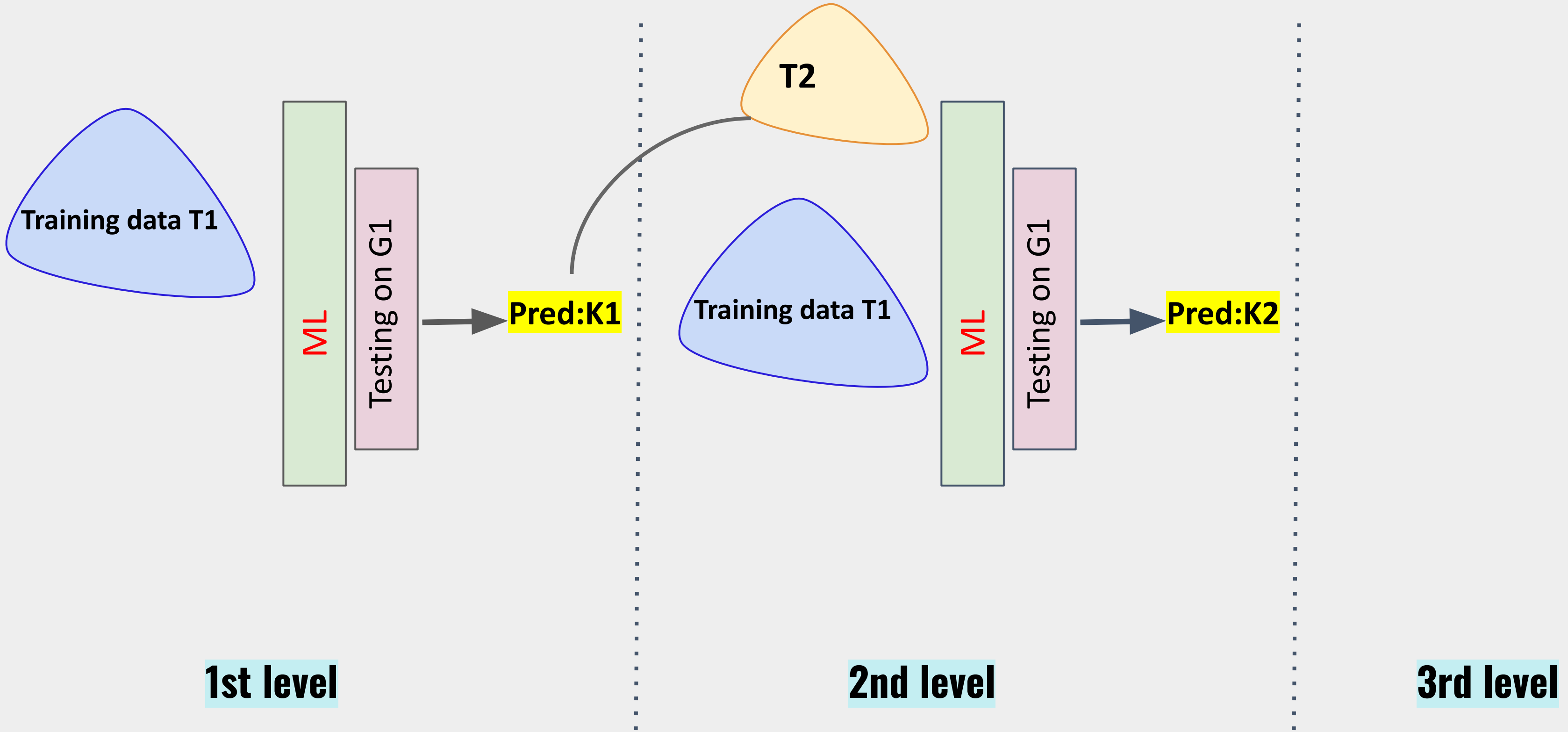
- # Active Learning

$\mathbf{G1} = \{\alpha_u = 0.5,\ \beta_u = 2.5,\ \alpha_d = 0.1,\ \beta_d = 3.0\}$



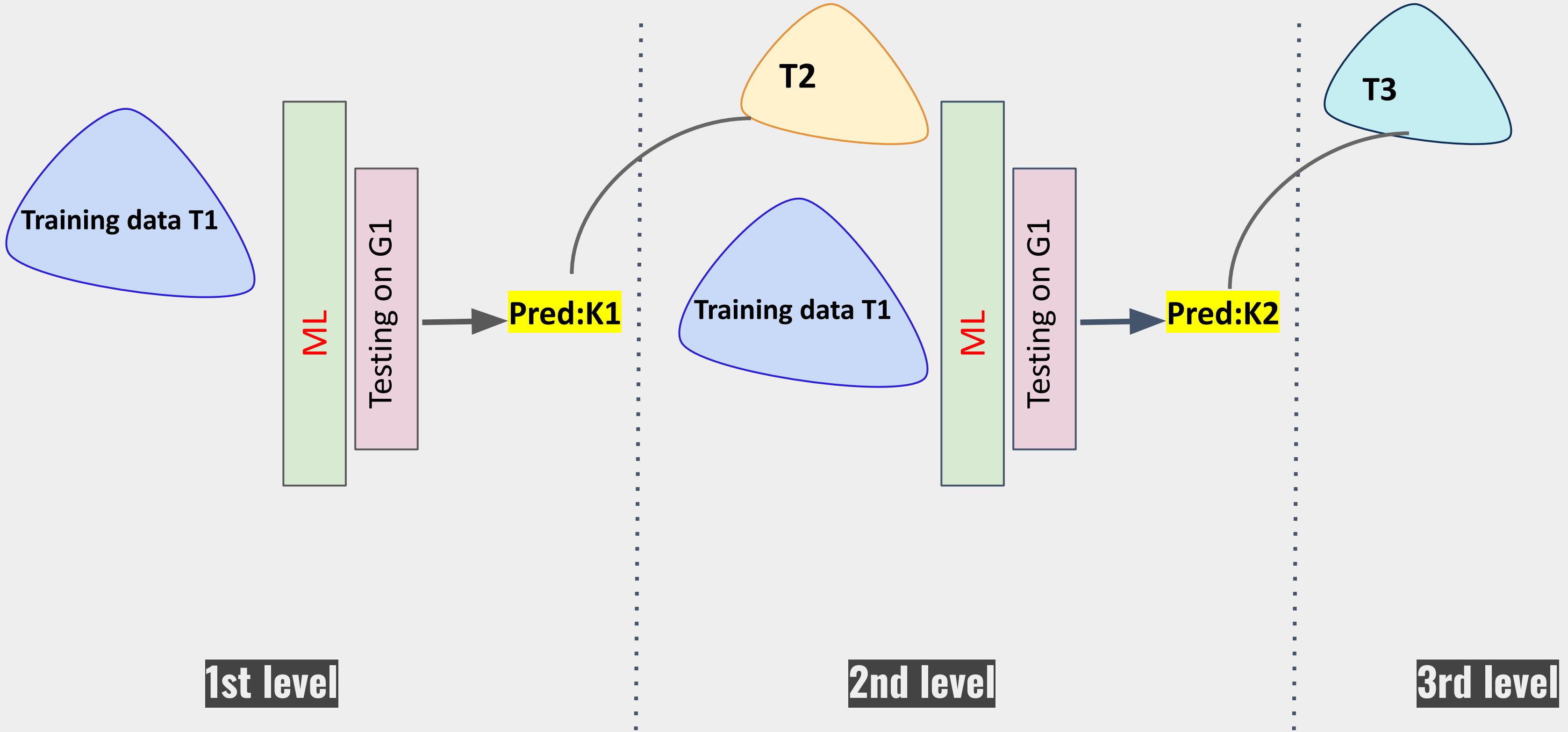
- ## Active Learning

G1 $=\{\alpha_u = 0.5, \beta_u = 2.5, \alpha_d = 0.1, \beta_d = 3.0\}$



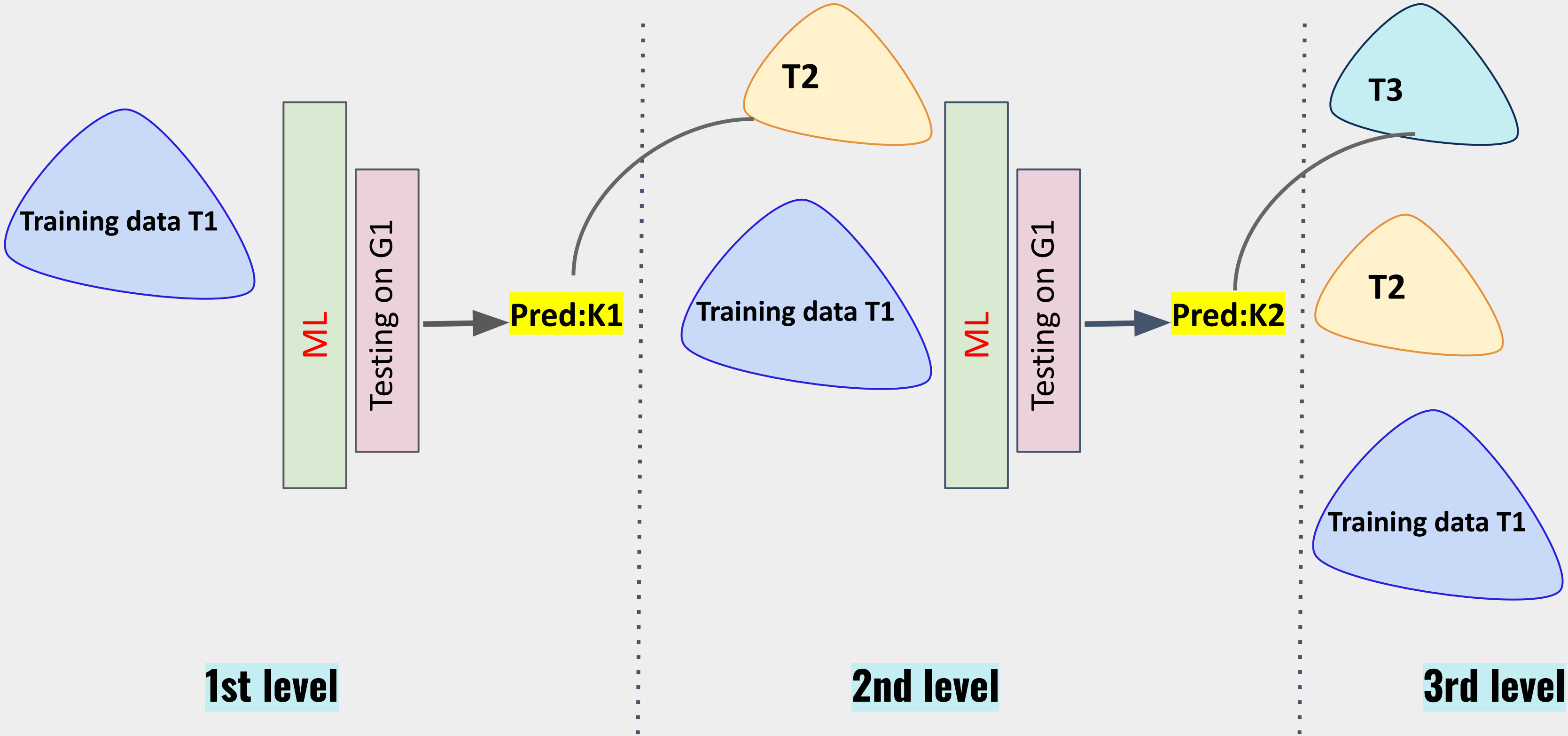
• Active Learning

G1 $=\{\alpha_u = 0.5, \beta_u = 2.5, \alpha_d = 0.1, \beta_d = 3.0\}$



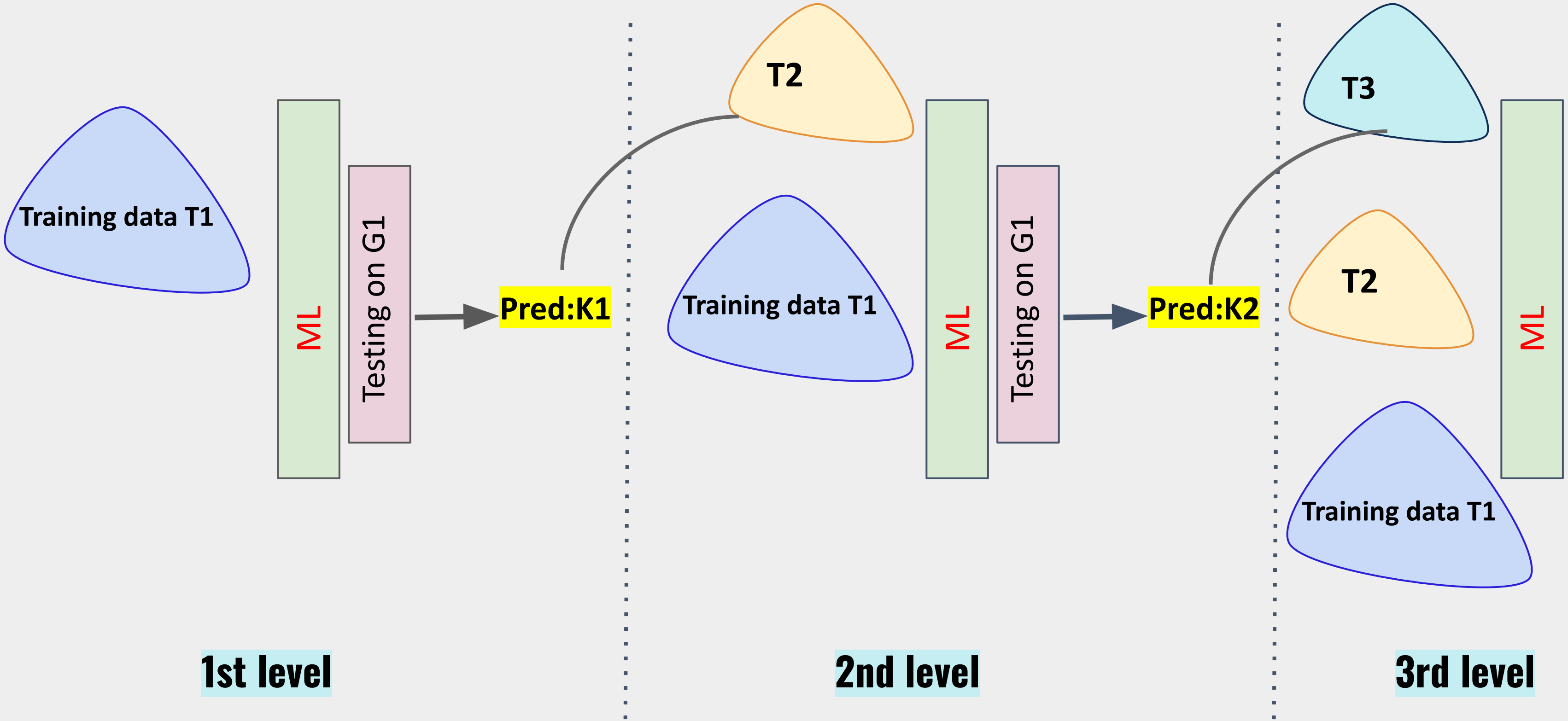
• Active Learning

G1 $=\{\alpha_u = 0.5, \beta_u = 2.5, \alpha_d = 0.1, \beta_d = 3.0\}$



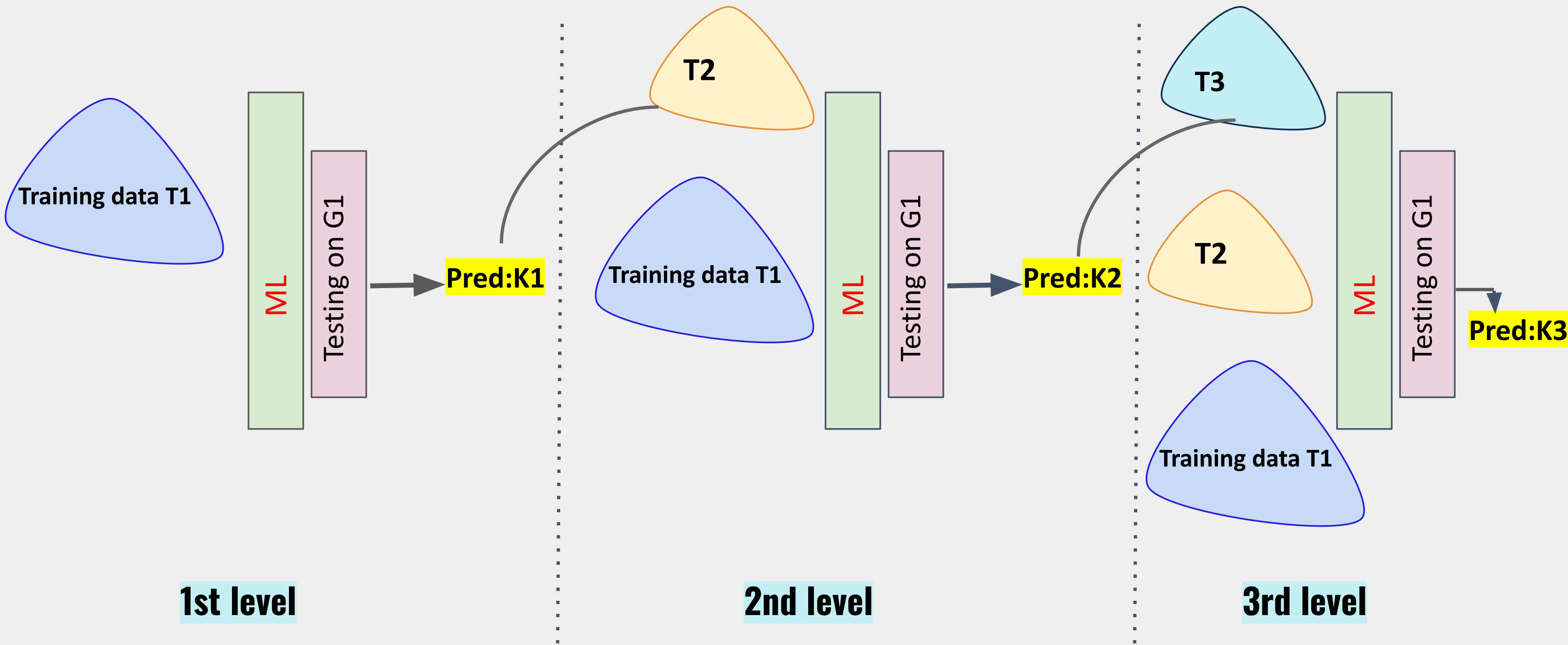
• Active Learning

G1 $=\{\alpha_u = 0.5, \beta_u = 2.5, \alpha_d = 0.1, \beta_d = 3.0\}$

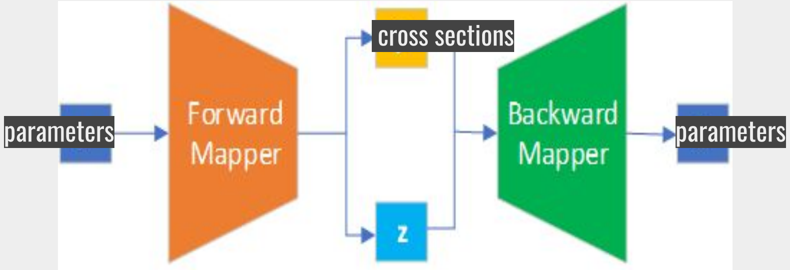


- # Active Learning

G1 =$\{\alpha_u = 0.5, \beta_u = 2.5, \alpha_d = 0.1, \beta_d = 3.0\}$

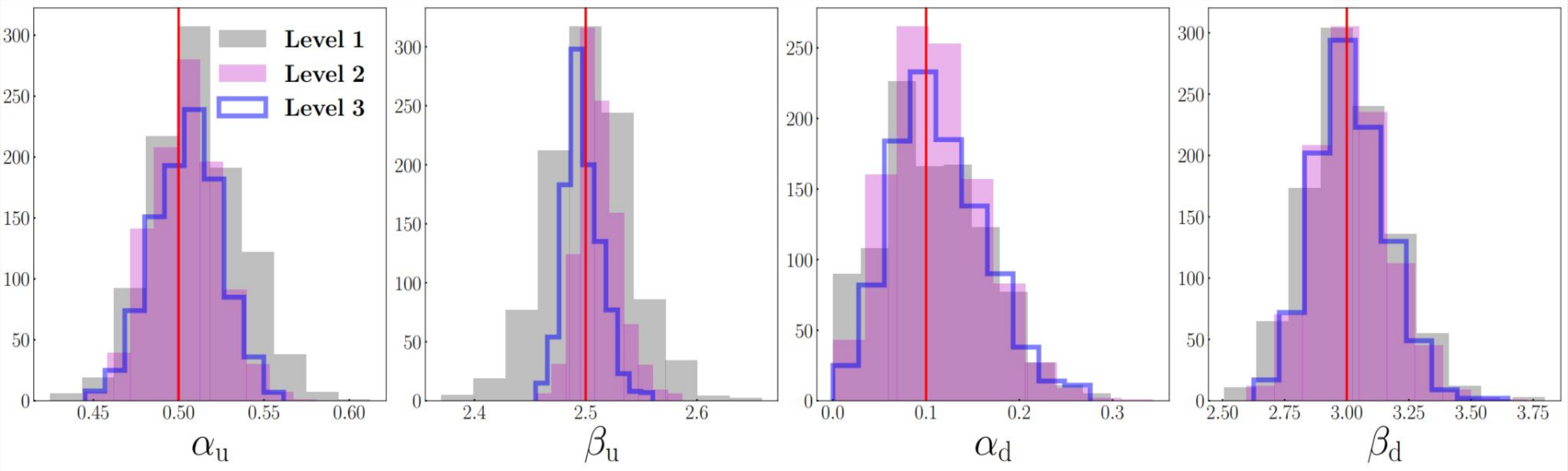


• Active Learning



➔ Applied to extracting the parton distribution function (PDF) parameters from the cross sections

G1 $=\{\alpha_u = 0.5, \beta_u = 2.5, \alpha_d = 0.1, \beta_d = 3.0\}$

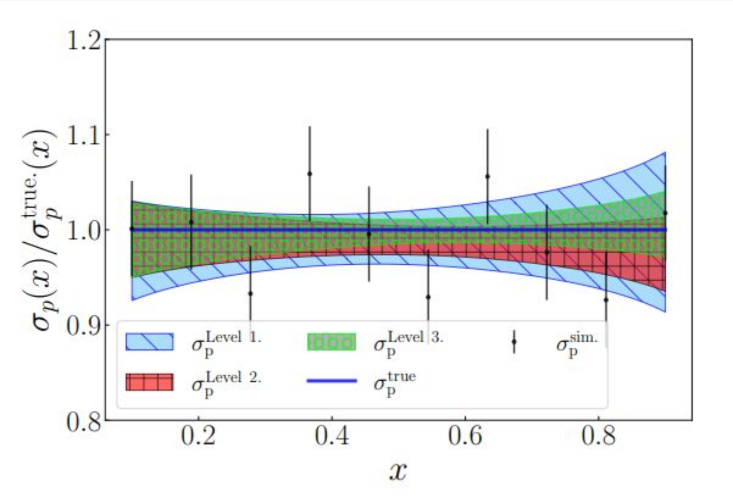


Active Learning

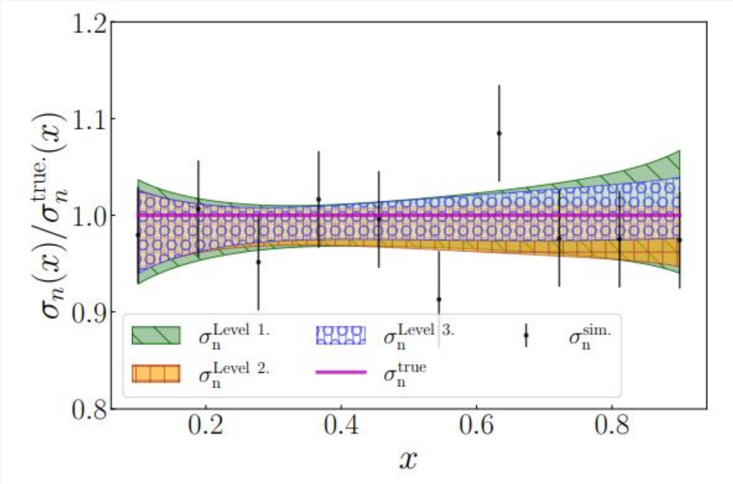
G1 $=\{\alpha_u = 0.5, \beta_u = 2.5, \alpha_d = 0.1, \beta_d = 3.0\}$

➔ Applied to extracting the parton distribution function (PDF) parameters from the cross sections

Reconstructed cross sections



σ_p



σ_n

Summary

- ➢ Uncertainty quantification methods: Dropout, Deep Ensembles and BNNs.
- ➢ UQ on the DVCS cross sections
- ➢ UQ on the CFFs
- ➢ Reducing the model uncertainty: Active Learning

Acknowledgement

DVCS + CFFs

Yaohang Li

Simonetta Liuti, Joshua Hoskins, Jake Grigsby.

Brandon Kriesten

Huey-Wen Lin

Active Learning

Nobuo Sato, W. Melnitchouk, Yasir Alanazi

References

1. Gawlikowski, J., Tassi, C.R., Ali, M., Lee, J., Humt, M., Feng, J., Kruspe, A.M., Triebel, R., Jung, P., Roscher, R., Shahzad, M., Yang, W., Bamler, R., & Zhu, X. (2021). A Survey of Uncertainty in Deep Neural Networks. *ArXiv, abs/2107.03342.*

2. Caldeira, J. o., & Nord, B. (2020). Deeply uncertain: comparing methods of uncertainty quantification in deep learning algorithms. *Machine Learning: Science and Technology*, *2*(1), 015002. doi:10.1088/2632-2153/aba6f3.

3. M. Almaeen, Y. Alanazi, N. Sato, W. Melnitchouk, M. P. Kuchera and Y. Li, "Variational Autoencoder Inverse Mapper: An End-to-End Deep Learning Framework for Inverse Problems," *2021 International Joint Conference on Neural Networks (IJCNN)*, Shenzhen, China, 2021, pp. 1-8, doi: 10.1109/IJCNN52387.2021.9534012.

4. Almaeen, M., Grigsby, J., Hoskins, J., Kriesten, B., Li, Y., Lin, H.-W., & Liuti, S. (2022). Benchmarks for a Global Extraction of Information from Deeply Virtual Exclusive Scattering. *arXiv [Hep-Ph]*. Retrieved from http://arxiv.org/abs/2207.10766.