Machine Learning for Inverse Problems in Femtography

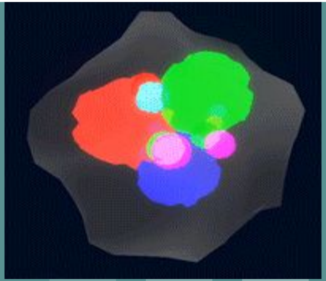
B. Kriesten

Argonne NATIONAL LABORATORY

Brookhaven National Laboratory

U.S. DEPARTMENT OF ENERGY | Office of Science

EXCLAIM



Workshop on Generalized Parton Distributions for Nucleon Femtography in the EIC era
Brookhaven National Laboratory
January 17, 2024

Outline

- Physics Motivation
- DNN Framework
- Uncertainty Quantification
- Conclusions

Outline

- Physics Motivation
- DNN Framework
- Uncertainty Quantification
- Conclusions

Femtography of the nucleon

Femtography - is **data driven visualizations** of the phase space (momentum and spatial) distribution of the quarks and gluons inside of the proton using a variety of **deeply virtual exclusive processes**.

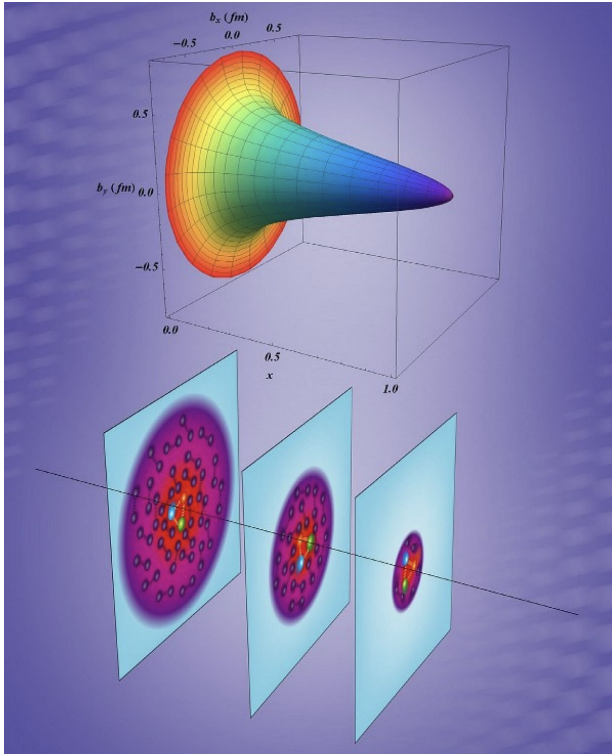
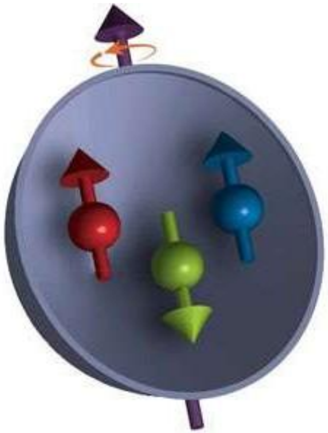


Image credit: **Rafael Dupre**

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** (phase space distributions) that can describe the consequences of orbital dynamics of partons in the nucleon are the 3D generalized parton distributions (GPDs).

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

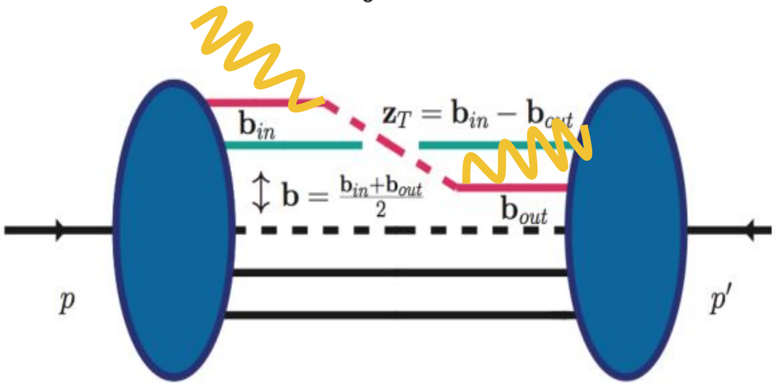


Image credit: A. Rajan, M. Engelhardt, S. Liuti **PRD 98 (2018)**

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

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

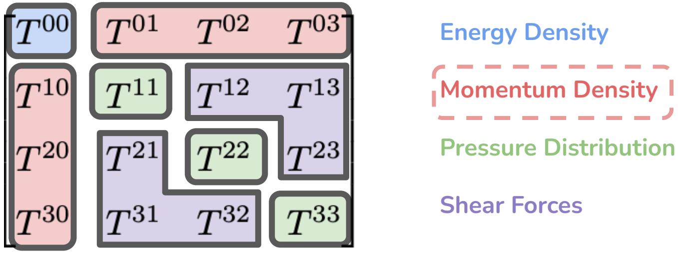
New information on parton dynamics!

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

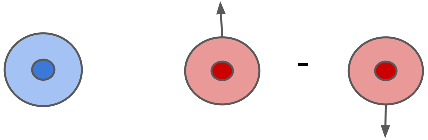
D. Muller, et. al. **(1994)**
M. Diehl **Phys.Rep. (2003)**

What are the possibilities?

$$T^{\mu\nu}_{QCD} = \frac{1}{4}\,\overline{\psi}\gamma^{(\mu}D^{\nu)}\psi + Tr\left\{F^{\mu\alpha}F^{\nu}_{\alpha} - \frac{1}{2}g^{\mu\nu}F^2\right\}$$



There is a connection between GPDs and the EMT of QCD, meaning GPDs can describe the **total angular momentum.**



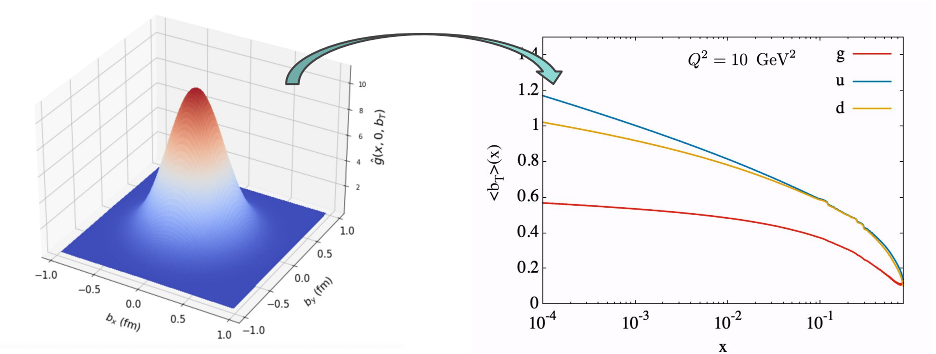
$$J^{q,g} = \frac{1}{2}\int_{-1}^{+1} dx\,x\left(H^{q,g}(x,0,0) + E^{q,g}(x,0,0)\right)$$

X. Ji, W. Melnitchouk, X. Song **PRD 56 (1997)**
X. Ji, **PRD. 55 (1997)**

Spatial Densities

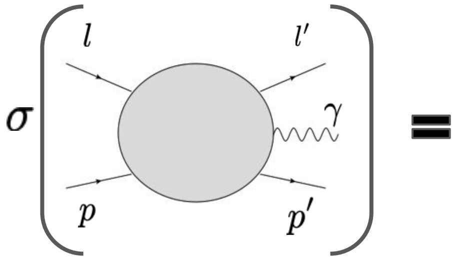
$$\rho^q_{\Lambda\lambda}(\mathbf{b}) = H_q(\mathbf{b}^2) + \frac{b^i}{M}\epsilon_{ij}S^j_T\frac{\partial}{\partial b}E_q(\mathbf{b}^2) + \Lambda\lambda\widetilde{H}_q(\mathbf{b}^2)$$

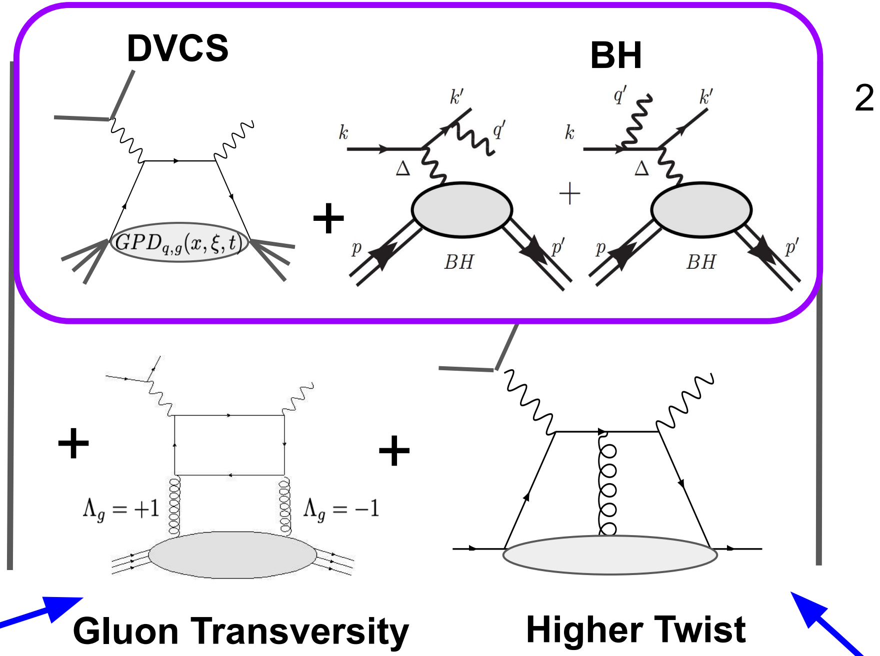
Through Fourier transform of the momentum transfer, we have access to the **spatial distribution** of the partons in the hadron.



M. Burkardt **PRD. 62 (2000)**
M. Burkardt ***Int.J.Mod.Phys.A* 18 (2003)**

How to measure GPDs? Deeply virtual Compton scattering





DVCS is parameterized by **generalized parton distributions** and is also accompanied by various background processes.

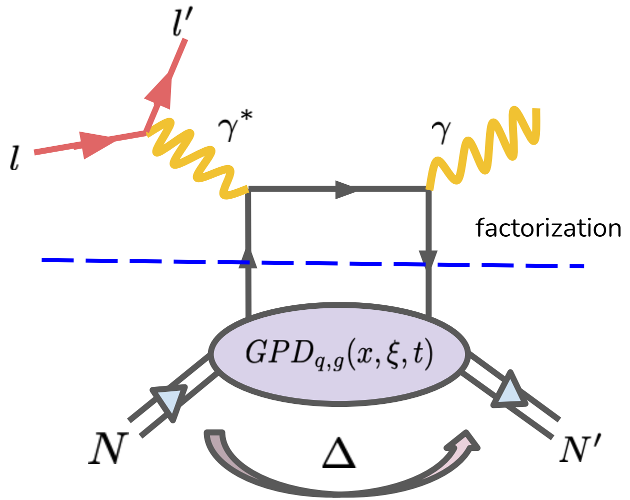
X. Ji, **PRD. 55 (1997)**

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

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

Linear CFFs: 3

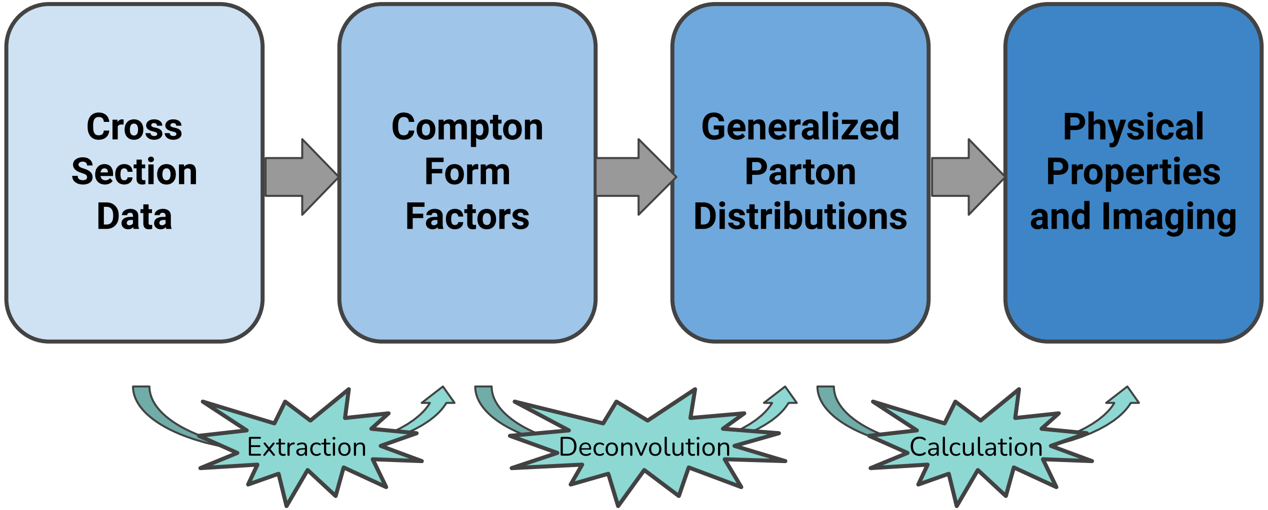
$$\begin{aligned}\sigma_{DVCS}(x_{Bj}, t, Q^2, E_b, \phi) = \frac{\Gamma}{Q^2}\frac{2}{1-\epsilon}\Big[&(1-\xi^2)\Big[(\Re e\mathcal{H})^2 + (\Im m\mathcal{H})^2 + (\Re e\widetilde{\mathcal{H}})^2 + (\Im m\widetilde{\mathcal{H}})^2\Big] \\ &+ \frac{t_o - t}{4M^2}\Big[(\Re e\mathcal{E})^2 + (\Im m\mathcal{E})^2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

Quadratic CFFs: 8

Very complicated to disentangle all of the various pieces!

So … it's a really difficult problem!

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



I believe we can exploit machine learning techniques to give **more information than a model**.

There are many more constraints on GPDs

Physics Constraints

- Polynomiality property (lattice QCD data)
- Positivity constraints
- Forward limit constraints of GPDs from PDFs
- Dispersion relations with threshold effects between CFFs
- Evolution constraints at the perturbative scale

Can we tell the difference in the data?

ϕ (deg)	$x_B = 0.343$, $Q^2 = 1.820$ GeV^2, t = -0.172 GeV^2
7.5	$111.6 \pm 4.1 \, {}^{+2.3}_{-0.0}$
22.5	$117.6 \pm 3.9 \, {}^{+0.7}_{-0.3}$
37.5	$99.8 \pm 3.5 \, {}^{+0.3}_{-0.5}$
52.5	$94.1 \pm 3.3 \, {}^{+0.0}_{-3.4}$
67.5	$88.2 \pm 3.0 \, {}^{+0.0}_{-2.1}$
82.5	$78.7 \pm 2.8 \, {}^{+0.4}_{-0.4}$
97.5	$67.2 \pm 2.6 \, {}^{+3.6}_{-0.0}$
112.5	$60.8 \pm 2.4 \, {}^{+0.5}_{-1.3}$
127.5	$57.5 \pm 2.3 \, {}^{+0.0}_{-2.3}$
142.5	$50.1 \pm 2.1 \, {}^{+0.2}_{-0.8}$

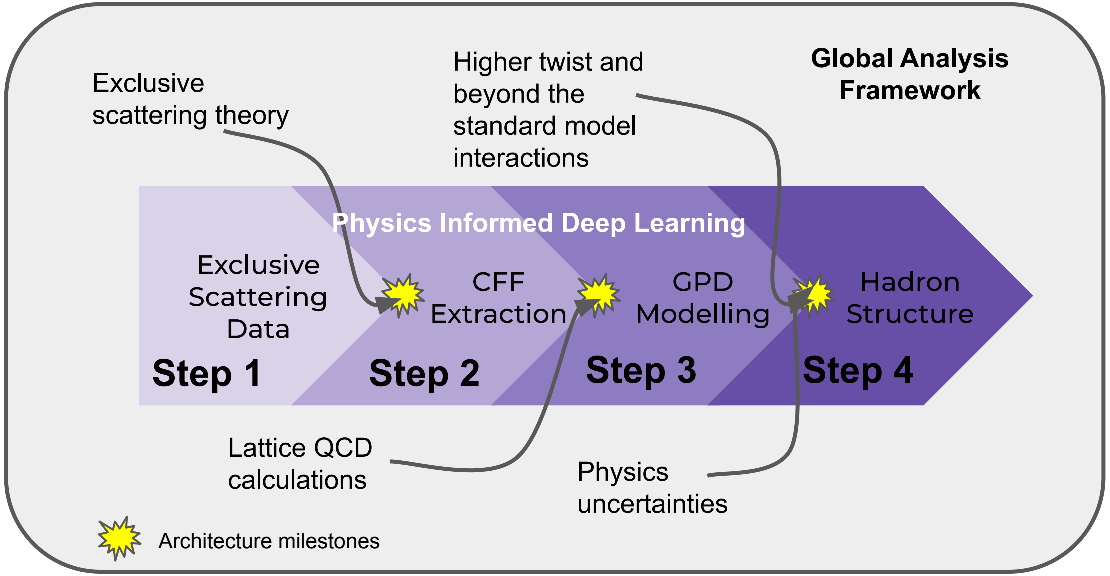
M. Defurne et al., **Phys. Rev. C92, 055202 (2015)**

So it's MUCH more than just a deconvolution problem! **How much information of the GPD is retained in the cross section measurement?**

Outline

- Physics Motivation
- DNN Framework
- Uncertainty Quantification
- Conclusions

Machine Learning as a Femtography Framework



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

Physics Constrained 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 architecture of the network so that certain physical properties are inherently satisfied in the network's predictions.

Hard v. Soft Constraints

Hard constraints are built into the architecture of the neural network itself.

- Always satisfied in network predictions
- Reduces the need for large quantities of data
- Improves generalization and possible extrapolation
- Are difficult to train and optimize
- Architectures can be difficult to develop
- Difficult to interpret how the condition enters

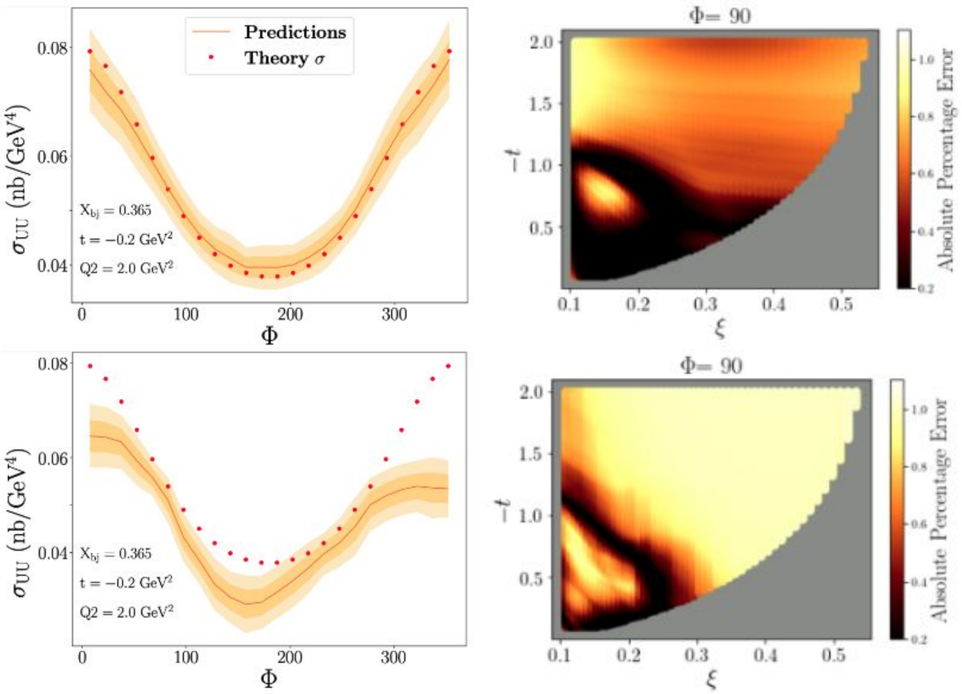
Soft constraints are built into the loss function and are optimized, but never really completely satisfied.

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.

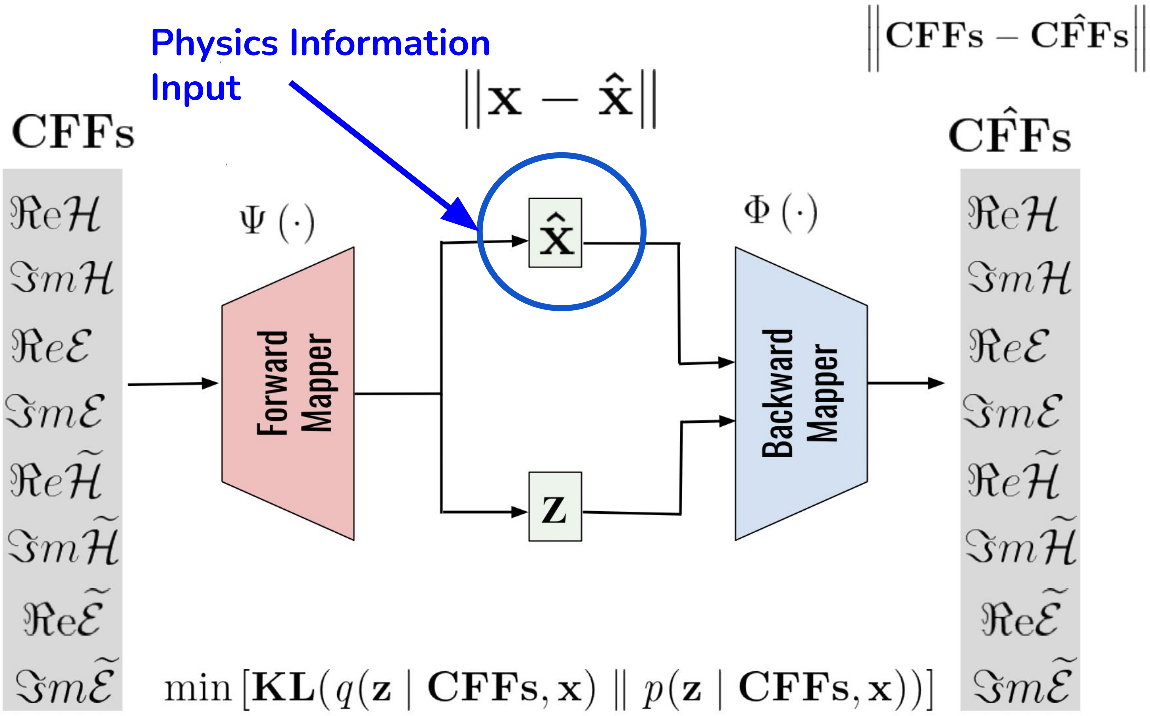
$$||f(x_{Bj}, t, Q^2, \phi, \epsilon) - f(x_{Bj}, t, Q^2, -\phi, \epsilon)||$$

SOFT!



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**

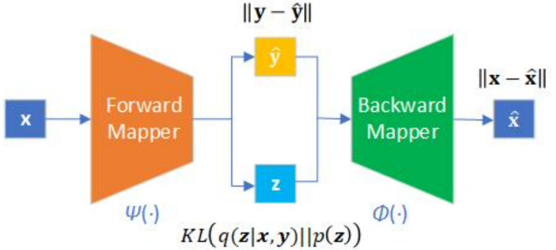
VAIM-CFF: A variational autoencoder framework



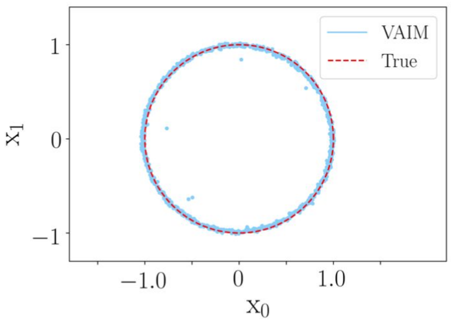
M. Almaeen, J. Grigsby, J. Hoskins, BK, Y. Li, H-W. Lin, S. Liuti **arXiv:2207.10766**

VAIM: a continuous toy problem

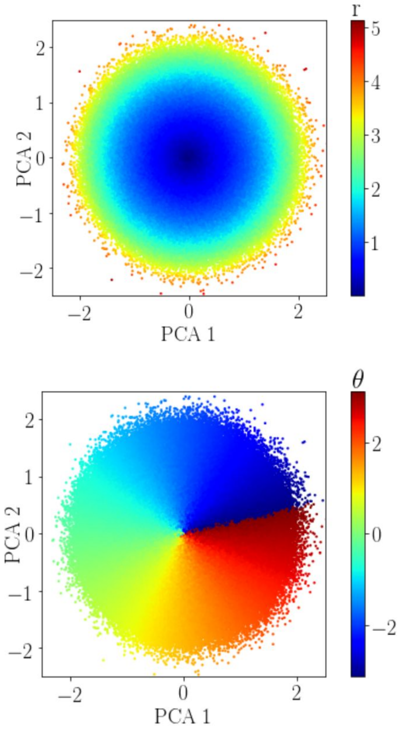
$$f(x_0, x_1) = x_0^2 + x_1^2 \quad x_0, x_1 \in [-2, 2]$$



Prediction



Latent Space

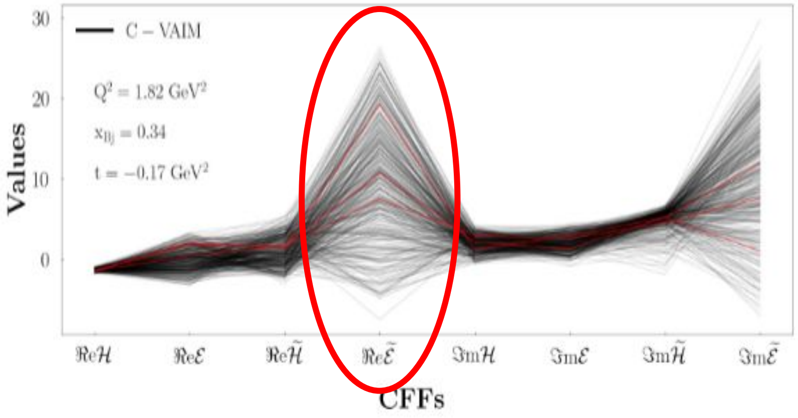


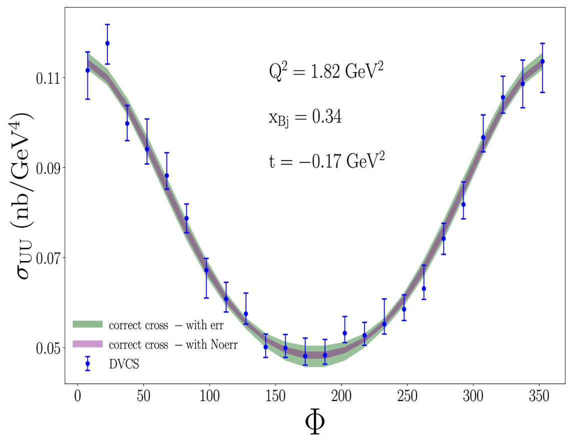
The VAIM framework is able to reconstruct the parameter space of x_0, x_1 such that f = 1.

Notice that a PCA analysis of the first two principal components shows that the lost information is the radial dependence and the azimuthal dependence.

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

VAIM-CFF results

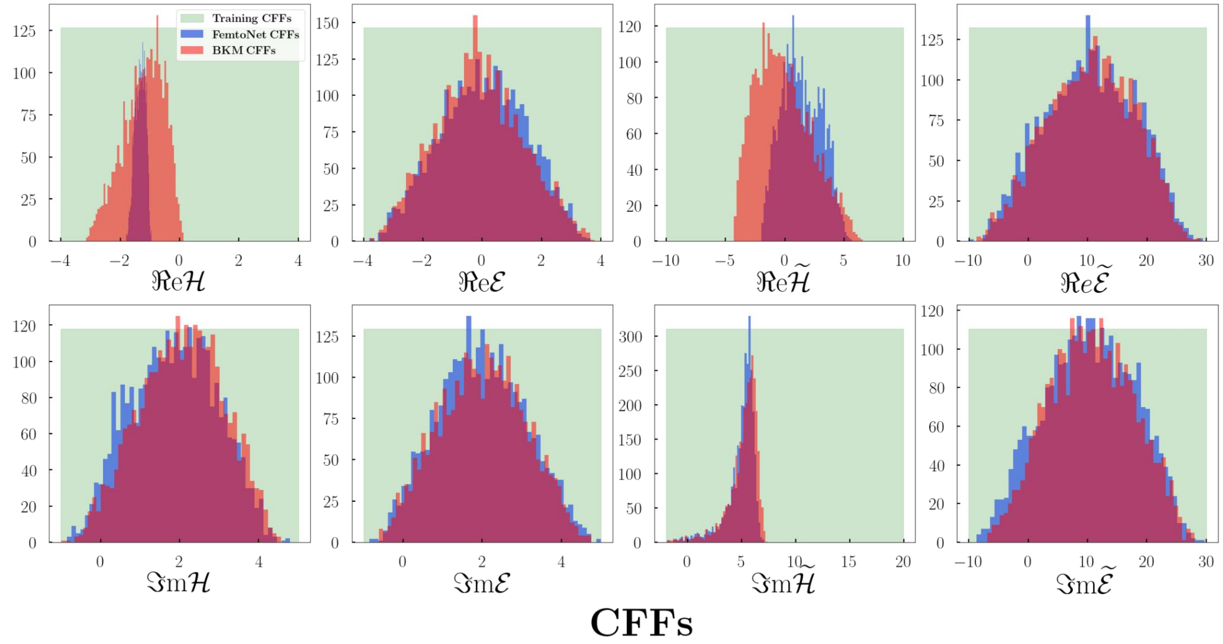




Some unconstrained CFFs, but notice that the large uncertainty in the CFF does not result in a large uncertainty in the cross section.

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

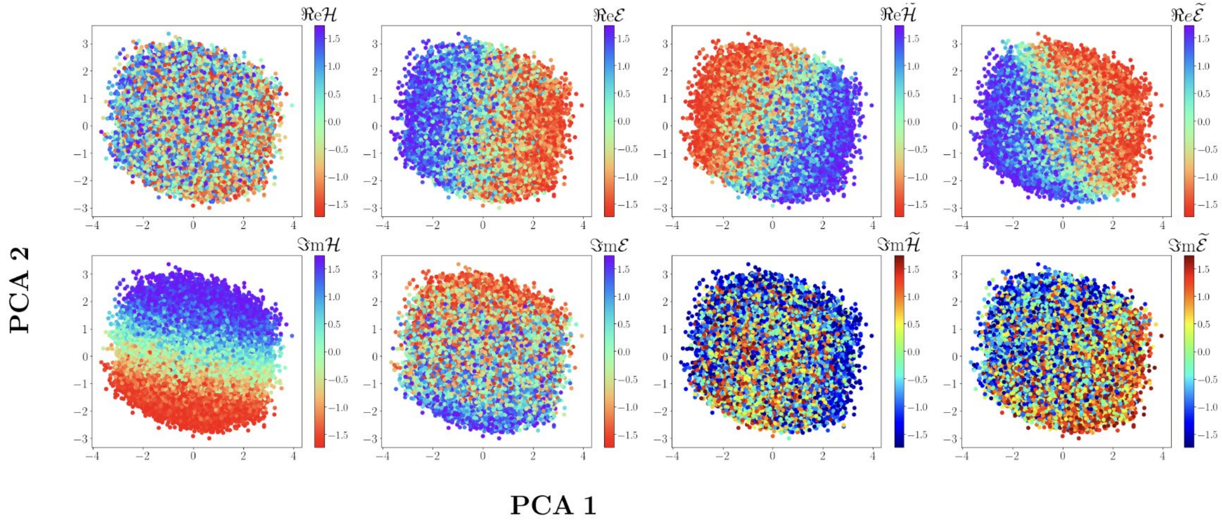
VAIM-CFF results



There exist multiple solutions to this problem (expected) however, the solution set to some of the CFFs seems to be bounded.

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

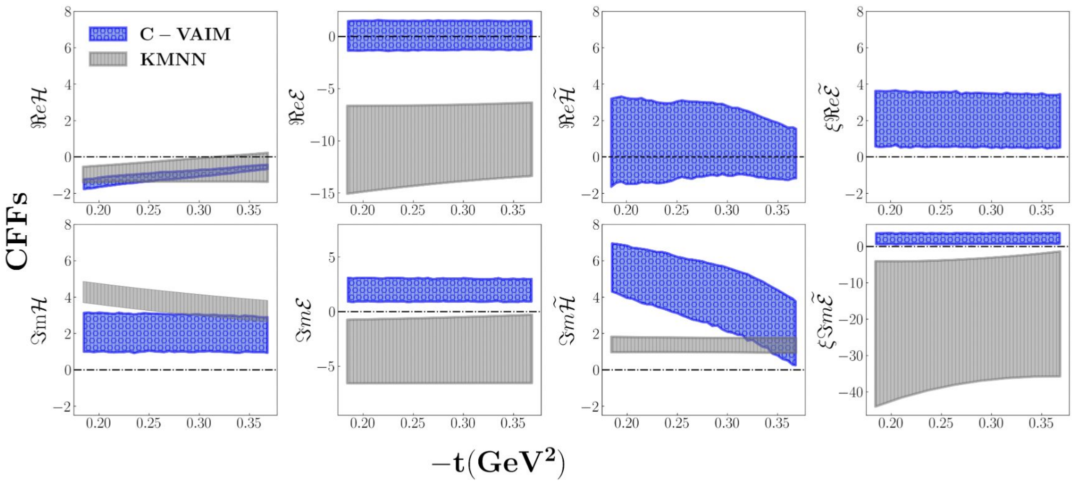
Interpretability of the Latent Distribution



It is still ongoing to study the interpretability of the latent space distribution, what is truly being learned by these models?

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

Conditional VAIM

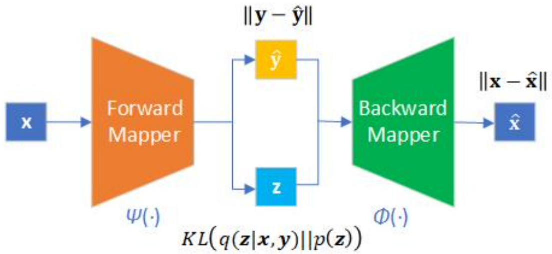


We can take out VAIM and predict across many kinematic variables, comparing with other ML-based extractions.

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

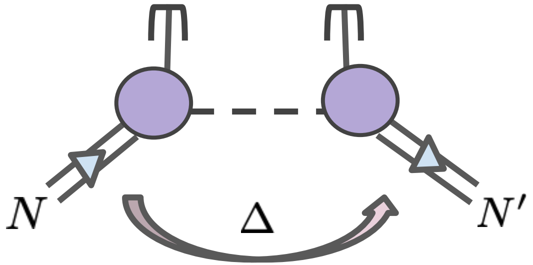
<u>Upcoming:</u> VAIM for GPDs, searching for signatures of AM

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)**

Outline

- Physics Motivation
- DNN Framework
- **Uncertainty Quantification**
- Conclusions

Uncertainty Quantification and DVCS Error Analysis

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 (aleatoric)

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

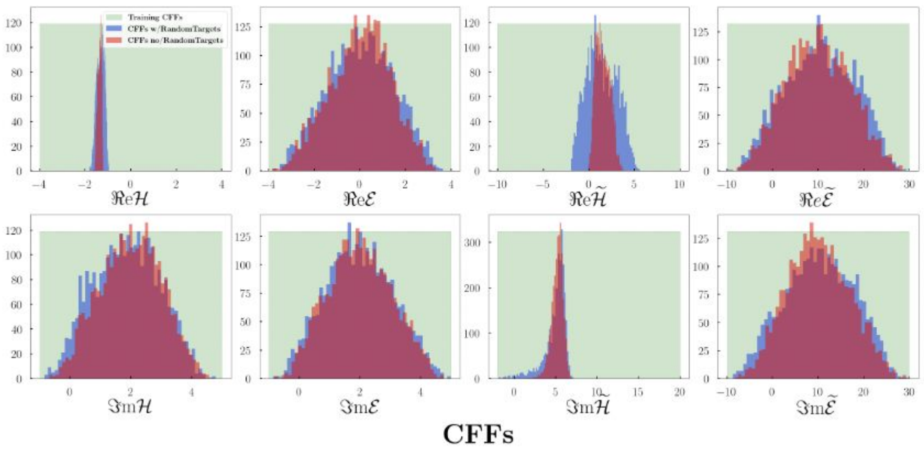
Reducible (epistemic)

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 do this through a method called **random targets**.

Random Targets Method

Dropout approximates a large ensemble of number of different networks.

The **random targets** resample the dataset within the error bars before they train each member of the ensemble of networks.



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

Outline

- Physics Motivation
- DNN Framework
- Uncertainty Quantification
- **Conclusions**

Conclusions

It has become extraordinarily apparent that it is impossible to extract the full x-dependence of GPDs from only 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 GPD from all the information we have 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.