# Machine learning for Deeply Virtual Compton Scattering

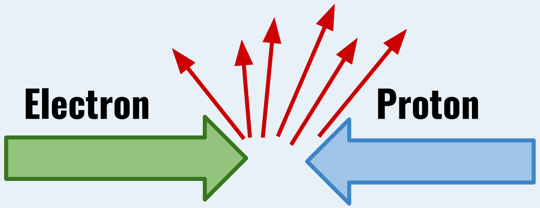
Manal Almaeen (malma004@odu.edu )
PhD Candidate
Old Dominion University, Department of Computer Science

Information and Statistics for Nuclear Experiment and Theory workshop (ISNET-9)

Washington University in St. Louis

# Physics Motivation

## How do we get **hadron structure**?




**Quantum Correlation Functions (QCF):**

➔ Parton distribution functions(PDF)

➔ Parton to hadron fragmentation functions (FFs)

➔ Transverse momentum dependent distributions (TMDs)

➔ **Generalized parton distributions (GPDs)**

# How to measure GPDs? Deeply virtual Compton scattering

- **DVCS** is known to probe generalized parton distributions
- **GPDs** extraction is a challenging problem! This problem can be divided into several levels from experimental cross sections to the physical properties of the hadron

Physics Constrained Neural Networks for DVCS cross sections

Step 1

Variational Autoencoder Inverse Mapper for Compton Form factors extraction

Step 2

GPDs Modeling

Step 3

Hadron Structure

Step 4

# How to measure GPDs? Deeply virtual Compton scattering

- **DVCS** is known to probe generalized parton distributions
- **GPDs** extraction is a challenging problem! This problem can be divided into several levels from experimental cross sections to calculating the physical properties of the hadron

Physics Constrained Neural Networks for DVCS cross sections

**Step 1**

Variational Autoencoder Inverse Mapper for Compton Form factors extraction

**Step 2**

GPDs Modeling

**Step 3**

Hadron Structure

**Step 4**

# Physics Constrained Neural Networks

# Physics Constrained Neural Networks (PCNNs):

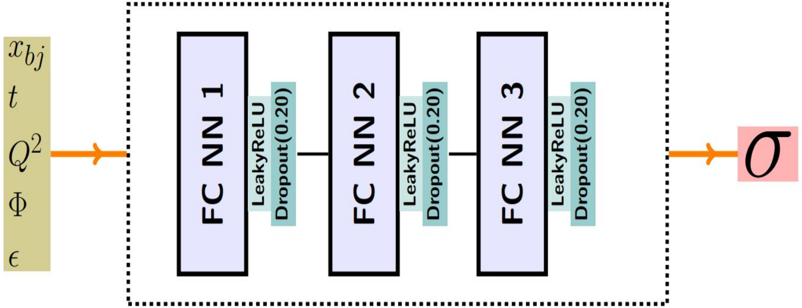
- **What is PCNNs?**

  NNs integrate data and physics knowledge

- **Why incorporating physics to ML?**
  - Data size
  - Generalization / performance

# PCNNs: Model Architecture

- **Physics constraints:**
  - DVCS Error bar :
    - Each DVCS data point includes a mean and a standard error indicating its uncertainty.
    - We treat this error as a source of information through the use of data augmentation.
  - Angular Symmetry
    - For the unpolarized angular symmetry, we use additional loss : $\| f(X_{bj}, t, Q^2, \phi, \epsilon) - f(X_{bj}, t, Q^2, -\phi, \epsilon) \|$



# PCNNs: Methods

- Data Description

❖ Each record of the dataset has [xbj, t, Q2,Eb, φ, L, σ, Δσ, δσ]

❖ where L is a label for polarization of beam and target -

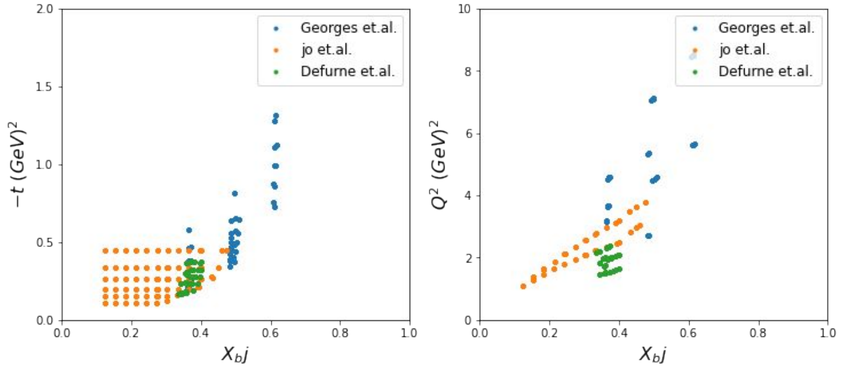
  ➢ 1 is unpolarized beam/unpolarized target

  ➢ 2 is polarized beam/unpolarized target

❖ σ is the cross section value

❖ Δσ is the statistical error

❖ δσ is the systematic error.

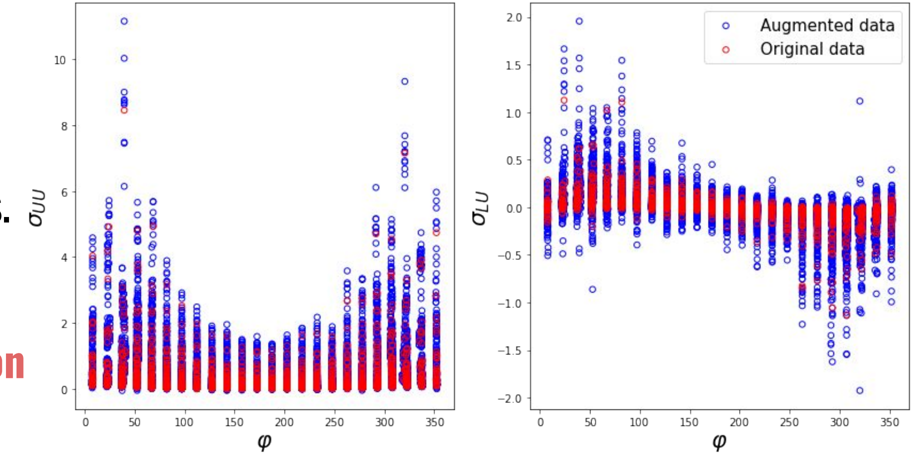



(Left) Kinematic region in xBj and t. (Right) xBj and Q 2 where experimental measurements used in this analysis have been taken at 6 GeV and 12 GeV.

# PCNNs: Methods

- Data Augmentation
  - We have 3,862 unpolarized and 3,884 polarized data points.
  - Data augmentation! Using statistical errors.

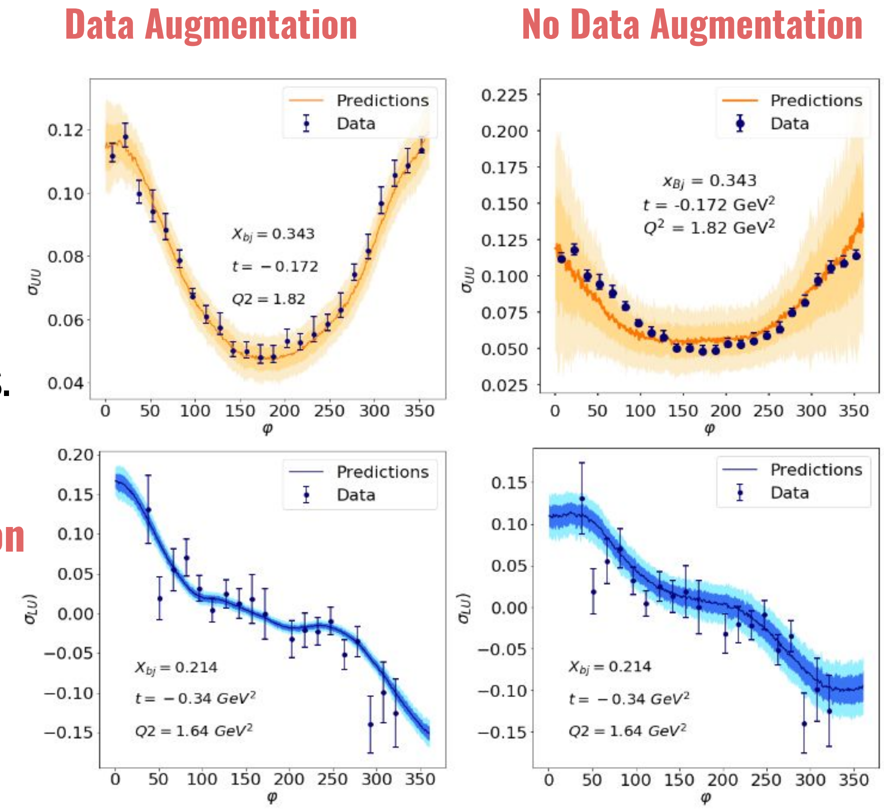
➔ This is additional source of physics information



# PCNNs: Methods

- Data Augmentation
  - We have 3,862 unpolarized and 3,884 polarized data points.
  - Data augmentation! Using statistical errors.

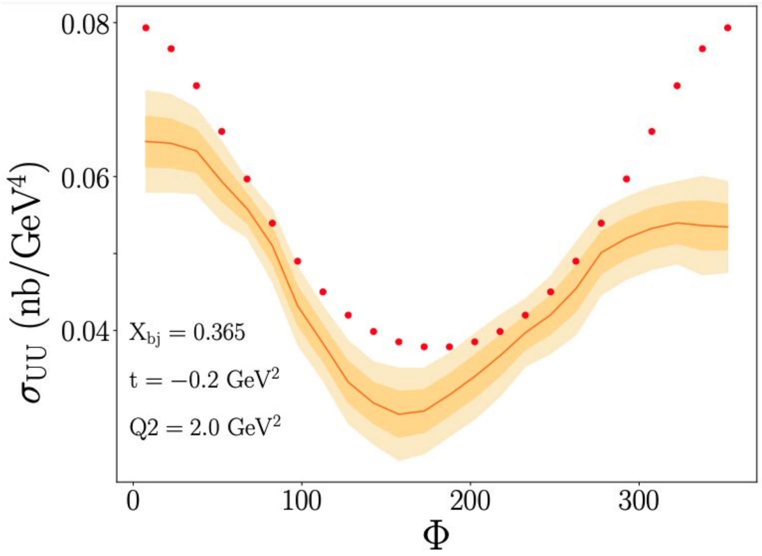
➔ This is additional source of physics information



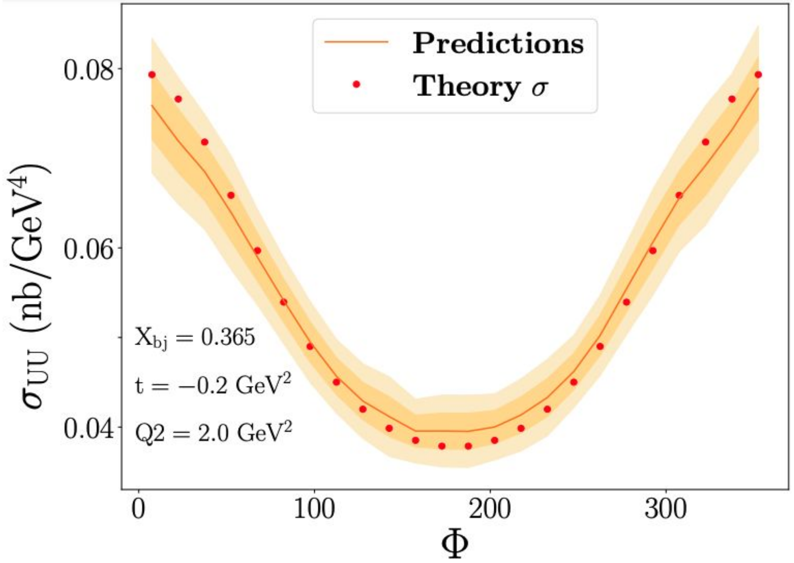

Visualization script obtained from FemtoNet group, UVA

# PCNNs: Results

- Testing the PCNNs ability to generalize to kinematics outside of the range covered in BH cross sections.



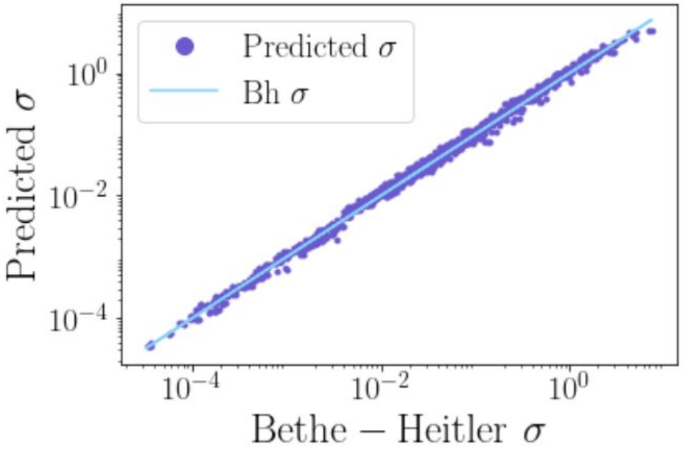

**NNs**




**PCNNs**

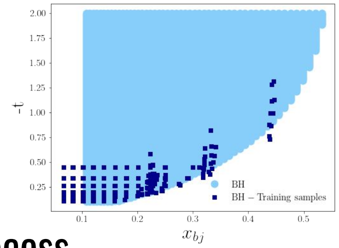
# PCNNs: Results

- Testing the PCNNs ability to generalize using pseudo data for the Bethe-Heitler (BH) process.

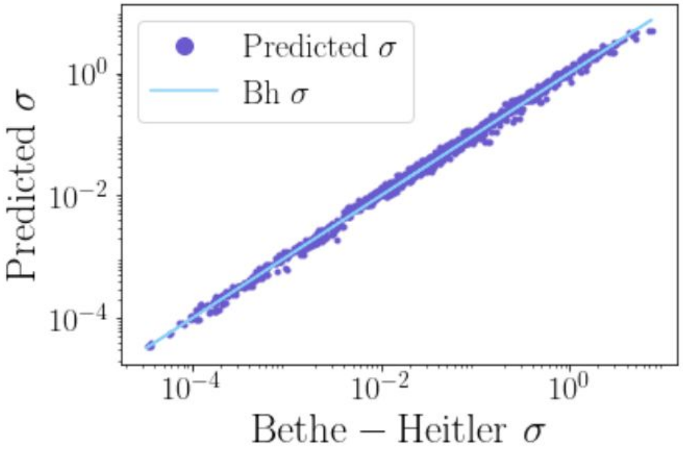



**PCNNs on the test set**

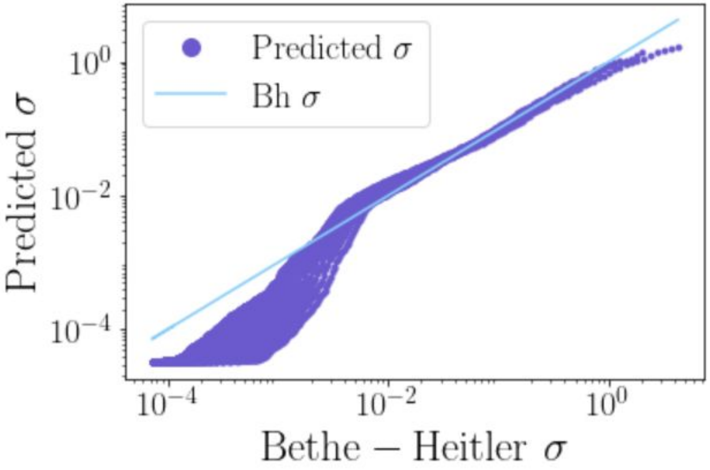
# PCNNs: Results




- Testing the PCNNs ability to generalize using pseudo data for the Bethe-Heitler (BH) process.




**PCNNs on the test set**

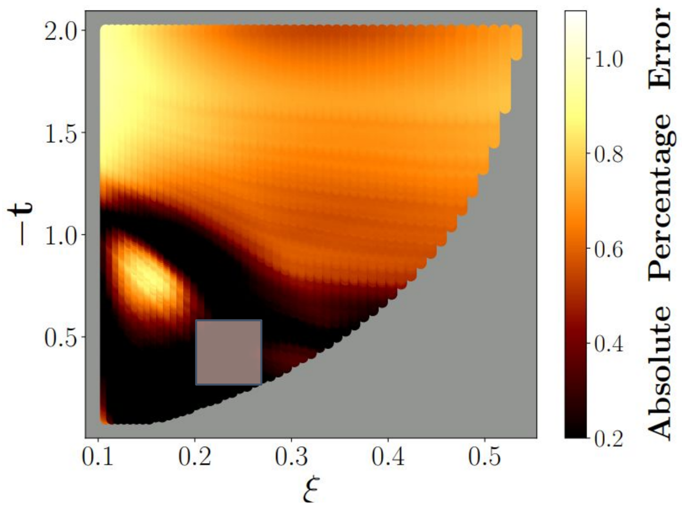



**PCNNs on generalization task with samples outside of the training sample**

# PCNNs: Results

Train on kinematics ($Q^2$ = 2, Eb =6 , $\xi$ = 0.2 - 0.3 and -t = 0.17-0.4)

- Testing the PCNNs ability to generalize using BH and DVCS data.

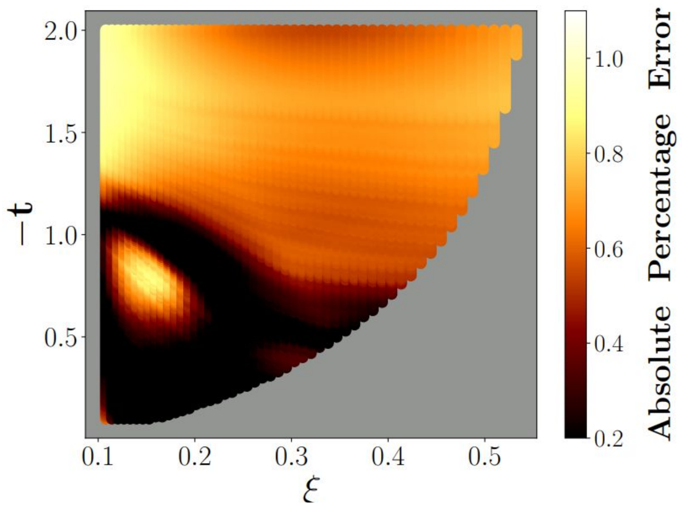



BH

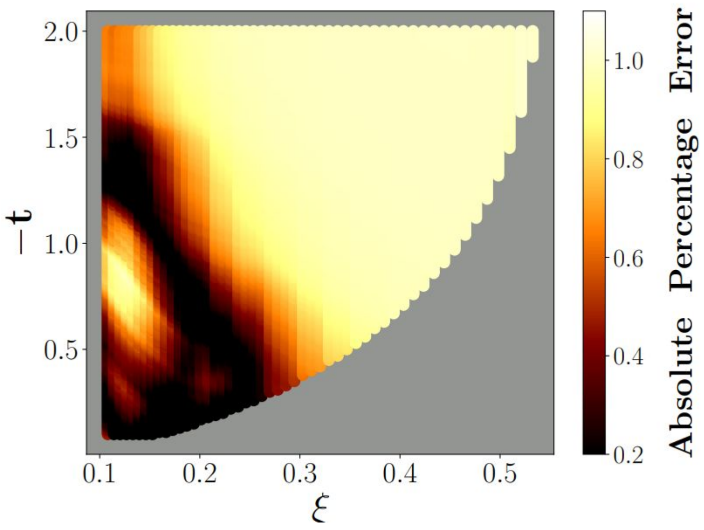
# PCNNs: Results

Train on kinematics ($Q^2$ = 2, Eb =6 , $\xi$ = 0.2 - 0.3 and -t = 0.17-0.4)

- Testing the PCNNs ability to generalize using BH and DVCS data.




BH

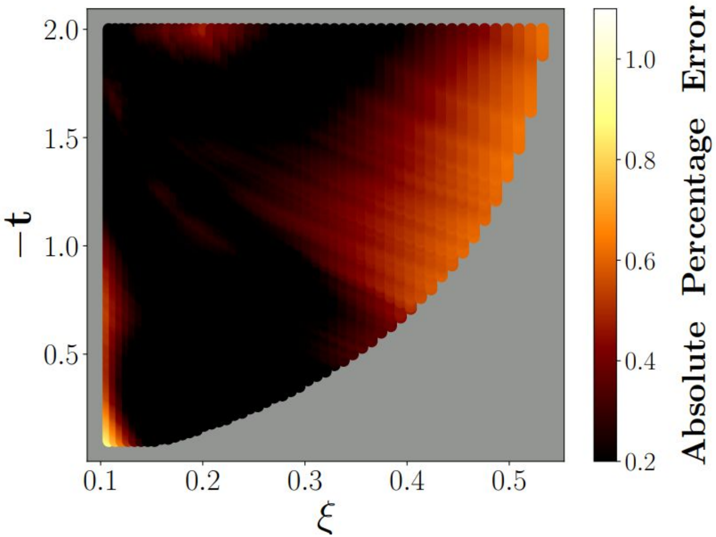



DVCS

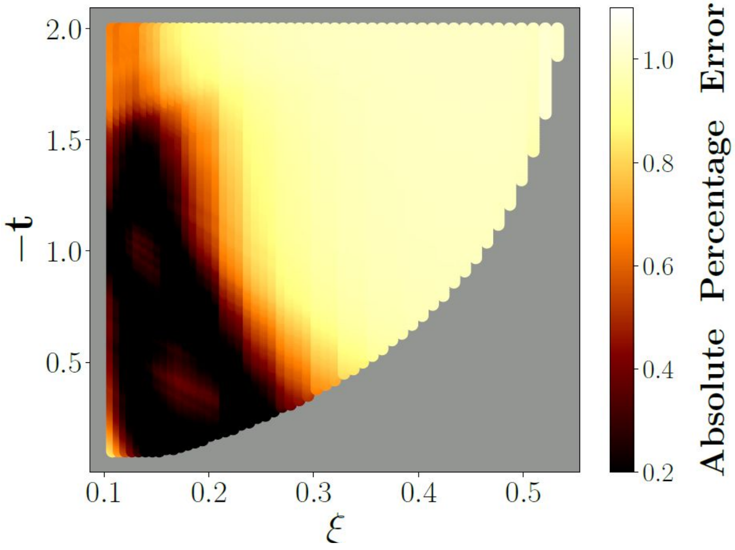
# PCNNs: Results

$$\mathcal{L} = \frac{1}{n}\sum_{\theta_t}\frac{(y - f_{\theta_t}(x))^2}{R_{\theta_t}}$$

- Incorporating the **statistical error** to BH and DVCS training.




**BH-**combine the statistical error to the loss function

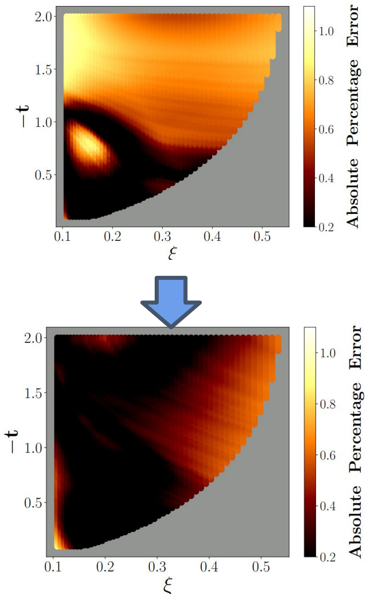



**DVCS-** combine the statistical error to the loss function

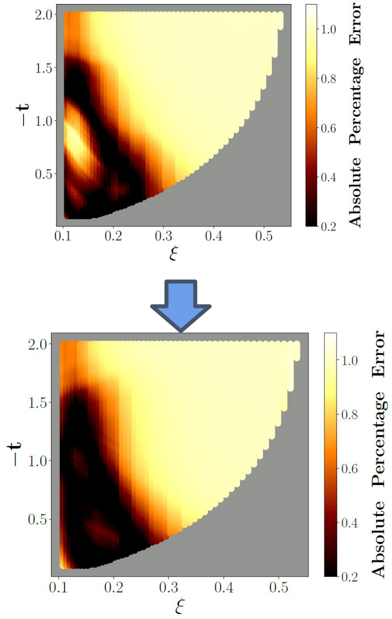
# PCNNs: Results

$$\mathcal{L} = \frac{1}{n} \sum_{\theta_t} \frac{(y - f_{\theta_t}(x))^2}{R_{\theta_t}}$$

- Incorporating the **statistical error** to BH and DVCS training.




**BH-**combine the statistical error to the loss function




**DVCS-** combine the statistical error to the loss function

# Compton Form Factors Extractions

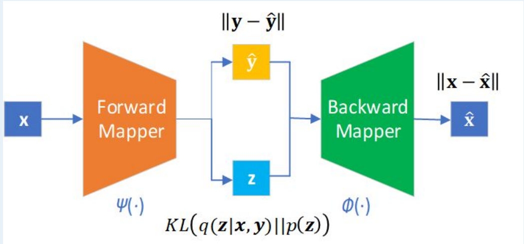
# CFFs extraction:

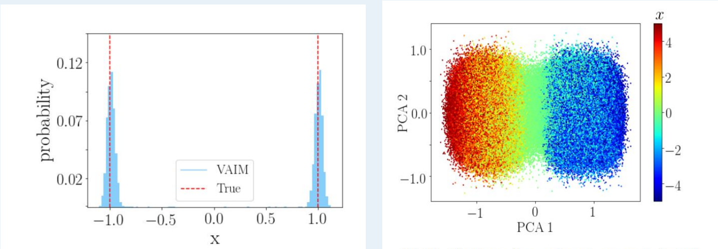
DVCS has been identified as the "golden channel" for the extraction of information on partonic 3D dynamics in the nucleon.

## Inverse problem:

**cross sections** → **CFFs**

## Variational Autoencoder Inverse Mapper (VAIM)[1]




(a) Predicted distribution of $x$ when $f(x) = 1$.

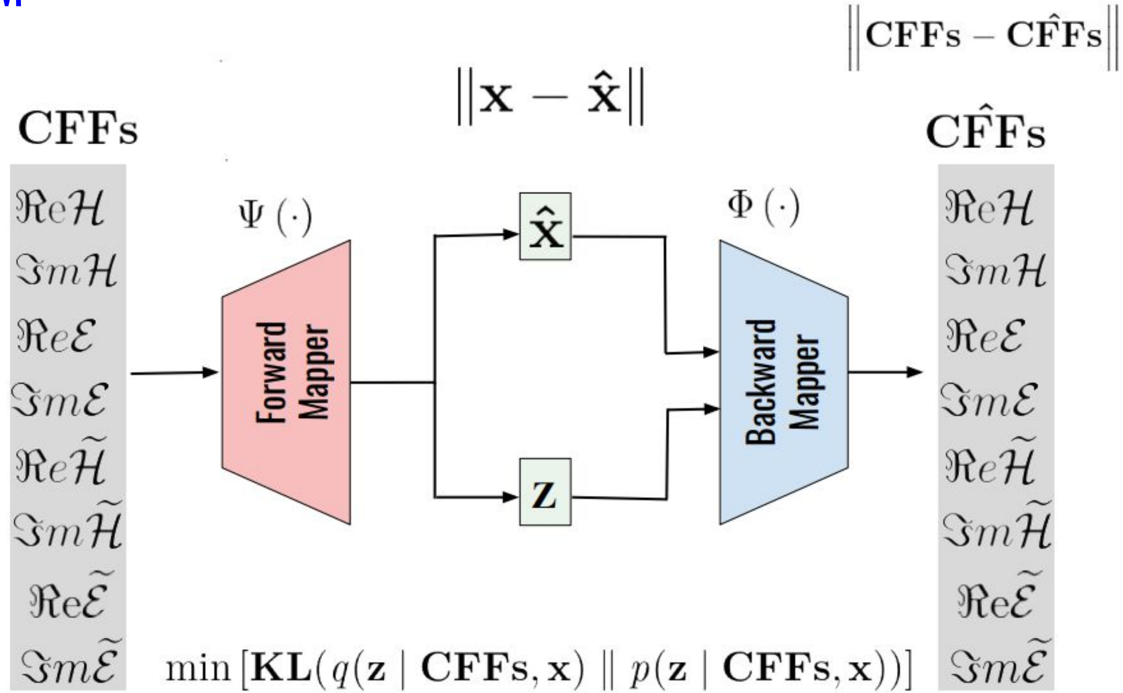
(b) Distribution of $x$ on latent space projected on the first two principle components.


[1] M. Almaeen, Y. Alanazi, N. Sato, et al., "Variational Autoencoder Inverse Mapper: An End-to-End Deep Learning Framework for Inverse Problems", in 2021 International Joint Conference on Neural Networks (IJCNN) (2021).

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

# CFFs extraction: Methods

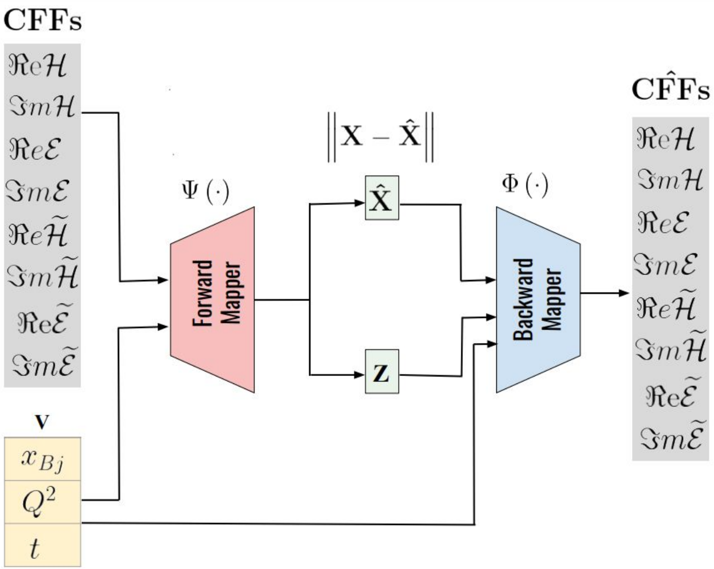
## A. VAIM



# CFFs extraction: Methods

| | $x_bj$ | $t$ | $Q^2$ |
|---|---|---|---|
| 1 | 0.343 | -0.172 | 1.820 |
| 2 | 0.368 | -0.232 | 1.933 |
| 3 | 0.375 | -0.278 | 1.964 |
| 4 | 0.379 | -0.323 | 1.986 |
| 5 | 0.381 | -0.371 | 1.999 |

## B. Conditional-VAIM (C-VAIM)



# CFFs extraction: Methods

## B. Conditional-VAIM (C-VAIM)

- Learn an approximate distribution $q(\mathbf{z} \mid \mathbf{cff}, \mathbf{v}, \mathbf{x})$

$$q(\mathbf{z} \mid \mathbf{cff}, \mathbf{v}, \mathbf{x}) \sim p(\mathbf{z} \mid \mathbf{cff}, \mathbf{v}, \mathbf{x})$$

- Minimize the Kullback-Leibler (KL) divergence

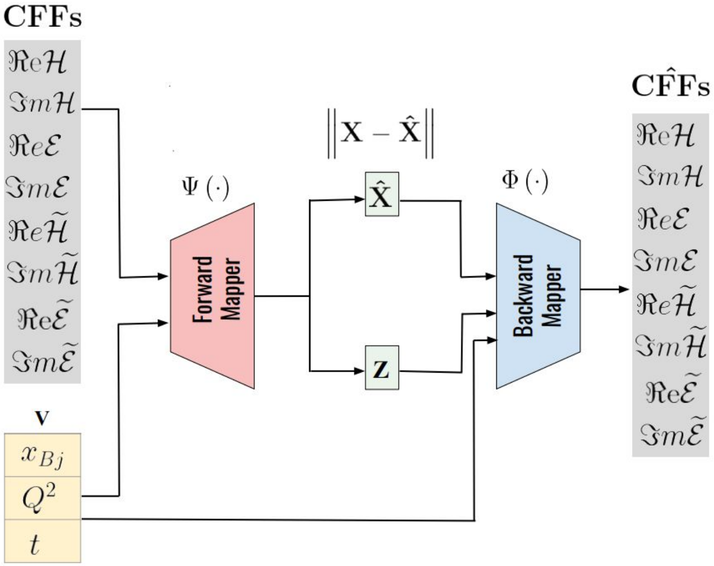
$$\min \left[\mathbf{KL}(q(\mathbf{z} \mid \mathbf{cff}, \mathbf{v}, \mathbf{x}) \parallel p(\mathbf{z} \mid \mathbf{cff}, \mathbf{v}, \mathbf{x}))\right]$$

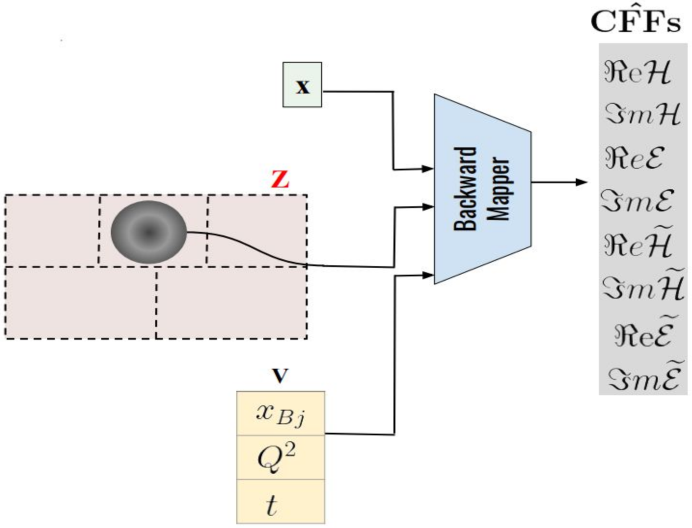
$$\rightarrow \min \left[\|\mathbf{x} - \hat{\mathbf{x}}\|_2^2 + \|\mathbf{cff} - \hat{\mathbf{cff}}\|_2^2 + \mathrm{KL}\left(q(\mathbf{z} \mid \mathbf{cff}, \mathbf{v}, \mathbf{x}) \parallel p(\mathbf{z} | \mathbf{v})\right)\right]$$

# CFFs extraction: Methods

| | $x_bj$ | $t$ | $Q^2$ |
|---|---|---|---|
| 1 | 0.343 | -0.172 | 1.820 |
| 2 | 0.368 | -0.232 | 1.933 |
| 3 | 0.375 | -0.278 | 1.964 |
| 4 | 0.379 | -0.323 | 1.986 |
| 5 | 0.381 | -0.371 | 1.999 |

## B. Conditional-VAIM (C-VAIM)

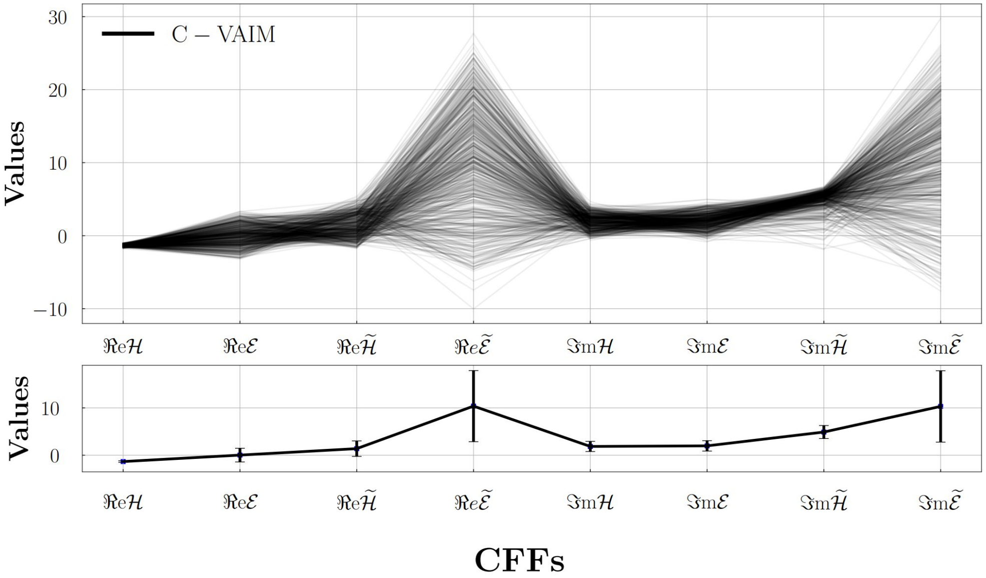


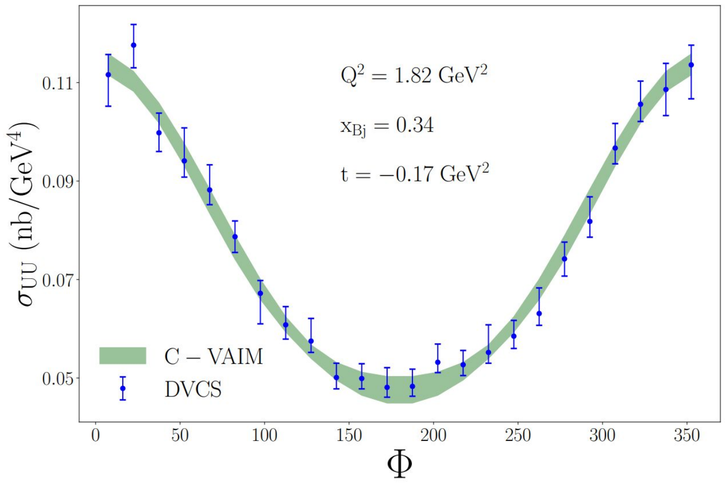
# CFFs extraction: Results

## B. C-VAIM on several kinematics sets

| | $x_{bj}$ | $t$ | $Q^2$ |
|---|---|---|---|
| 1 | 0.343 | -0.172 | 1.820 |
| 2 | 0.368 | -0.232 | 1.933 |
| 3 | 0.375 | -0.278 | 1.964 |
| 4 | 0.379 | -0.323 | 1.986 |
| 5 | 0.381 | -0.371 | 1.999 |




Reconstructed cross sections




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

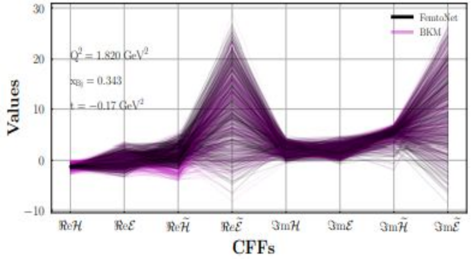
# CFFs extraction: Results

## B. C-VAIM on several kinematics sets

BKM cross sections : the cross sections as they are written in the literature[2].

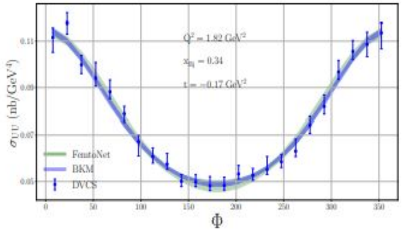
FemtoNet cross sections: UVA[3].

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

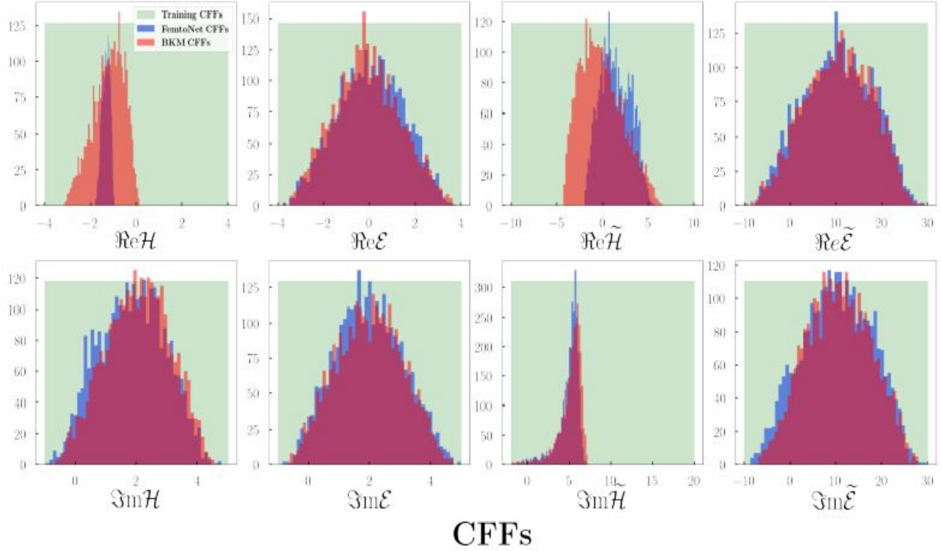



(a)

Reconstructed cross sections




(b)

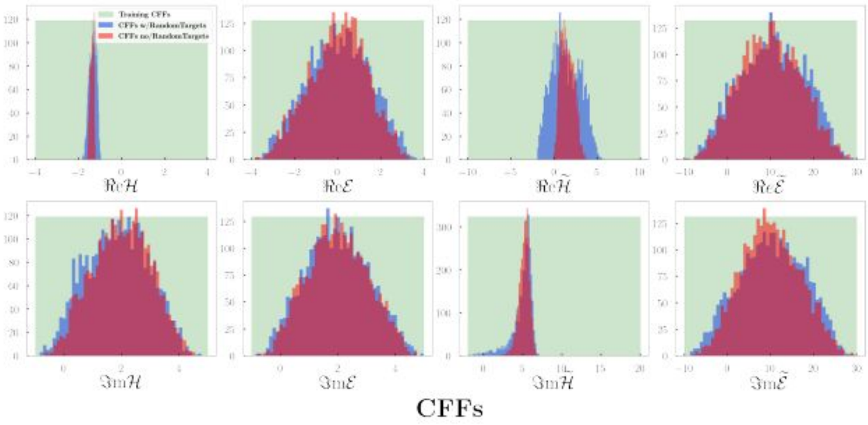


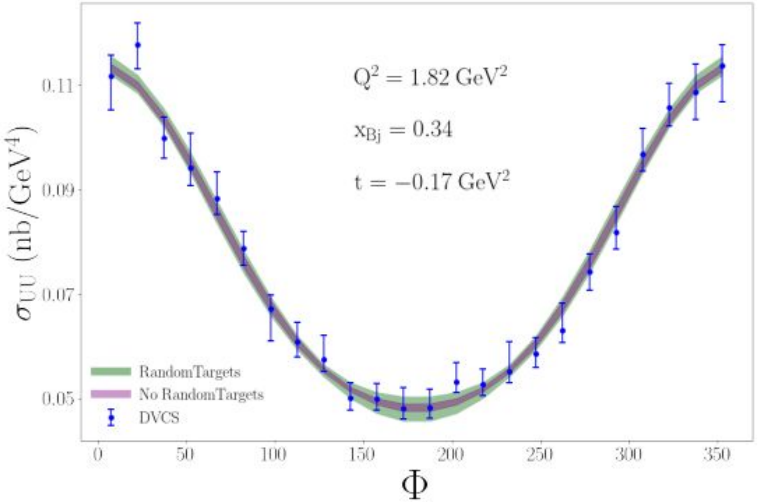

[2] A. V. Belitsky and D. Mueller, "Exclusive electroproduction revisited: treating kinematical effects," Phys. Rev., vol. D82, p. 074010, 2010.
[3 B. Kriesten et al. "Extraction of generalized parton distribution observables from deeply virtual electron proton scattering experiments," Phys. Rev. D, 2020.

# CFFs extraction: Results

## B. C-VAIM on several kinematics sets

The results of our analysis with and without the random targets method for the propagation of experimental uncertainties.




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

# Summary

- **PCNNs**
  - Incorporating Physics constraint such as angular symmetry to DVCS cross sections.
  - Generalization capability on BH and DVCS data.
- **C-VAIM**
  - Extracting the CFFs from cross section.
  - Applied to the DVCS experimental data.

# Publications:

- M. Almaeen, Y. Alanazi, N. Sato, W. Melnitchouk, M. Kunchera, and Y. Li. ”**Variational Autoencoder Inverse Mapper: An End-to-End Deep Learning Framework for Inverse Problems**”. International Joint Conference of Neural Networks IJCNN-2021.

- M. Almaeen, J. Grigsby, J. Hoskins, B. Kriesten, Y. Li, H. Lin, S. Liuti. “Benchmarks for a Global Extraction of Information from Deeply Virtual Exclusive Scattering”, arXiv:2207.10766.

- M. Almaeen, J. Hoskins, B. Kriesten, Y. Li, H. Lin, S. Liuti. “VAIM - CFF: A variational autoencoder inverse mapper solution to Compton form factor extraction from deeply virtual exclusive reactions”, In progress.

# FemtoNet

## Phenomenology

**Simonetta Liuti**
Univ. of Va.

**Brandon Kriesten**
CNF

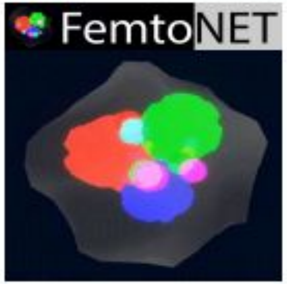
## Lattice QCD

**Huey-Wen Lin**
MSU

## Machine Learning

**Yaohang Li**
ODU

**Joshua Hoskins**
Univ. of Va.




## Students

**Jake Grigsby**
Univ. of Va.

**Manal Almaeen**
ODU

**Zaki Panjsheeri**
Univ. of Va.

**Joshua Bautista**
Univ. of Va.