Deep learning models for deeply virtual exclusive processes

B. Kriesten

CENTER for NUCLEAR FEMTOGRAPHY

FemtoNET

APS Topical Group on Hadronic Physics

April 12, 2023

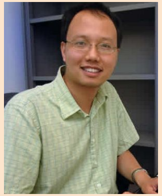
FemtoNet

Phenomenology

Simonetta Liuti
Univ. of Va.

Brandon Kriesten
CNF

Machine Learning

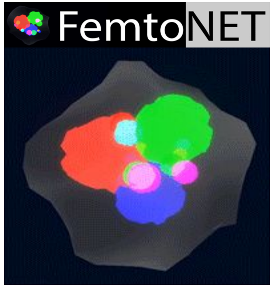


Yaohang Li
ODU

Joshua Hoskins
Univ. of Va.

Lattice QCD

Huey-Wen Lin
MSU

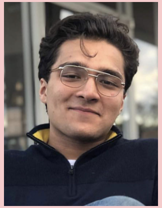


Students

Jake Grigsby
Univ. of Va.

Manal Almaeen
ODU

Zaki Panjsheeri
Univ. of Va.



Joshua Bautista
Univ. of Va.

FemtoNet Publications

Machine Learning

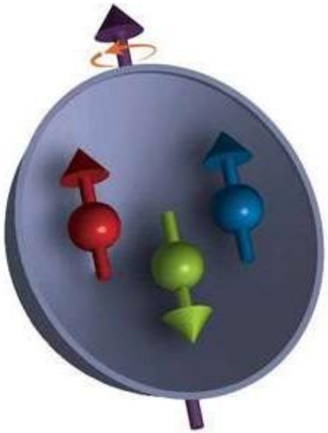
- Deep Learning Analysis of Deeply Virtual Exclusive Photoproduction **PRD 104 (2021)**
- Benchmarks for a Global Extraction of Information from Deeply Virtual Exclusive Scattering Experiments **arXiv:2207.10766**
- VAIM - CFF: A variational autoencoder inverse mapper solution to Compton form factor extraction from deeply virtual exclusive reactions **(in progress)**
- Deep learning partonic angular momentum through VAIM **(in progress)**

Phenomenology

- Extraction of generalized parton distribution observables from deeply virtual electron proton scattering experiments **PRD 101 (2020)**
- Theory of deeply virtual Compton scattering off the unpolarized proton **PRD 105 (2022)**
- Novel Rosenbluth extraction framework for Compton form factors from deeply virtual exclusive experiments **PLB 829 (2022)**
- Parametrization of quark and gluon generalized parton distributions in a dynamical framework **PRD 105 (2022)**
- Deeply virtual Compton scattering from fixed target to collider settings **(in progress)**

Some physics motivation …

Spin as an emergent phenomena of QCD dynamics



The naive parton model cannot explain the origin of hadronic properties such as spin.

Orbital motion (dynamics) of the quarks and gluons could be the answer.

How do we describe <u>orbital angular momentum</u> of partons?

Generalized Parton Distributions

It was shown that the **quantum correlation functions** that can describe the consequences of orbital dynamics of partons in the nucleon are the 3D generalized parton distributions (GPDs).

$$F^{[\Gamma]}_{\Lambda,\Lambda'}(x,\xi,t) = \frac{1}{2}\int \frac{dz^-}{2\pi} e^{ixP^+z^-} \langle p',\Lambda'|\overline{\psi}\Big(-\frac{z}{2}\Big)\Gamma\mathcal{W}\Big(-\frac{z}{2},\frac{z}{2}|n\Big)\psi\Big(\frac{z}{2}\Big)|p,\Lambda\rangle\Big|_{z^+=z_T=0}$$

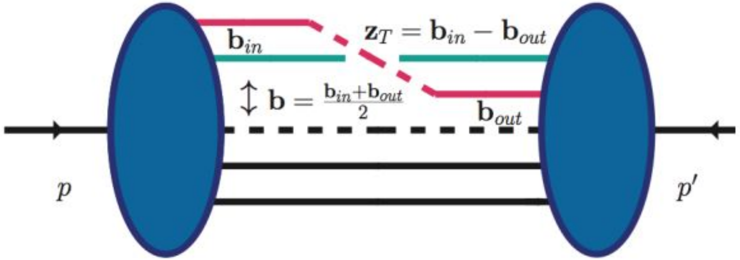


Image credit: A. Rajan, M. Engelhardt, S. Liuti **PRD 98 (2018)**

$$t = (p' - p)^2$$

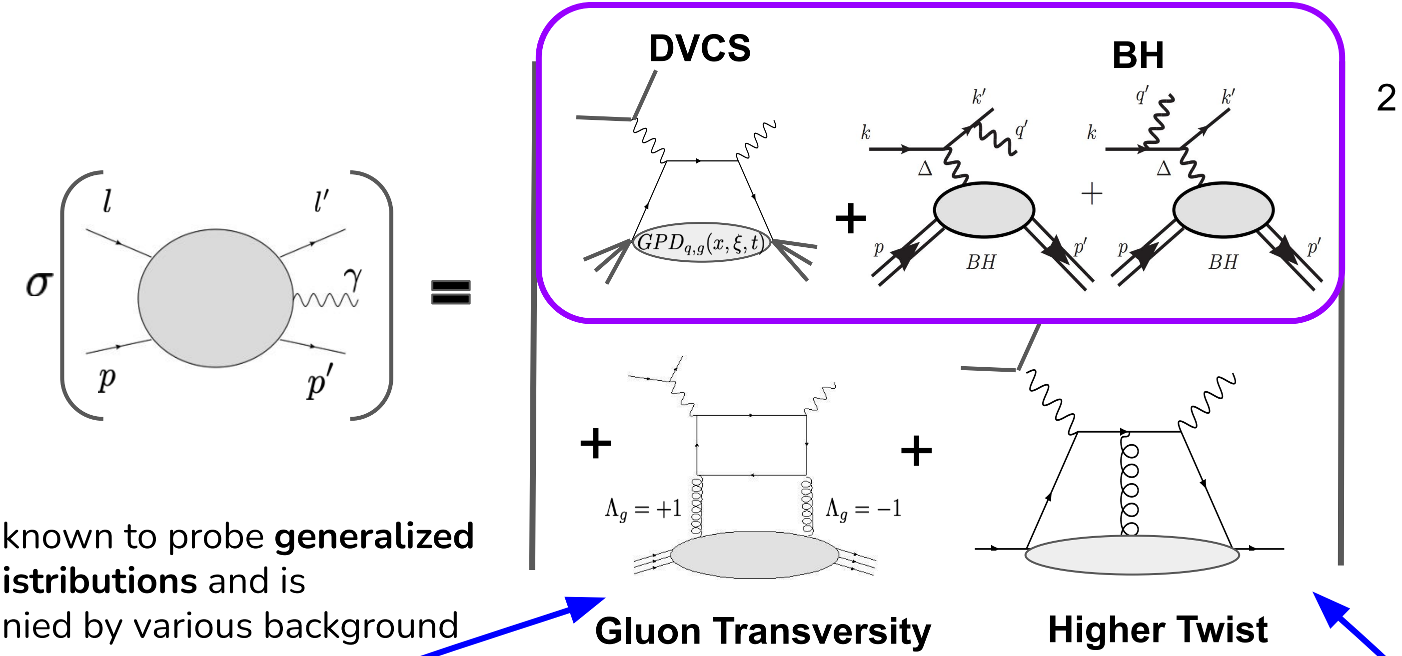
$$\xi = \frac{(p'-p)^+}{(p'+p)^+}$$

New information on parton dynamics!

X. Ji **PRL. 78 (1997)**
A. Radyushkin **PRD. 56 (1997)**

D. Muller, et. al. **(1994)**
M. Diehl **Phys.Rep. (2003)**

How to measure GPDs? Deeply virtual Compton scattering



DVCS is known to probe **generalized parton distributions** and is accompanied by various background processes.

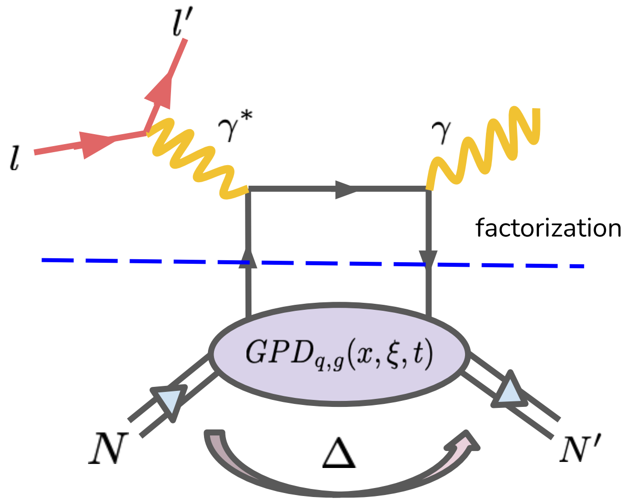
X. Ji, **PRD. 55 (1997)**

B.Kriesten, S.Liuti, et. al. **PRD. 101 (2020)**

Important, but reserved for later ...

However … there's a catch!

In the DVCS cross section, **GPDs** come convoluted with Wilson coefficient functions (Compton Form Factors) meaning we only have experimental access to integrals (ReCFF) or specific points in x (ImCFF) of these distributions.



$$\mathcal{H}^q(\xi,t) = e_q^2 P.V. \int_{-1}^{+1} dx \left[\frac{1}{\xi - x} - \frac{1}{\xi + x}\right] H^q(x,\xi,t) + i\pi e_q^2 H^{q+\bar{q}}(\xi,\xi,t)$$

Not the same integral for angular momentum!

What does the DVCS cross section look like?

$$\sigma = \sigma_{BH} + \sigma_{DVCS} + \sigma_{\mathcal{I}}$$

The cross section has three components that contribute to leading order

$$\sigma_{BH}(x_{Bj}, t, Q^2, E_b, \phi) = \frac{\Gamma}{t}\left[A_{UU}^{BH}(F_1^2 + \tau F_2^2) + B_{UU}^{BH}\tau G_M^2(t)\right]$$

No CFFs

$$\sigma_{\mathcal{I}}(x_{Bj}, t, Q^2, E_b, \phi) = \frac{\Gamma}{Q^2 t}\Big[A_{\mathcal{I}}(x_{Bj}, t, Q^2, E_b, \phi)\Big(F_1(t)\Re e\mathcal{H}(x_{Bj}, t, Q^2) + \tau F_2(t)\Re e\mathcal{E}(x_{Bj}, t, Q^2)\Big)$$
$$+ B_{\mathcal{I}}(x_{Bj}, t, Q^2, E_b, \phi)G_M(t)\Big(\Re e\mathcal{H}(x_{Bj}, t, Q^2) + \Re e\mathcal{E}(x_{Bj}, t, Q^2)\Big)$$
$$+ C_{\mathcal{I}}(x_{Bj}, t, Q^2, E_b, \phi)G_M(t)\Re e\widetilde{\mathcal{H}}(x_{Bj}, t, Q^2)\Big]$$

Linear CFFs: 3

$$\sigma_{DVCS}(x_{Bj}, t, Q^2, E_b, \phi) = \frac{\Gamma}{Q^2}\frac{2}{1-\epsilon}\Big[(1-\xi^2)\Big[(\Re e\mathcal{H})^2 + (\Im m\mathcal{H})^2 + (\Re e\widetilde{\mathcal{H}})^2 + (\Im m\widetilde{\mathcal{H}})^2\Big]$$
$$+ \frac{t_o - t}{4M^2}\Big[(\Re e\mathcal{E})^2 + (\Im m\mathcal{E})^2 + \xi^2(\Re e\widetilde{\mathcal{E}})^2 + \xi^2(\Im m\widetilde{\mathcal{E}})^2\Big]$$
$$- 2\xi^2\Big(\Re e\mathcal{H}\,\Re e\mathcal{E} + \Im m\mathcal{H}\Im m\mathcal{E} + \Re e\widetilde{\mathcal{H}}\,\Re e\widetilde{\mathcal{E}} + \Im m\widetilde{\mathcal{H}}\Im m\widetilde{\mathcal{E}}\Big)\Big]$$

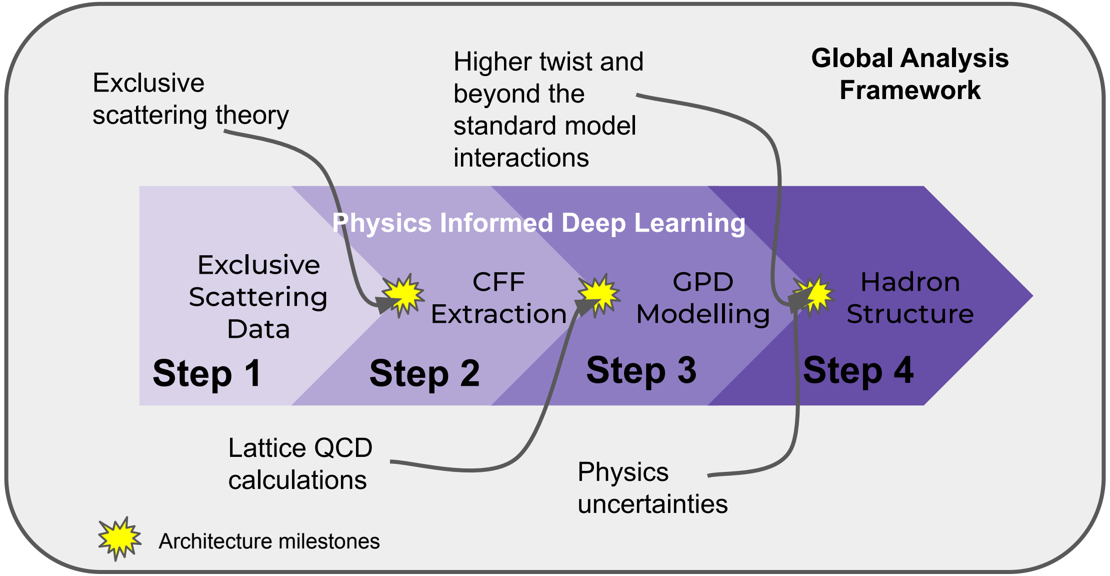
Quadratic CFFs: 8

All 8 CFFs enter into a single polarization observable!

A Physics Informed Deep Learning Framework …

GPD extraction is a really difficult problem!

There are many levels of abstraction going from experimental cross sections to calculating the physical properties of the hadron.



Strategic applications of **ML techniques** in four phases as a framework to pass from cross section data to the physical properties of interest.

Physics Informed Deep Learning Models

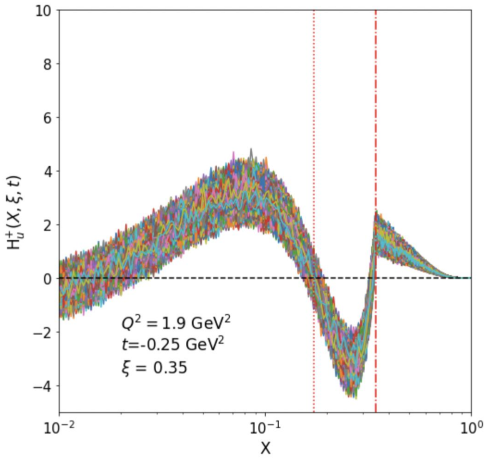
DNN models can spend a lot of computational resources to learn physical laws from data. To **reduce computation time** and **improve network performance/generalization**, we can incorporate those laws into the network so that certain physical properties are learned/inherently satisfied in the network's predictions.

Physics Constraints

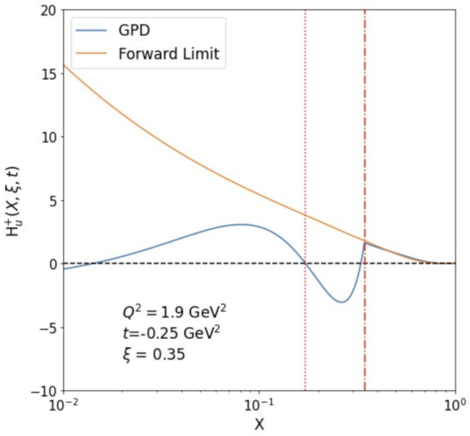
- Cross section structure built into the loss function
- Experimental error bars
- Lorentz invariance - polynomiality property
- Positivity constraints
- Forward limit constraints of GPDs
- Dispersion relations with threshold effects
- Evolution constraints

Physics constraints (ex. positivity) may **look different at each step of the analysis**.

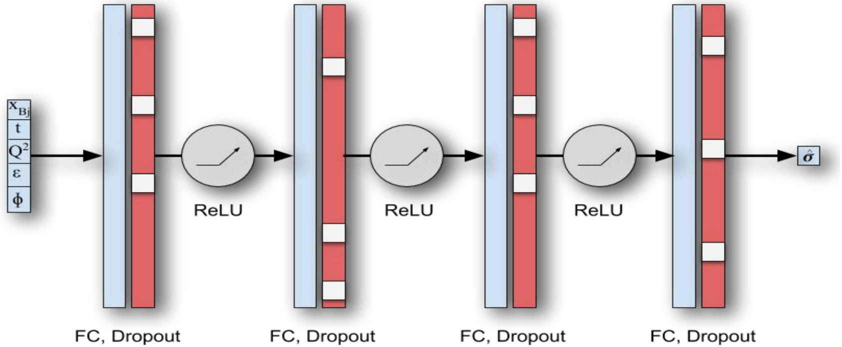
Step 2

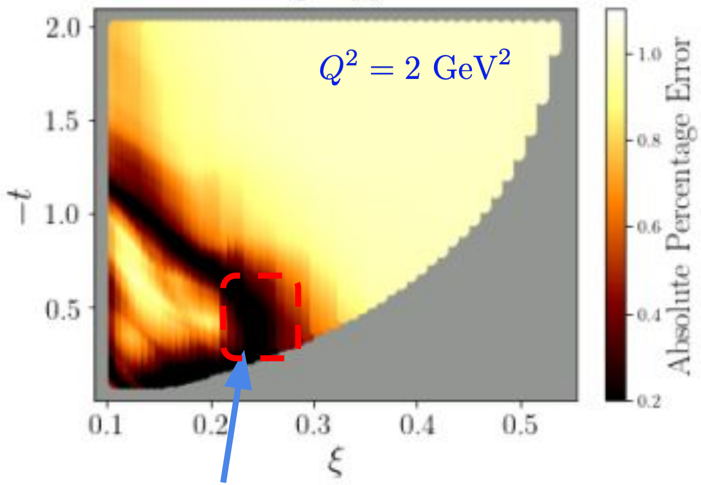


Step 3



Phase 1. Deep Learning DVCS Data





Current Data

Experiment	Q^2 (GeV^2)	$-t$ (GeV^2)	x_{Bj}	E_{beam} (GeV)	# points
Hall B	1.11 - 3.77	0.11 - 0.45	0.13 - 0.48	5.75	1931
Hall A	1.45 - 2.38	0.17 - 0.37	0.34 - 0.40	5.75	468
Hall A	1.51 - 2.00	0.18 - 0.36	0.36	3.36 - 5.55	383
Hall A	2.71 - 8.51	0.21 - 1.28	0.34 - 0.61	4.49 - 10.99	1080

Why do we need a deep neural network?

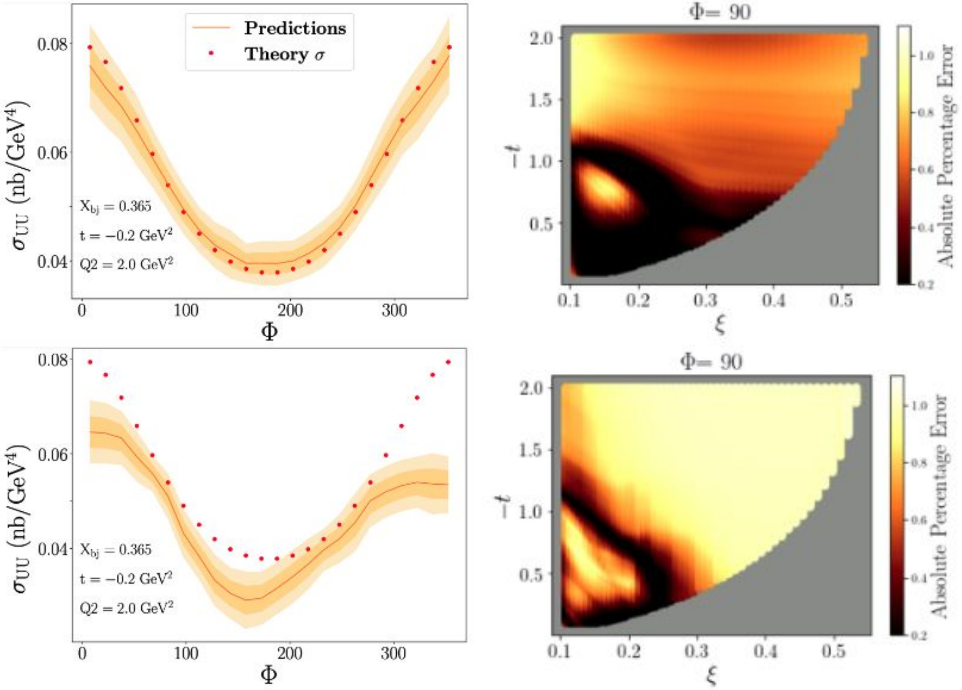
- DNN provide **efficient** and **accurate** predictions of the cross section while squeezing as much information from data as possible.

J. Grigsby, BK, S. Liuti, et. al. **PRD 104 (2021)**
M. Almaeen, J. Grigsby, J. Hoskins, BK, Y. Li, H-W. Lin, S. Liuti **arXiv:2207.10766**

Physics constrained cross section predictions

Simple physics constraints such as symmetry properties of the unpolarized cross section in the loss function lead to increased generalization of the DNN predictions.

$$||f(x_{Bj}, t, Q^2, \phi, \epsilon) - f(x_{Bj}, t, Q^2, -\phi, \epsilon)||$$



J. Grigsby, BK, S. Liuti, et. al. **PRD 104 (2021)**
M. Almaeen, J. Grigsby, J. Hoskins, BK, Y. Li, H-W. Lin, S. Liuti **arXiv:2207.10766**

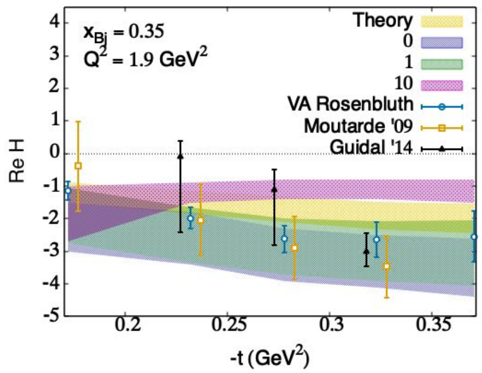
Phase 2a. Defining Benchmarks for Information Extraction

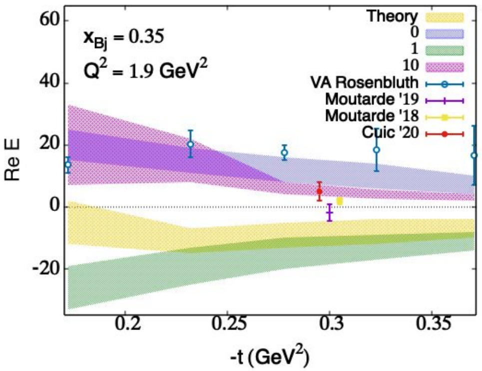
The idea of **benchmarks** is to establish *transparency* in the exact methodology of fitting, for *reproducibility* of results, and for *compatibility* of extracted quantities.

Deeply virtual exclusive processes are a class of process of their own, and therefore should be treated as such from the initial stages of their analysis. They require their own benchmarks!

There are 2 sets of benchmarks that one can establish:

Physics benchmarks - what physics is entering your fit?
ML / fitting benchmarks - what ML tools are you using, hyperparameters, and how are you validating your results?





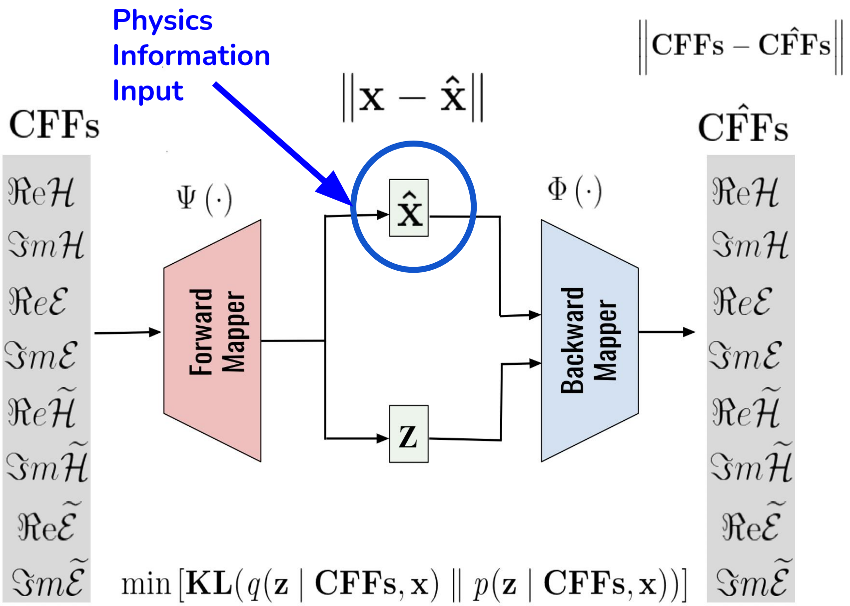
M. Almaeen, J. Grigsby, J. Hoskins, BK, Y. Li, H-W. Lin, S. Liuti **arXiv:2207.10766**

Phase 2b. VAIM-CFF: A variational autoencoder framework or Reframing the Extraction of Compton Form Factors

FemtoNet

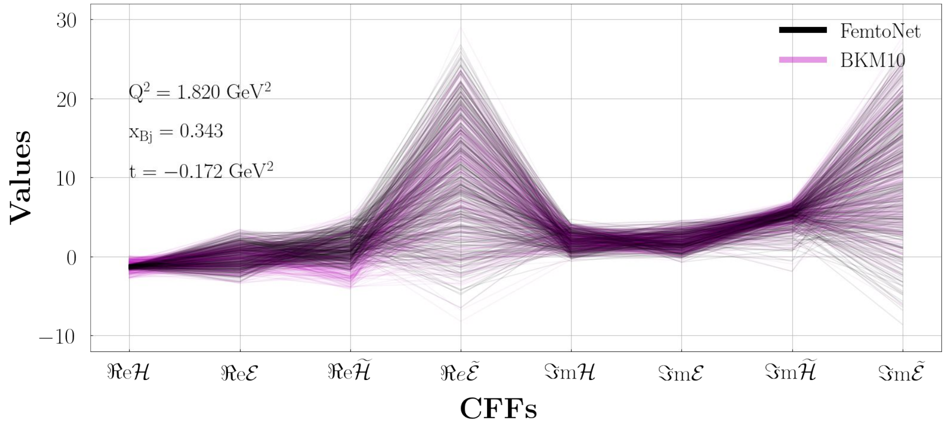
Extraction of 8 CFFs from a single polarization observable treated as an "inverse problem" of extracting 8 unknowns from a single equation.

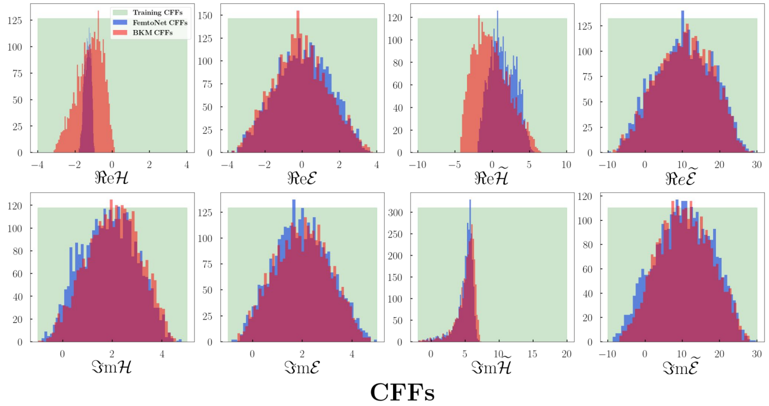
Quantification of **information that is possible to extract** from certain experiments. What is possible to be learned from current data?

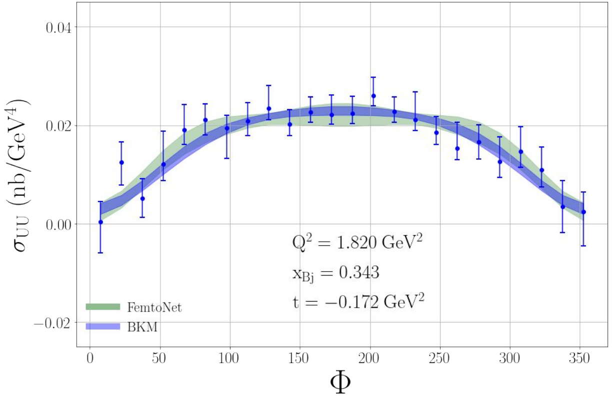


M. Almaeen, J. Hoskins, BK, Y. Li, H-W. Lin, S. Liuti (**in progress**)

VAIM-CFF results





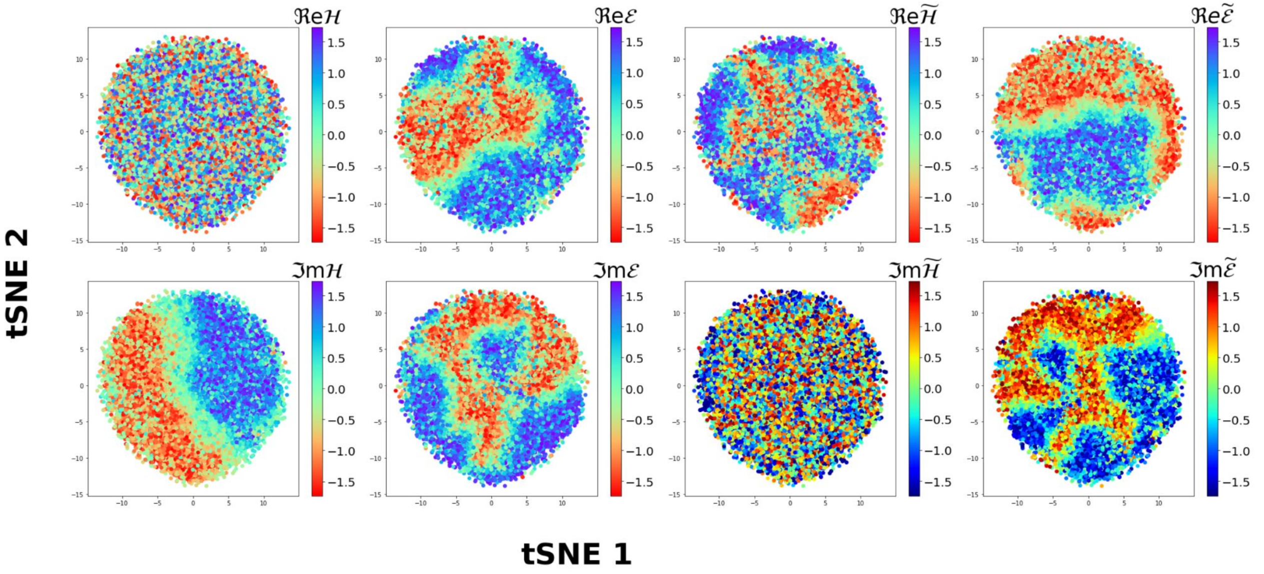


The solution set for some of the CFFs seems to be bounded (ex. ReH) while others are not bounded (ex. ReE tilde). Nevertheless there seems to exist a singular solution with an error bar for each CFF.

Ex. ReH is interesting, given the choice of (+) or (-) it seems to determine that it is always (-).

M. Almaeen, J. Grigsby, J. Hoskins, BK, Y. Li, H-W. Lin, S. Liuti **(in progress)**

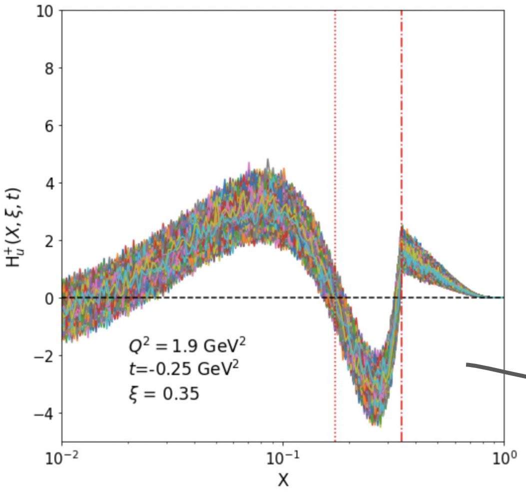
VAIM-CFF results

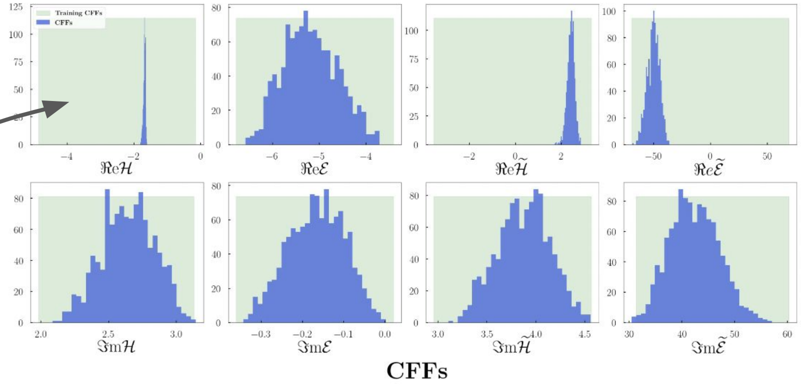


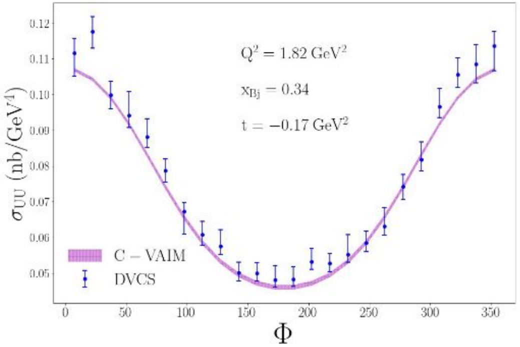
Dimensionality reduction techniques such as tSNE reveal structures that indicate the latent space is capturing some information that is lost in the encoding process.

M. Almaeen, J. Grigsby, J. Hoskins, BK, Y. Li, H-W. Lin, S. Liuti **(in progress)**

Physics informed VAIM-CFF results







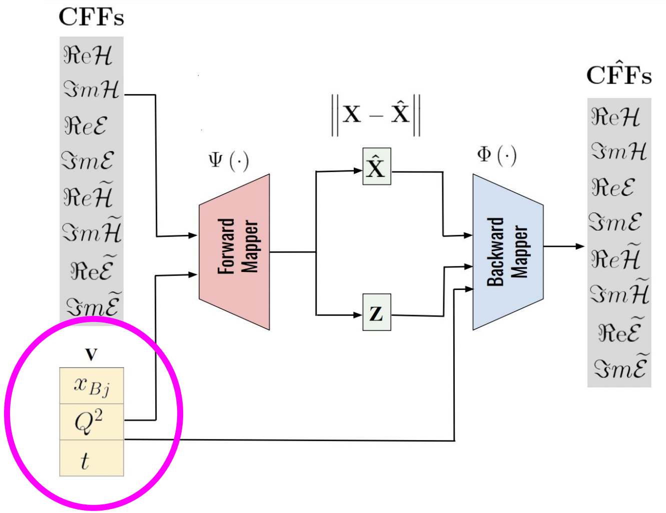
How can we include physics information?

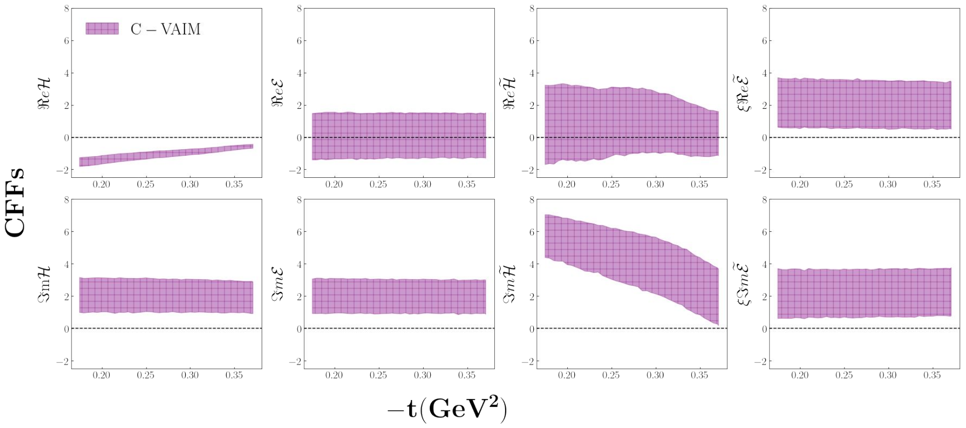
Construct envelopes of GPDs created from oscillating model parameters and adding in noise while ensuring physics constraints are still compatible.

M. Almaeen, J. Grigsby, J. Hoskins, BK, Y. Li, H-W. Lin, S. Liuti **(in progress)**

C-VAIM: Studying CFF trends

Conditional VAIM (C-VAIM) allows us to study trends in DVCS data by placing a kinematic condition on the training of the forward/backward mapper so during training it can learn on a conglomerate of DVCS data.





M. Almaeen, J. Grigsby, J. Hoskins, BK, Y. Li, H-W. Lin, S. Liuti **(in progress)**

Uncertainty Quantification

Uncertainty arises in many places when using ML algorithms, it is critical to make sure we understand how much we can trust the algorithms predictions. Four factors vital for understanding uncertainty are:

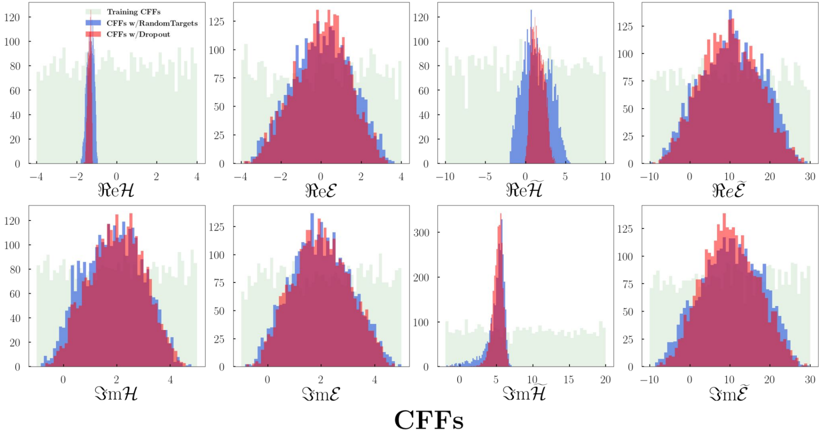
1. Statistical uncertainty from experimental measurements
2. Systematic uncertainties from physics measurements

Irreducible

3. Error in the ML model and its architecture
4. Errors in training procedures

Reducible

We have to make sure we are properly propagating irreducible errors through our DNN architectures and that we understand the size of our network errors. We can randomly sample the error bars of the data to create **random targets** for our neural network to train on.

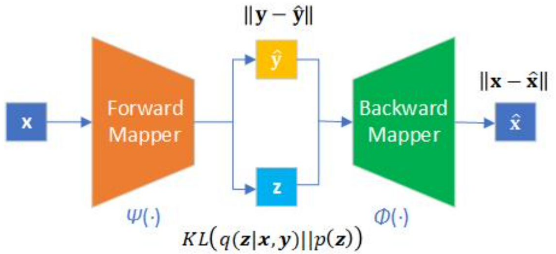


M. Almaeen, J. Hoskins, BK, Y. Li, H-W. Lin, S. Liuti (**in progress**)

What’s next …

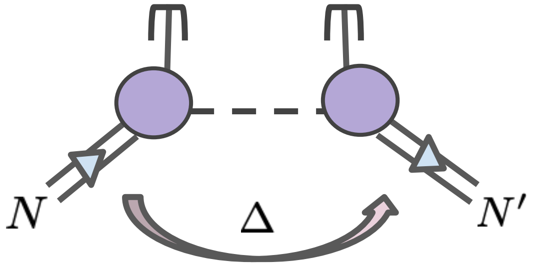
Upcoming: Phase 3a. VAIM for GPDs, searching for signatures of AM

Using VAIM, can we determine all possible model parameters for a solution set of GPDs that can be fit to theory constraints, lattice QCD calculations, and experimental data?



What are the various **outcomes/signatures of angular momentum** allowed by the data?

Parameterization developed theoretically in a spectator model



$$H(x,\xi,t)^i = \mathcal{N}_i x^{-\alpha_i} x^{-\alpha_i'(1-x)^{p_i} t} H^{M_{\Lambda,i}}_{M_{X,i},m_i}(x,\xi,\Delta_T)$$

8 parameters per GPD for q^i_v, q^i_s, g and an initial scale for pQCD evolution.

B.K, P. Velie, E. Yeats, F. Yepez-Lopez, S. Liuti **PRD 105 (2022)**

M. Almaeen, J. Grigsby, J. Hoskins, BK, Y. Li, H-W. Lin, S. Liuti **(in progress)**

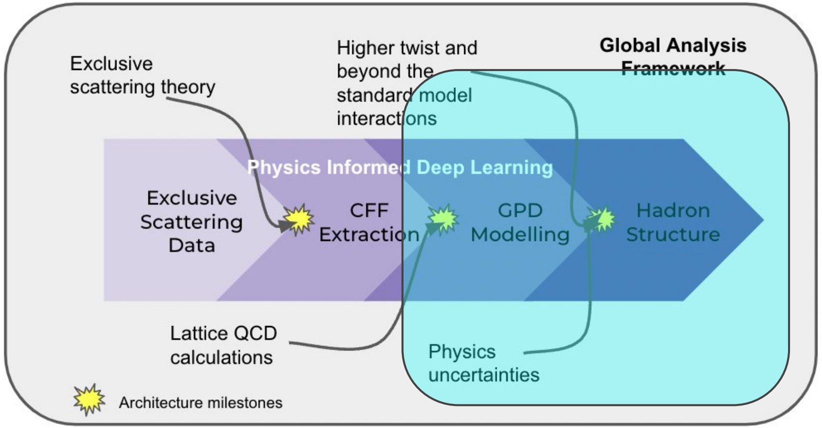
Conclusions

The extraction of the full x-dependence of GPDs requires more than just DVCS data alone.

- Lattice QCD calculation of moments
- Experimental measurements of elastic form factors
- Theoretical GPD properties (polynomiality, positivity, symmetries, forward limits)
- DVES data from multi-channel global analysis

This complicated reconstruction of the information from all the information we have on GPDs requires **new and innovative ML techniques**.

A suite of uncertainty quantification techniques must be applied to determine whether the physics of interest are contained in the networks predictions.



May require more sophisticated algorithms and model development beyond standardized techniques.

Thank you for your attention!