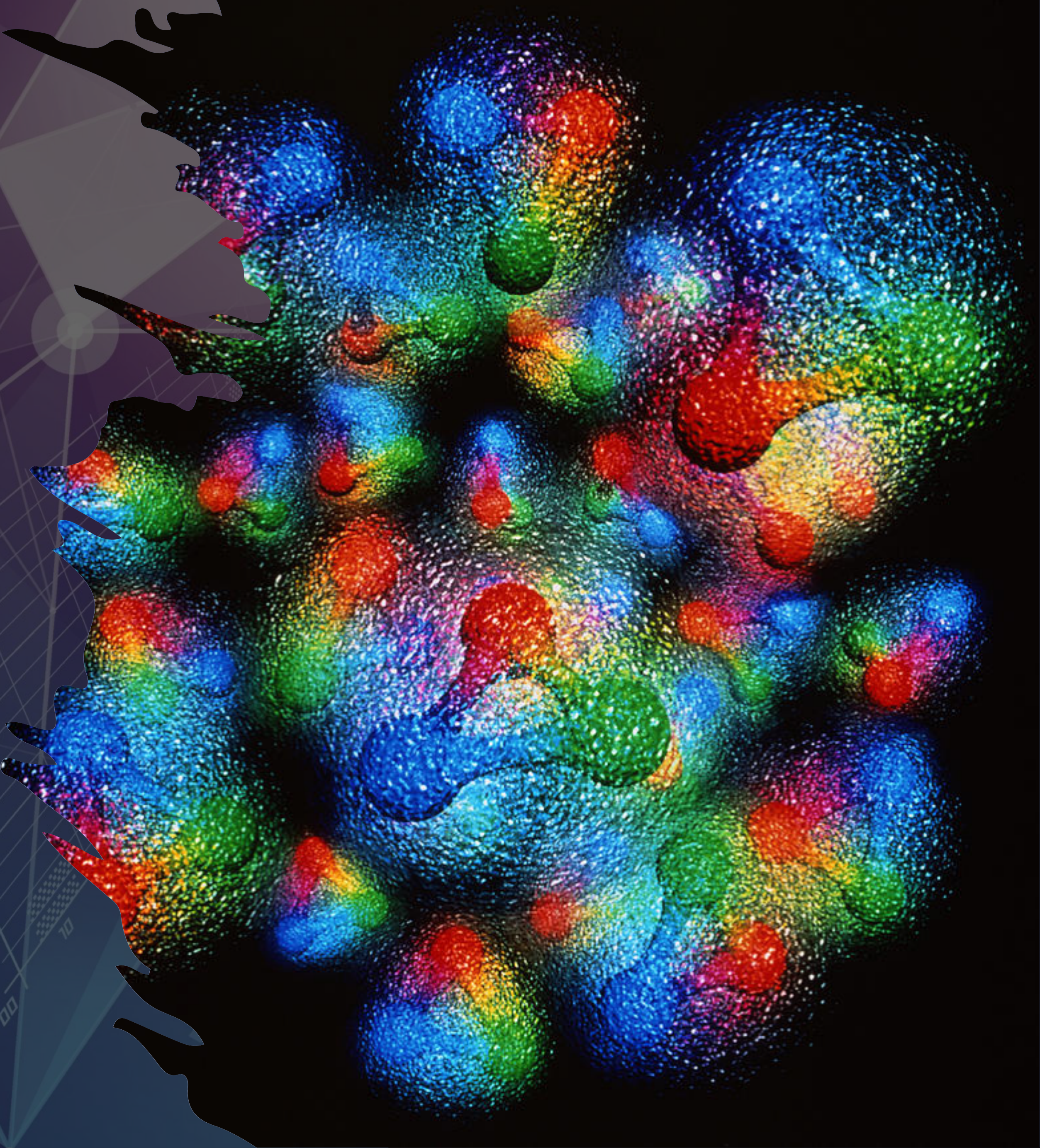
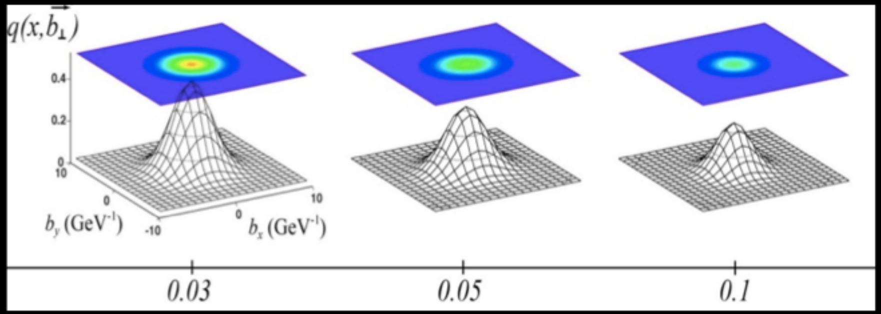
Entanglement Entropy in DVCS

Simonetta Liuti

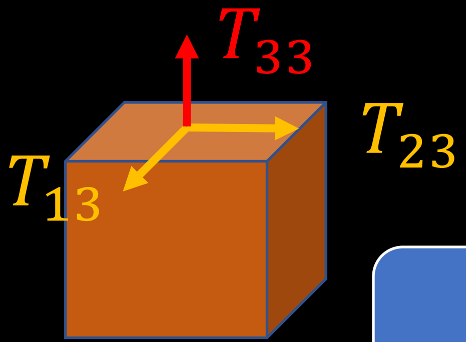


CENTER for NUCLEAR FEMTOGRAPHY



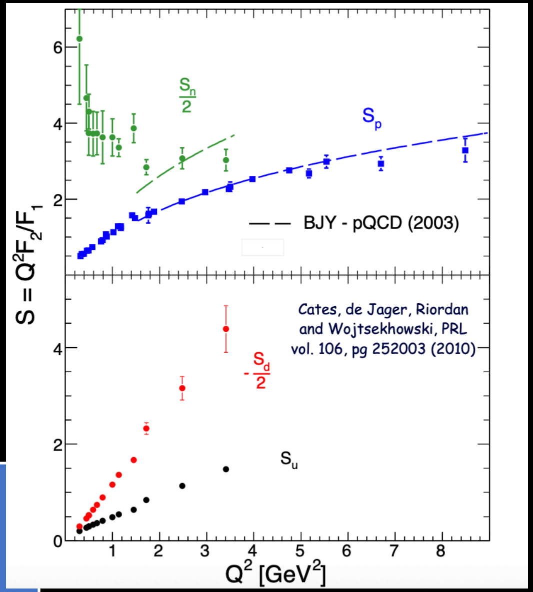
Fourier transforms of GPDs (M. Defurne)

Imaging the nuclear initial state



QCD Energy Momentum Tensor content

Gluon dynamics and flavor/color symmetries



Imaging the nuclear initial state

3D Coordinate Space Representation – Gluon Results

Zaki Panjsheeri's talk

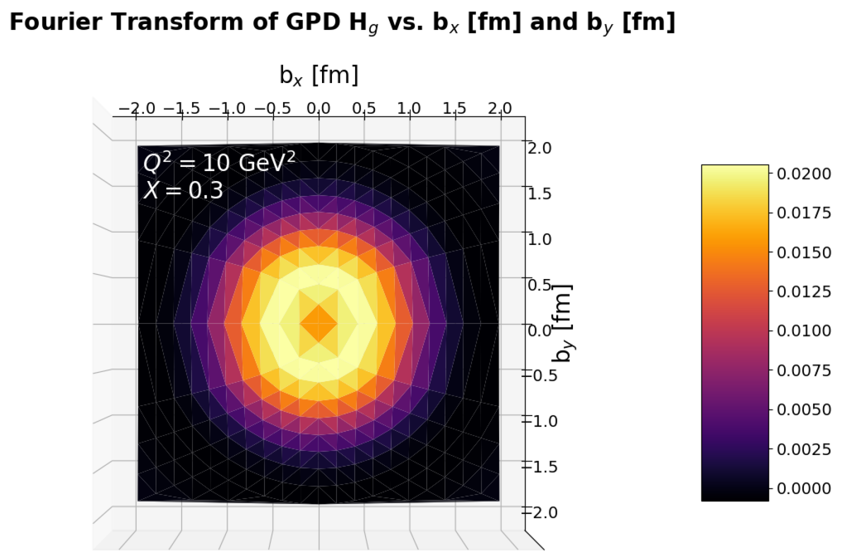
- GPDs can be Fourier transformed from momentum space into coordinate space, providing insight into matter, charge, and radial distributions of the quarks and gluons inside the proton.

M. Burkardt, Phys. Rev. D72, 094020 (2005), hepph/0505189.

$$\mathcal{H}^q(X,0,b_T) = \int \frac{d^2\Delta_T}{(2\pi)^2} H^q(X,0,\Delta_T) e^{-i\Delta_T \cdot b_T}$$

UVA's parametrization constrained by lattice QCD and experiment:

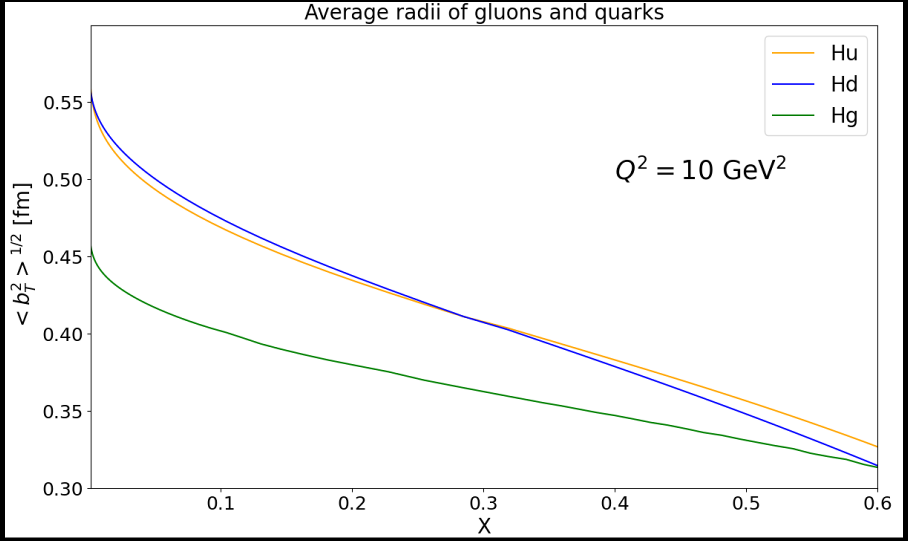
B. Kriesten. P. Velie, E. Yeats, F. Y. Lopez, & S. Liuti, *Phys.Rev.D* 105 (2022) 5, 056022



Average Distance spanned by quarks and gluons

- Expectation value of the transverse impact parameter distance
- The radius of the gluon matter density is smaller than the quark radius

$$< b_T^2 >^q (X) = \frac{\int_0^\infty d^2 b_T b_T^2 \mathcal{H}^q(X,0,b_T)}{\int_0^\infty d^2 b_T \mathcal{H}^q(X,0,b_T)}$$

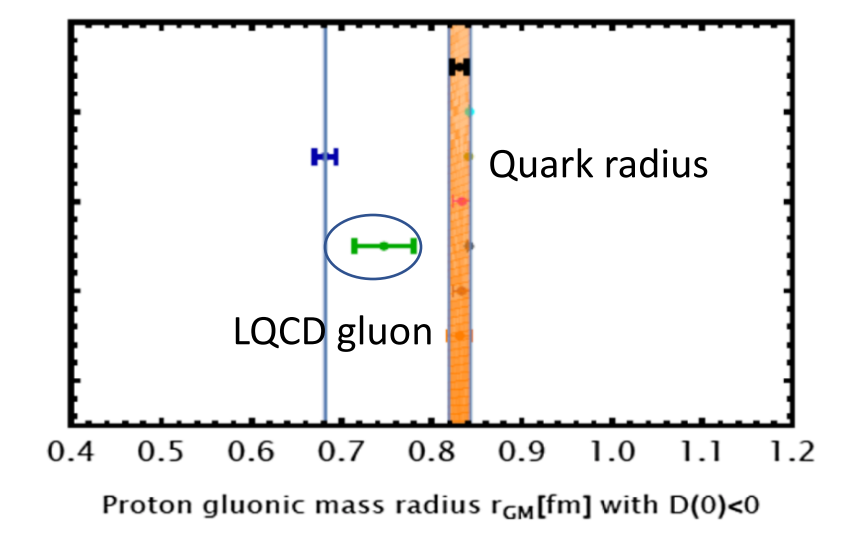


Compare to lattice and AdS/CFT results

K. Mamo and I. Zaeed

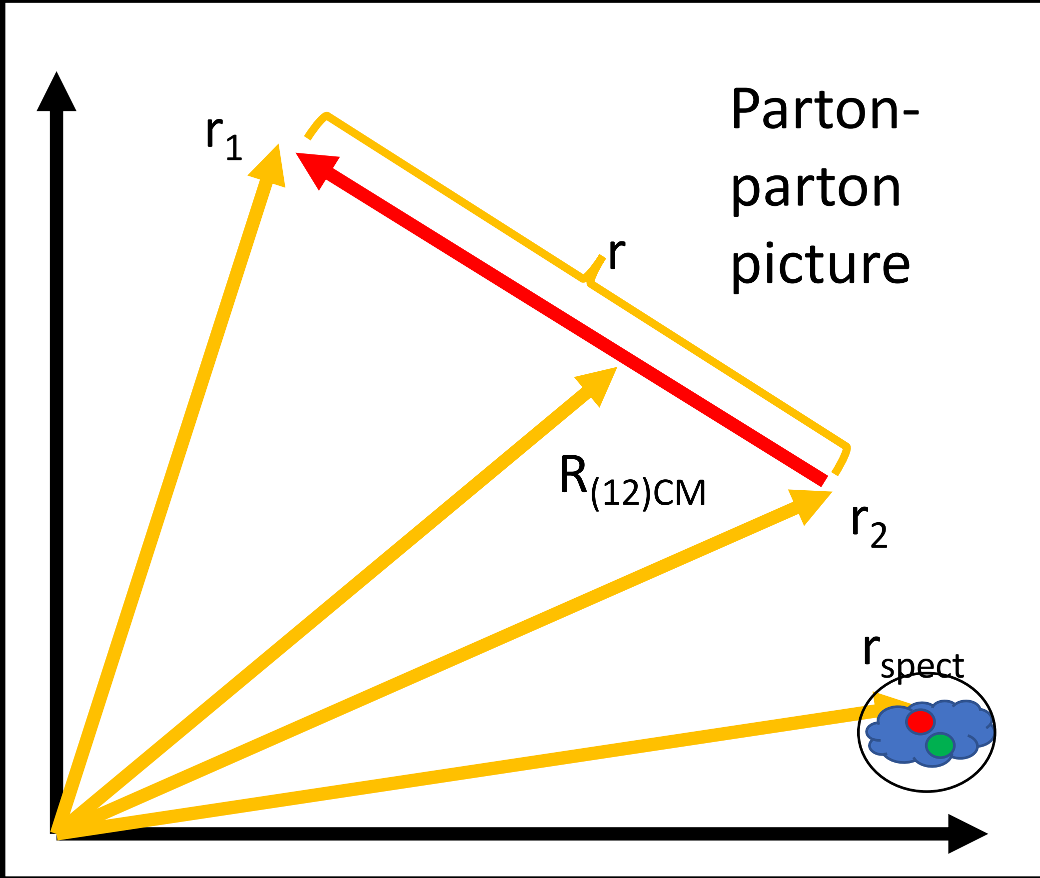
PRD106, 086004 (2022)

LQCD: Detmold and Shanahan

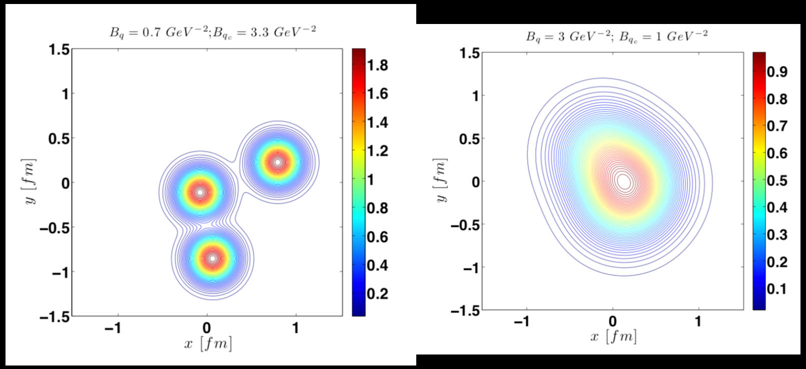


From one-body to two-body densities

- We can see a lot just from the one-body densities, but is that enough for imaging the proton's internal structure?
- We want to also understand how partons are situated relative to one another.



Quark and Gluon Imaging



Traini and Blaizot, PRD(2019)

Recent development

A more *differentia*l imaging, descriging the *event-by-event* quantum fluctuations in the wave function of the colliding hadron

H. Mantysaari, B. Schenke, F. Salazar et al.

arXiv 2001.10705 [hep-ph]

“Dynamic” images of gluon distributions forming hot-spots: can we connect them to GPDs?

- This emerging picture supports the idea of the gluons being at the core of the nucleon and carrying baryon number

D. Kharzeev

QCD Energy Momentum Tensor content

Deeply virtual exclusive experiments: quark and gluon angular momentum and the origin of the spin crisis

$$J_q + J_g = L_q + \frac{1}{2}\Sigma_q + J_g = \frac{1}{2}$$

This sum rule has longitudinal and transverse components, how do we access them through observables/quark and gluon distributions?

Fundamental role of deuteron through:

1. Extension of sum rule to spin 1 and interpretation of observables (SL)
2. Additional tensor-like observables

Quark OAM Integral Relation
for $\xi \longrightarrow 0$

$$
\begin{array}{ccccc}
J_L & = & L_L & + & S_L \\
\frac{1}{2}\int dx\, x(H+E) & = & \int dx\, x(\widetilde{E}_{2T}+H+E) & + & \frac{1}{2}\int dx\, \widetilde{H} \\
 & = & -\int dx\, F_{14}^{(1)} & + & \frac{1}{2}\int dx\, \widetilde{H}
\end{array}
$$

- A. Rajan, A. Courtoy, M. Engelhardt, S.L., PRD (2016)
- A. Rajan, M. Engelhardt, S.L., PRD (2018)
- Twist 3 GPD notation from Meissner, Metz and Schlegel, JHEP(2009)

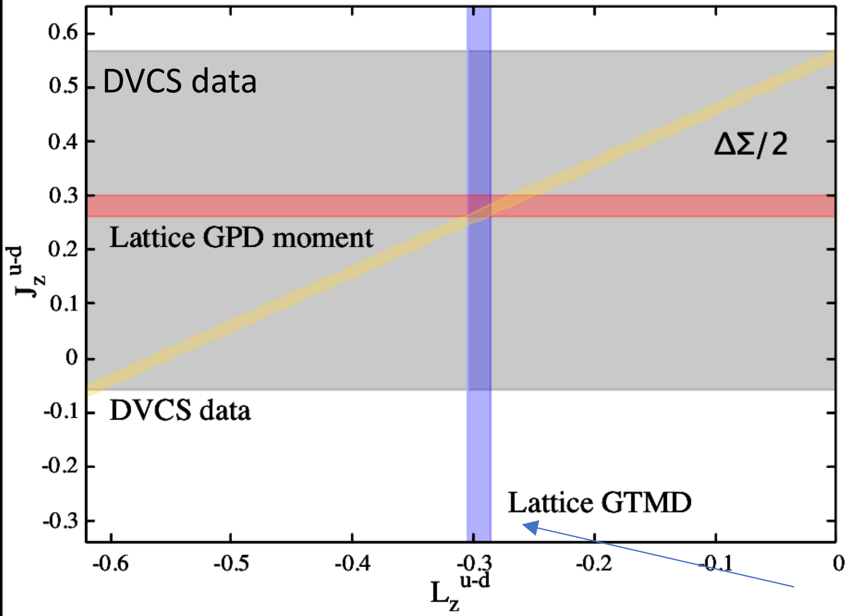
For $\xi \neq 0$

$$\frac{d}{dx}F_{14}^{(1)} = H + E + \widetilde{E}_{2T} - \xi E_{2T}$$

Generalized Lorentz Invariance Relation (LIR)

Putting this all together: what we know from measurements and lattice

m_π = 320 MeV

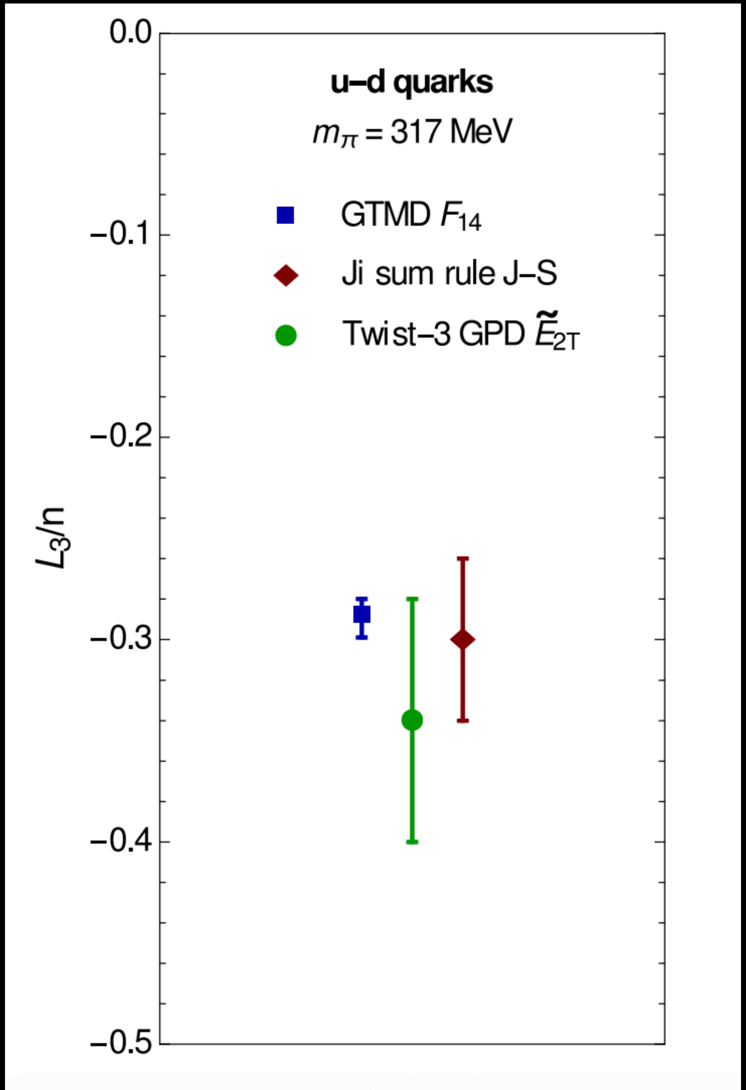


Mazouz et al.(Jlab Hall A) PRL (2007)

Engelhardt et al., Phys. Rev. D102, (2020)

q=u-d

$$J_q = L_q + \frac{1}{2}\Delta\Sigma_q$$



M. Engelhardt, Lattice 2023

How do we separate twist two and twist three components?

Twist 3 GPDs Physical Interpretation

	GPD	$P_q P_p$	TMD	Ref. 1
	$H^{\perp}$	UU	$f^{\perp}$	$2\widetilde{H}_{2T} + E_{2T}$
	$\widetilde{H}_L^{\perp}$	LL	$g_L^{\perp}$	$2\widetilde{H}'_{2T} + E'_{2T}$
J_L	$H_L^{\perp}$	UL	$f_L^{\perp\,(*)}$	$\widetilde{E}_{2T} - \xi E_{2T}$
J_L	$\widetilde{H}^{\perp}$	LU	$g^{\perp\,(*)}$	$\widetilde{E}'_{2T} - \xi E'_{2T}$
J_T	$H_T^{(3)}$	UT	$f_T^{(*)}$	$H_{2T} + \tau \widetilde{H}_{2T}$
J_T	$\widetilde{H}_T^{(3)}$	LT	g'_T	$H'_{2T} + \tau \widetilde{H}'_{2T}$

1/Q correction to H

1/Q correction to $\widetilde{\mathrm{H}}$

NEW!! Orbital Angular Momentum L

NEW!! Spin Orbit correlation $L \cdot S$

NEW!! Transverse OAM L_T

Transverse spin

(*) T-odd

[1] Meissner, Metz and Schlegel, JHEP(2009)

A. Rajan, A. Courtoy, M. Engelhardt, S.L., PRD (2016)

A. Rajan, M. Engelhardt, S.L., PRD (2018)

A. Rajan, O. Alkassasbeh, M. Engelhardt, S.L., (2023)

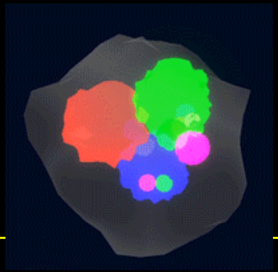
8/8/23 20

Defining the Benchmarks for a Global Analysis of Deeply Virtual Exclusive Experiments:

M. Almaeen et al. arXiv 2207.10766

The EXCLAIM project

(EXCLusives with Artificial Intelligence and Machine learning)



DOE funded collaboration, Co-PIs:

Computer Science: Gia Wei Chern, Yaohang Li

Experiment: Marie Boer

Lattice QCD: Michel Engelhardt, Huey Wen Lin

Phenomenology/ Theory: Gary Goldstein, S.L., Matt Sievert

Affiliates:

Aurore Courtoy, Tanja Horn, Brandon Kriesten, Pawel Nadolsky, Dennis Sivers

UVA students: Joshua Bautista, Adil Khawaja, Zaki Panjsheeri

In the process of hiring several postdocs!

OUR PROGRAM

1. To develop *physics informed* networks that include *theory constraints* in *deep learning* models.

2. ML is not treated as a set of "black boxes" whose working is not fully controllable

3. Utilize concepts in *information theory and quantum information theory* to interpret the working of ML algorithms necessary to extract information from data

4. At the same time, use ML methods as a testing ground for the working of quantum information theory in deeply virtual exclusive processes, as well as for inclusive processes

1. Theory constraints

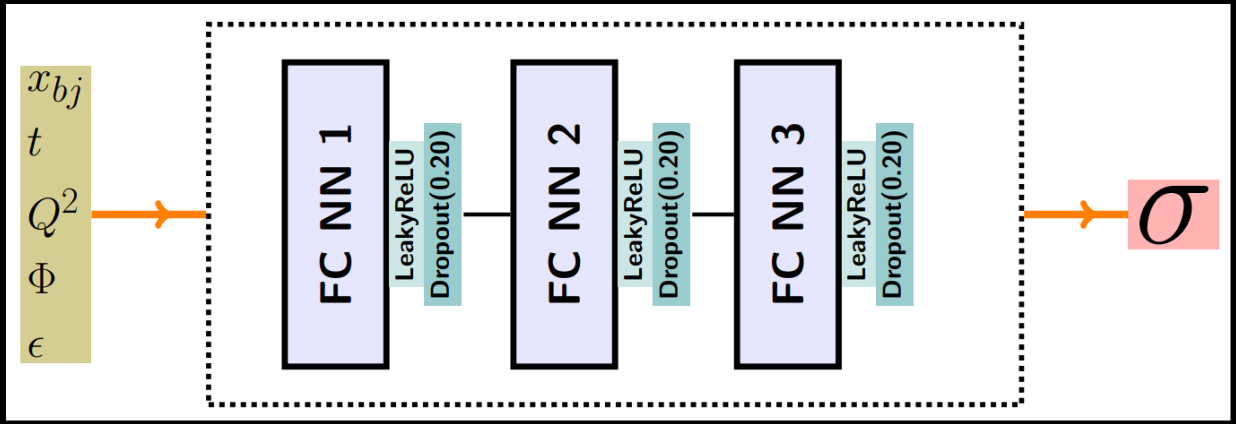
Hard constraints
"built into the architecture of the network"

- network invertibility
- choice of **activation functions**
- defining customized neural network layers

Soft constraints
"adding additional terms to the loss function that can be learned to minimize and generate physics weighted parameters"

1. Cross section structure
2. Lorentz invariance
3. Positivity constraints
4. Forward kinematic limit, defined by ξ, $t \to 0$, to PDFs, when applicable
5. $\Re$e-$\Im$m connection of CFFs through dispersion relations with proper consideration of threshold effects

Standard Feed Forward Deep Neural Network (DNN)



Loss function

$$\mathcal{L}_{\theta_t} = \frac{1}{2}(y - \boxed{f_{\theta_t}(x))^2}$$

$$\theta_{t+1} = \theta_t - \alpha \nabla_{\theta_t} \frac{1}{B} \sum_{i=0}^{B} \mathcal{L}_{\theta_t}(x^{(i)}, y^{(i)})$$

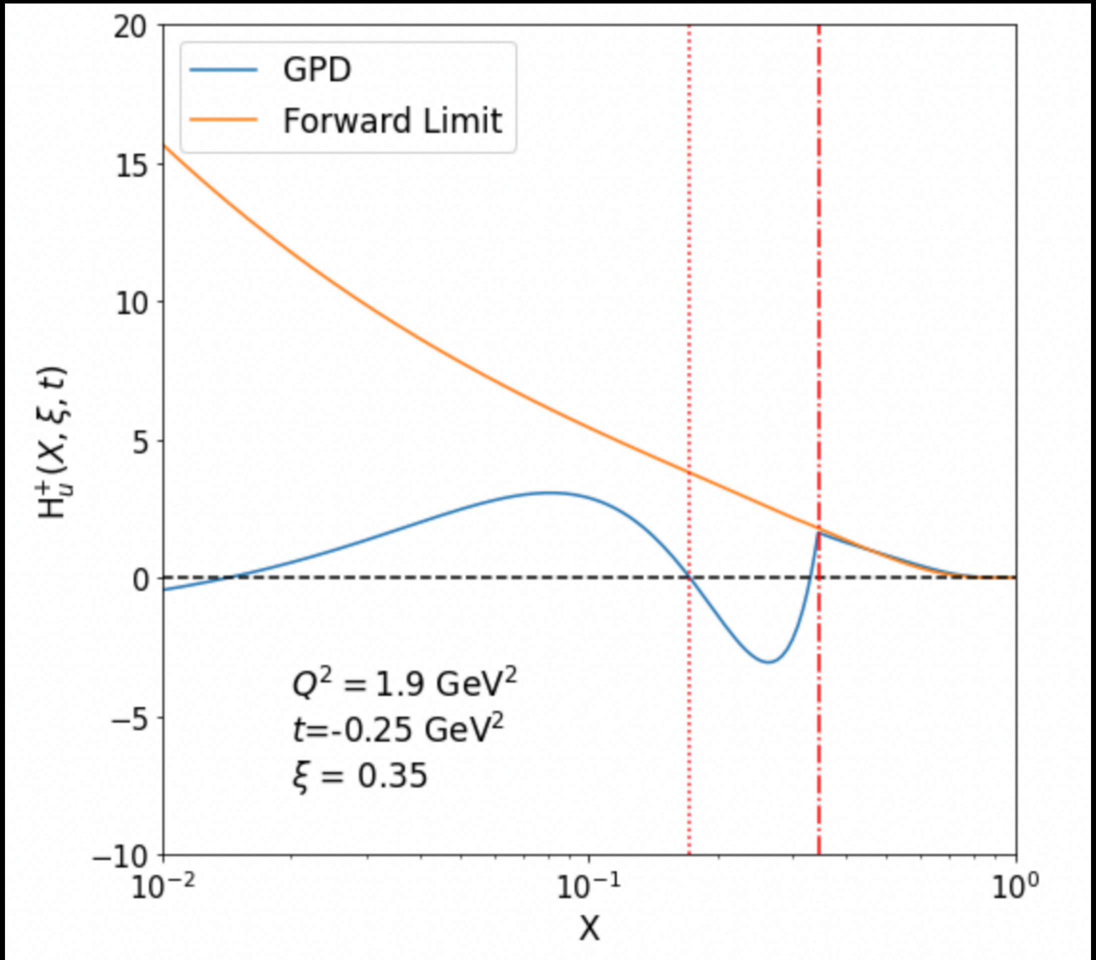
θ_i Hyperparameters (weights)

x Input Data

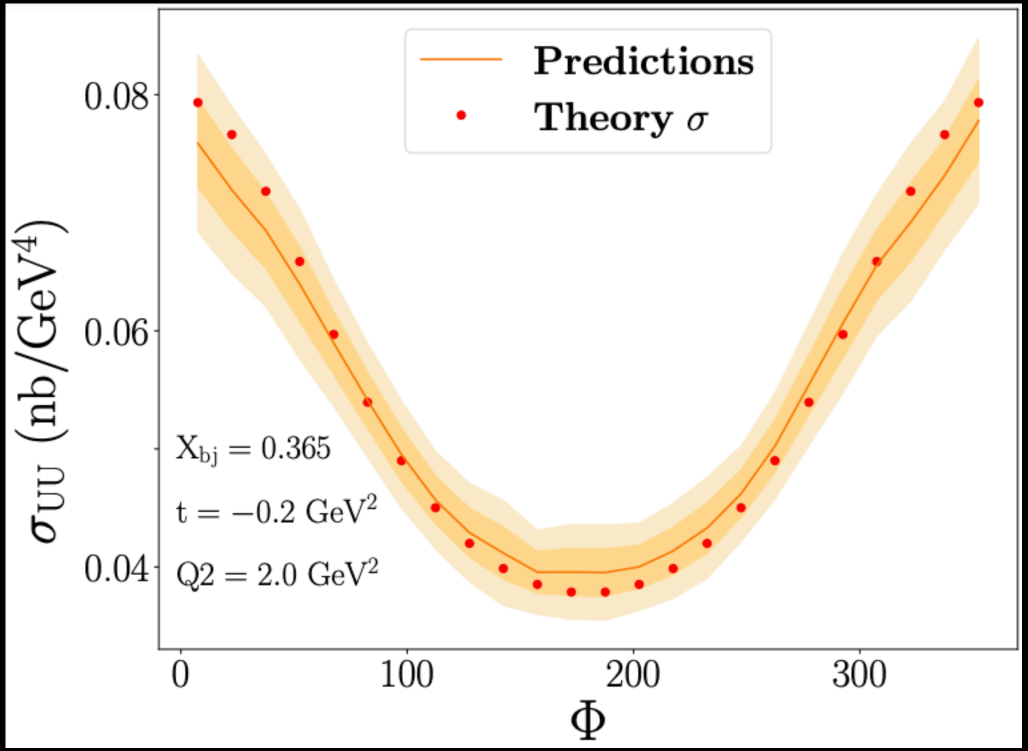
y Labels = predictions

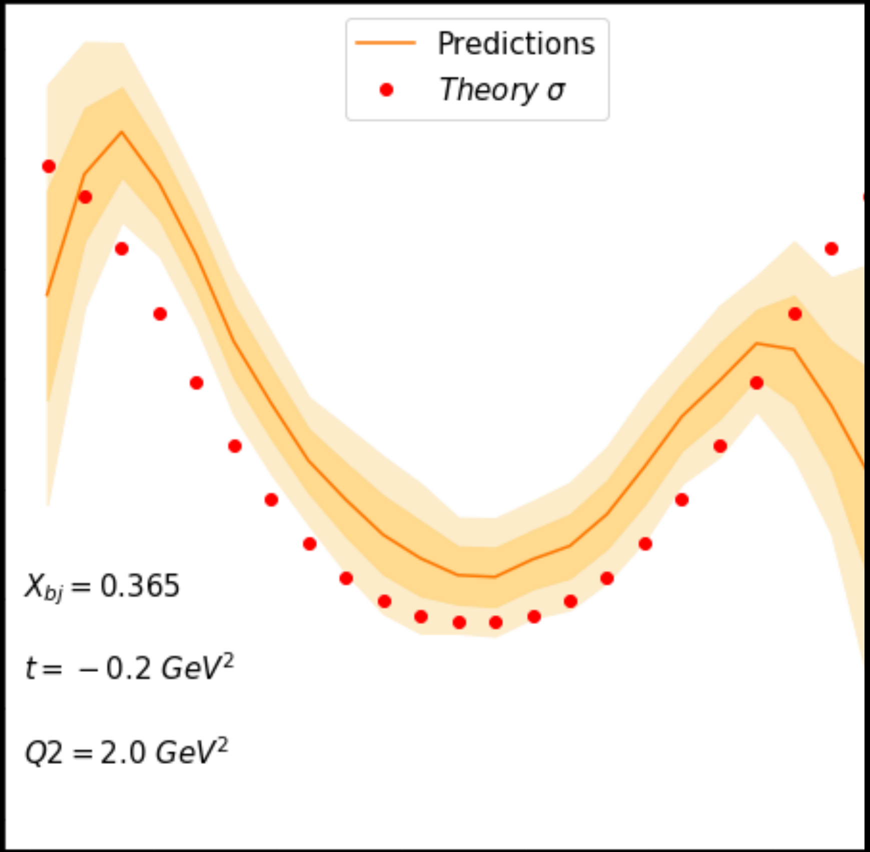
- Backpropagation = minimization through Stochastic Gradient Descent (SGD) controlled by learning rate α
- Hyperparameters shift in a direction that has a lower prediction error

Examples of soft constraints

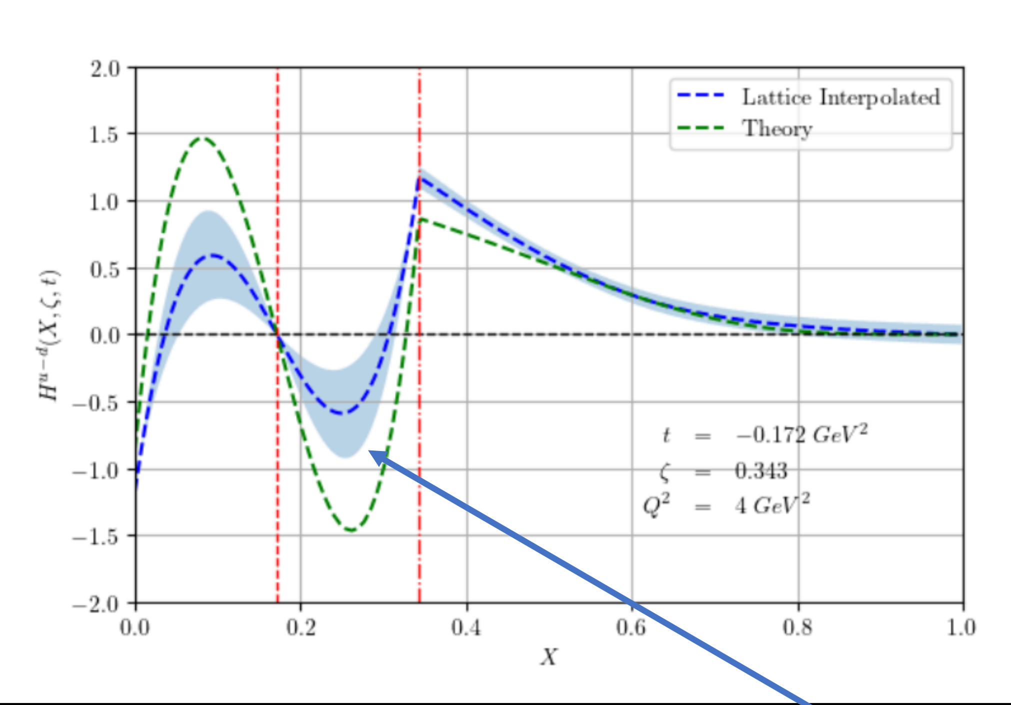


Bethe-Heitler contribution with and without parity constraint





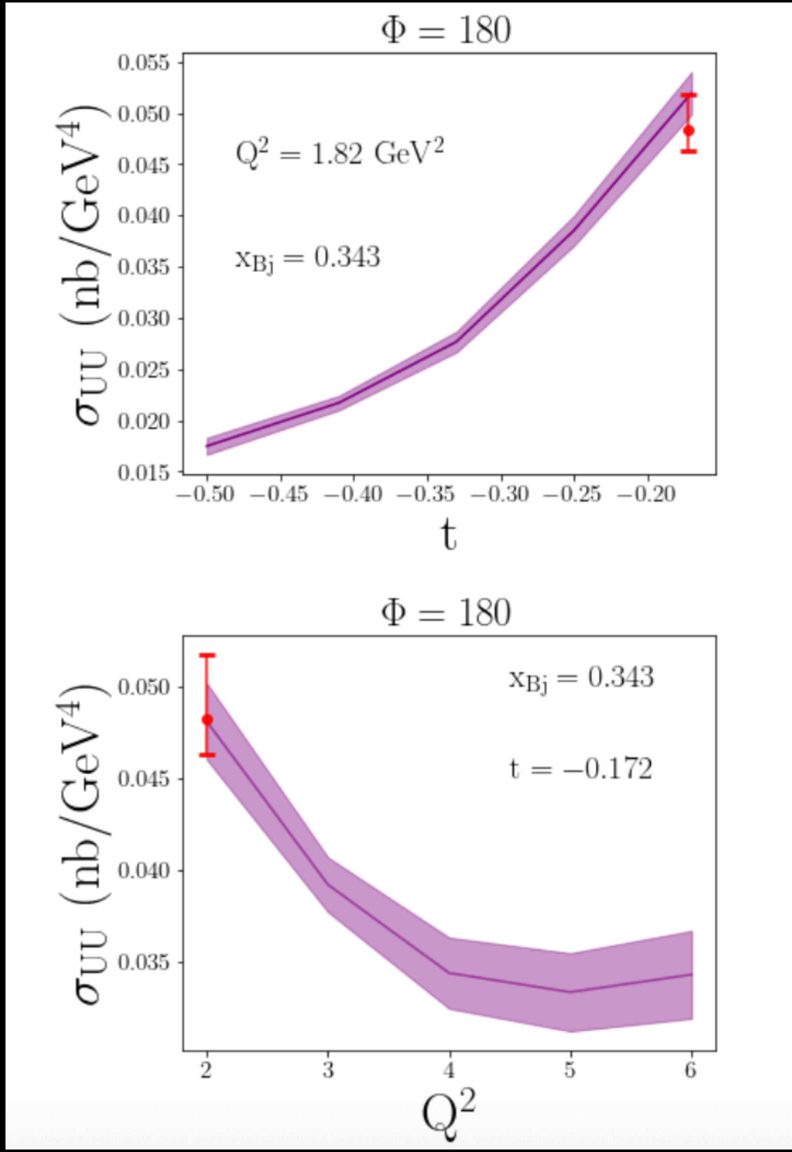
Using lattice moments to constrain ERBL rregion



$$M_2^q(\zeta,t) = \int_{\zeta/2}^{1} \frac{dX}{1-\zeta/2}\left(\frac{X-\zeta/2}{1-\zeta/2}\right) H^+(X,\zeta,t)$$
$$= A_{2,0}^q(t) + 4\left(\frac{\zeta}{2-\zeta}\right)^2 C_{2,0}^q(t)$$

With lattice constraint

How our DNN generalizes trends in t and Q^2



Parametrization of DVCS cross section

$$|T^{BH}_{UU}|^2 = \frac{\Gamma}{t}\Big[A^{BH}_{UU}\,(F_1^2+\tau F_2^2) + B^{BH}_{UU}\,\tau G_M^2(t)\Big]$$

$$|T^{\mathcal{I}}_{UU}|^2 = \frac{\Gamma}{Q^2 t}\Big[A^{\mathcal{I}}_{UU}\,\Re e\,(F_1\mathcal{H}+\tau F_2\mathcal{E}) + B^{\mathcal{I}}_{UU} G_M\,\Re e\,(\mathcal{H}+\mathcal{E}) + C^{\mathcal{I}}_{UU} G_M\,\Re e\widetilde{\mathcal{H}}\Big]$$

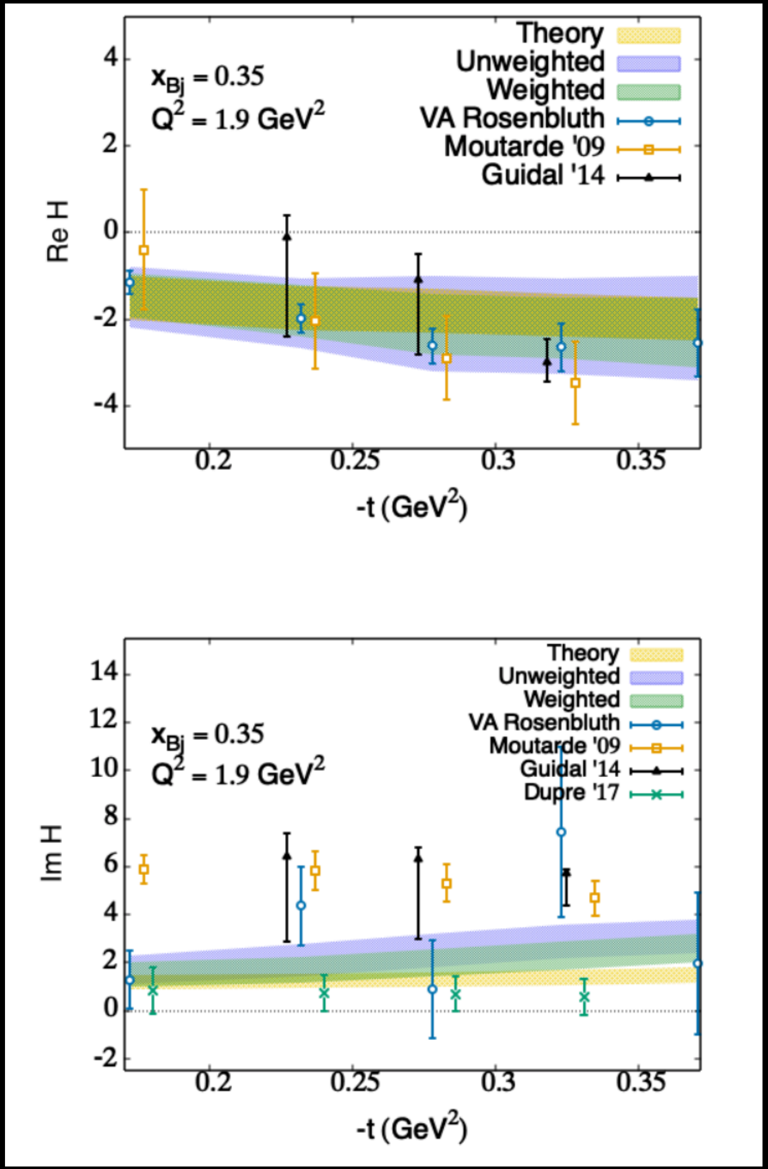
$$|T^{\mathcal{I}}_{LU}|^2 = \frac{\Gamma}{Q^2 t}\Big[A^{\mathcal{I}}_{LU}\,\Im m\,(F_1\mathcal{H}+\tau F_2\mathcal{E}) + B^{\mathcal{I}}_{LU} G_M\,\Im m\,(\mathcal{H}+\mathcal{E}) + C^{\mathcal{I}}_{LU} G_M\,\Im m\widetilde{\mathcal{H}}\Big]$$

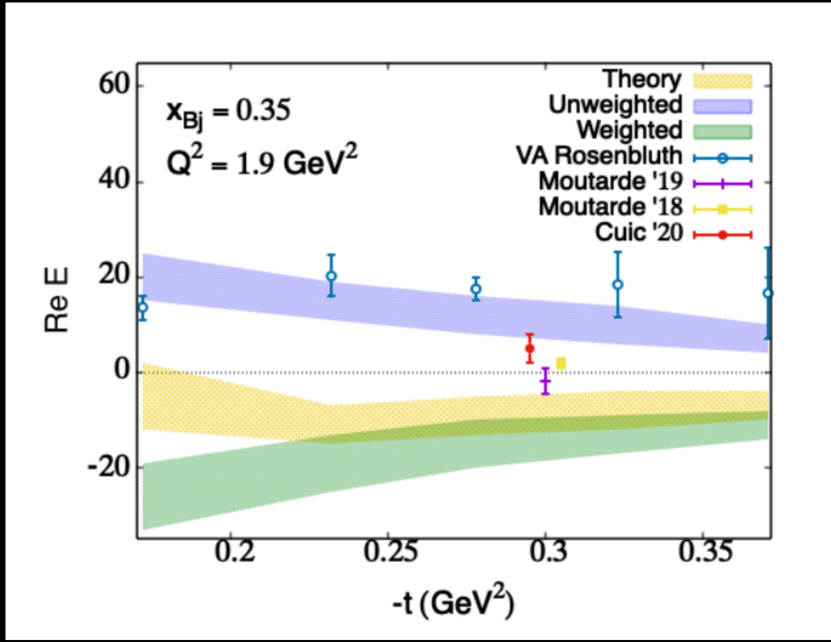
$$\begin{aligned}|T^{DVCS}_{UU}|^2 = \frac{\Gamma}{Q^2}\frac{2}{1-\epsilon}\Big[&(1-\xi^2)\Big[(\Re e\mathcal{H})^2+(\Im m\mathcal{H})^2+(\Re e\widetilde{\mathcal{H}})^2+(\Im m\widetilde{\mathcal{H}})^2\Big]\\
&+\frac{t_o-t}{4M^2}\Big[(\Re e\mathcal{E})^2+(\Im m\mathcal{E})^2+\xi^2(\Re e\widetilde{\mathcal{E}})^2+\xi^2(\Im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

Differs from BKM!!

- B. Kriesten et al, *Phys.Rev. D* 101 (2020)
- B. Kriesten and S. Liuti, *Phys.Rev. D105 (2022),* arXiv 2004.08890
- B. Kriesten and S. Liuti, Phys. Lett. B829 (2022), arXiv:2011.04484

Extraction of CFFs

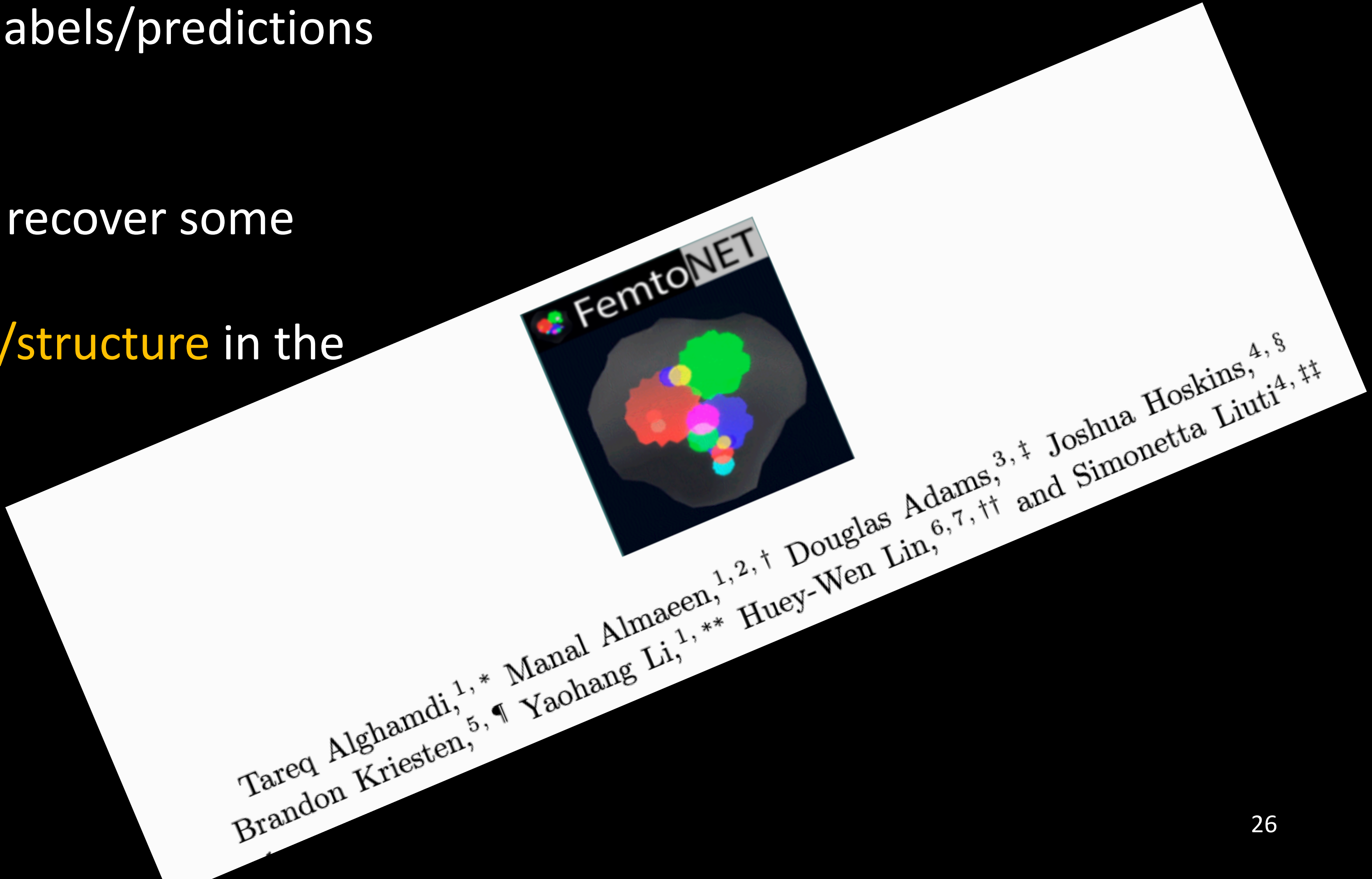




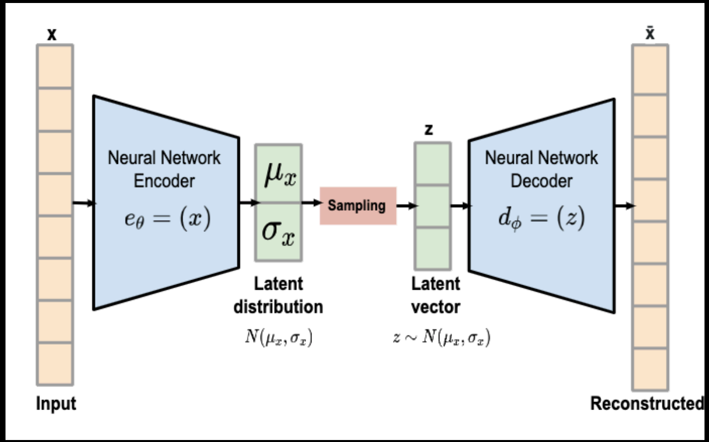
2./ 3. Using information theory to understand Machine Learning
Variational Autoencoder Inverse Mapper (VAIM)

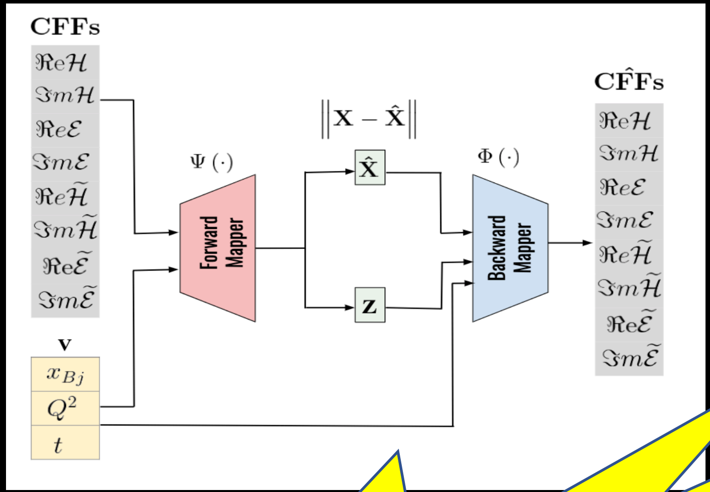
- Unsupervised learning: learning problems using input data, but no labels/predictions are used.

- The goal of learning is to recover some underlying, nontrivial, information/correlations/structure in the dataset.



2./ 3. Using information theory to understand Machine Learning
Variational Autoencoder Inverse Mapper (VAIM)





- ➢ Layers in neural networks keep enough information about the input so that the output labels can be predicted
- ➢ … but they discard the unnecessary information and focus on the essential learned representation

Entropy enters VAIM based approach at every step!

New: Missing Information can be described also in this case introducing the concept of ***entropy***

Entropy of a random variable, (X)

$$H(X) = -\sum_{x \in X} p(x) \log p(x) = \boldsymbol{E}\left(\frac{1}{\log p(x)}\right)$$

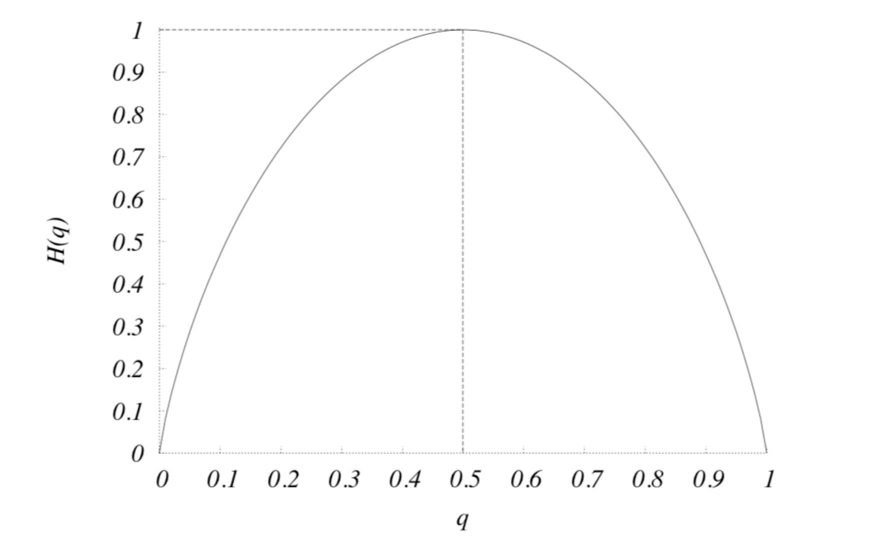
- the larger the entropy, the less a priori information one has on the value of the random variable.

Shannon Entropy: random variable, (X) takes values 0 with probability q and 1 with probability (1-q)

$$H(X) = -q \log_2 q - (1-q) \log_2 (1-q)$$

$$0 \leq H(X) \leq \log_2 d$$

d is the number of values that X can take



von Neumann Entropy

$$S(\hat{\rho}) = -\mathrm{Tr}\hat{\rho}\log\hat{\rho}$$

Spectral representation →

$$\hat{\rho} = -\sum_i \lambda_i \mid \lambda_i\rangle\langle\lambda_i \mid$$

Evaluation in terms of eigenvalues →

$$S(\hat{\rho}) = -\sum_i \lambda_i \log \lambda_i$$

Von Neumann Entropy of pure state = 0 (by definition)

Maximum von Neuman entropy

$$S(\hat{I}/d) = \log d$$

Entanglement Entropy: Given a bipartite quantum state of the composite system, it is possible to obtain a reduced density matrix describing knowledge of the state of a subsystem:

$$S(\hat{\rho}_{AB}) = -\sum_{i,j} p_i \rho_i \otimes \sigma_j$$

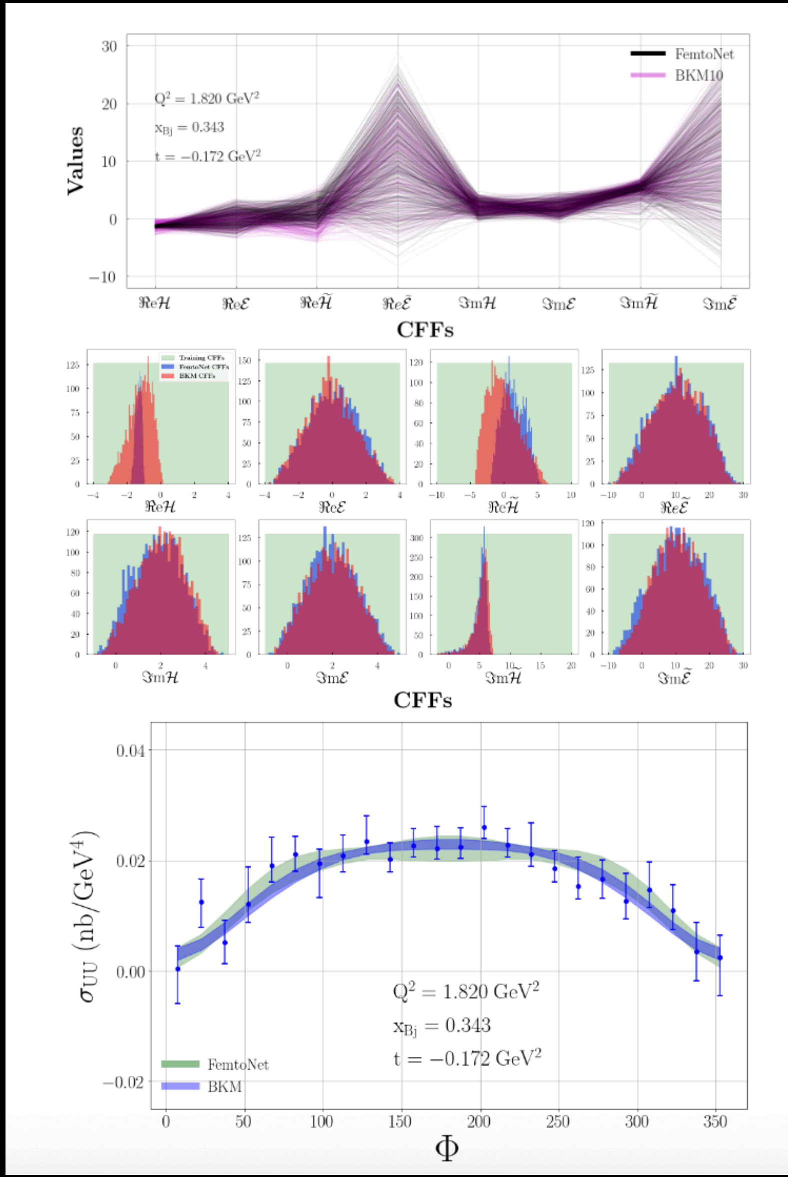
This entanglement entropy measures the degree of quantum entanglement between two subsystems constituting a two-part composite quantum system

Entanglement Entropy in DIS

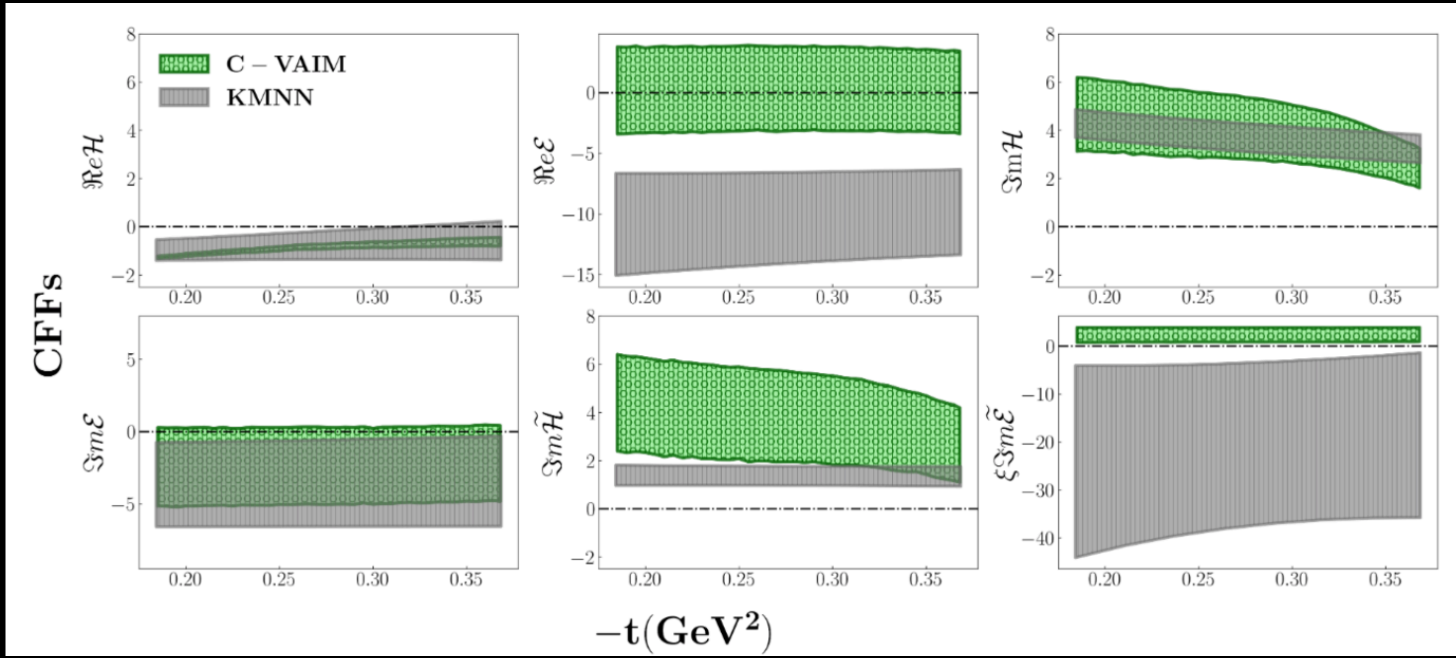
D. Kharzeev and E. Levin, PRD95 (2017)114008

- The proton as probed by DIS becomes a bipartite system with :
  A a localized region of radius ~1/Q
  B the region outside
- The von Neuman entropy of the system is: $\boxed{S(x) = \log x\, G(x)}$
- All microstates of the system are equally probable, and the von Neumann entropy is maximal -> parton saturation.
- In this case, for x ->0 the system is maximally entangled
- How do we go about this? This final state entropy can be measured from the multiplicity distribution of the produced hadrons.

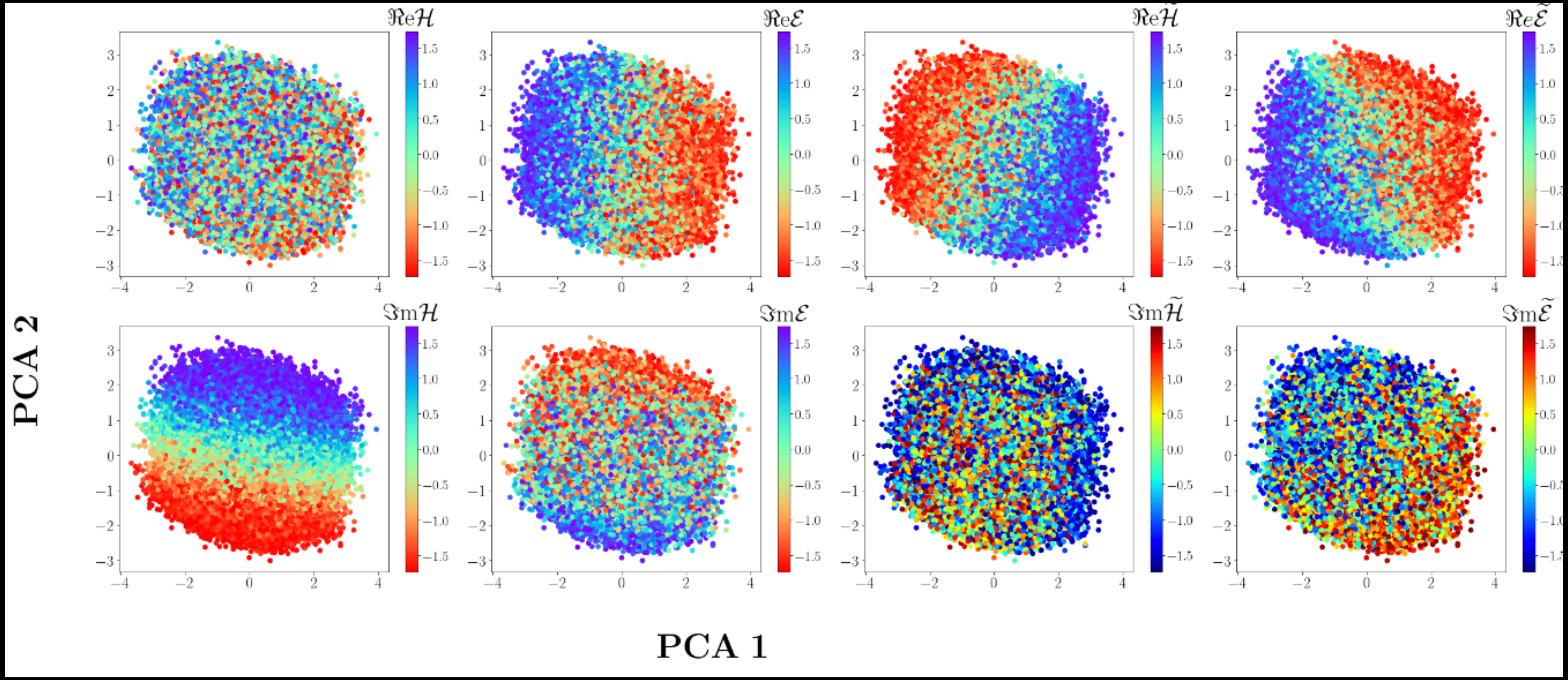
- Compton Form Factors at a specific kinematic point



- Point for discussion: in order to compare results from different approaches we need to define the **benchmarks!!!!**



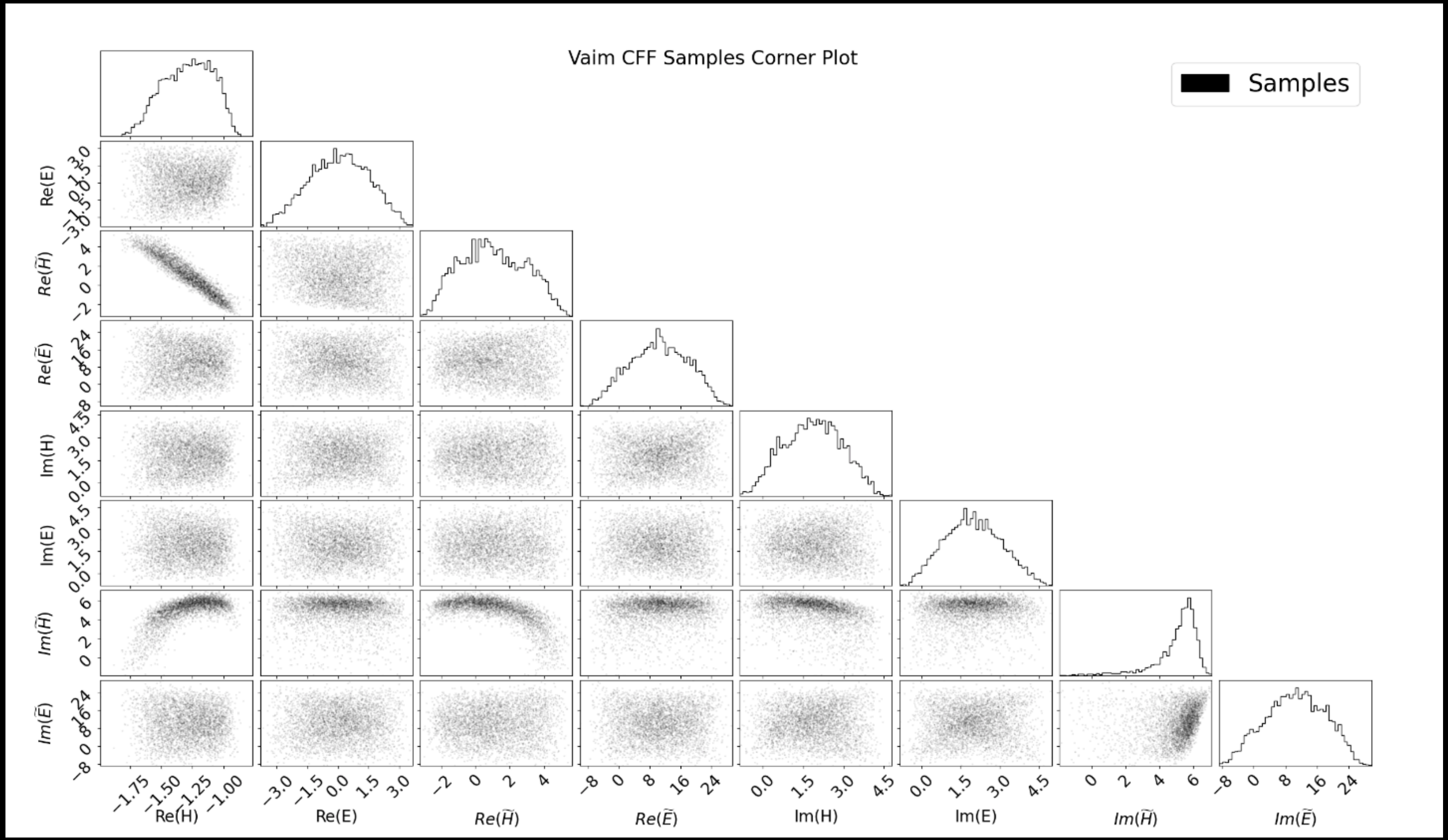
Entropy measurement from analysis of latent space



Manal Almaeen

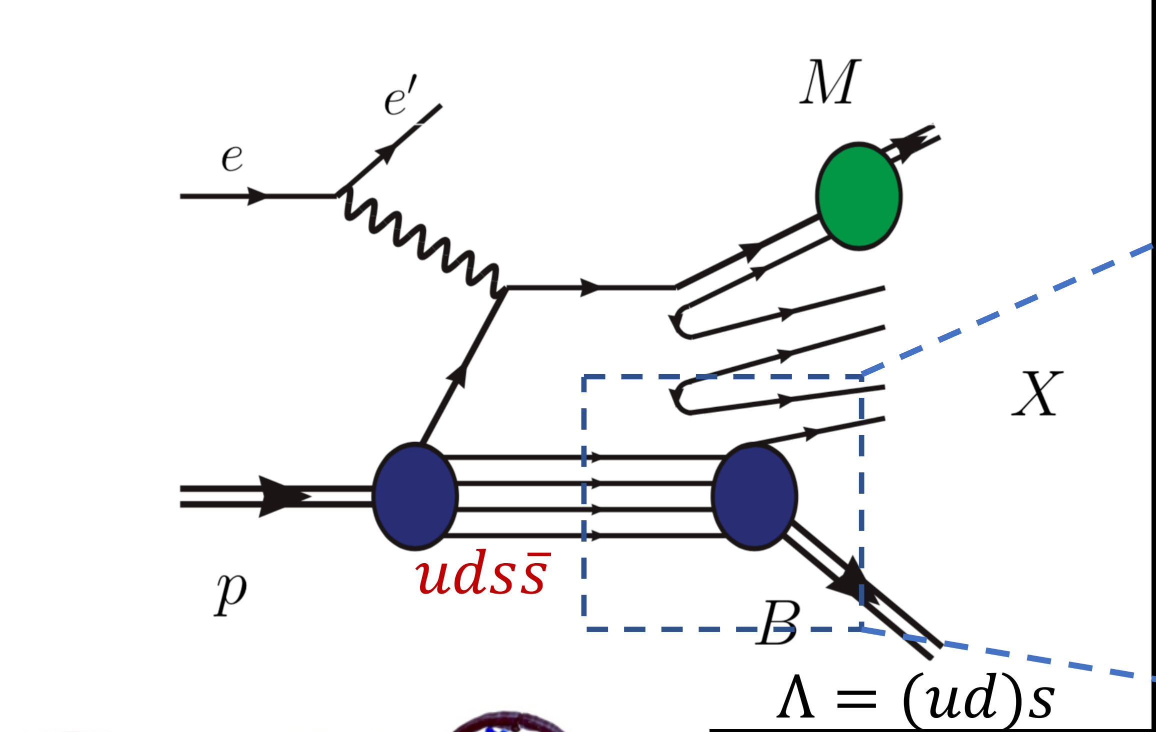
- some instances the PCA of latent space have clear structures that can be interpreted in terms of the symmetry of the theory.
- other cases there are structures that are formed but without a clear interpretation.
- Can entanglement entropy help with the interpretation?

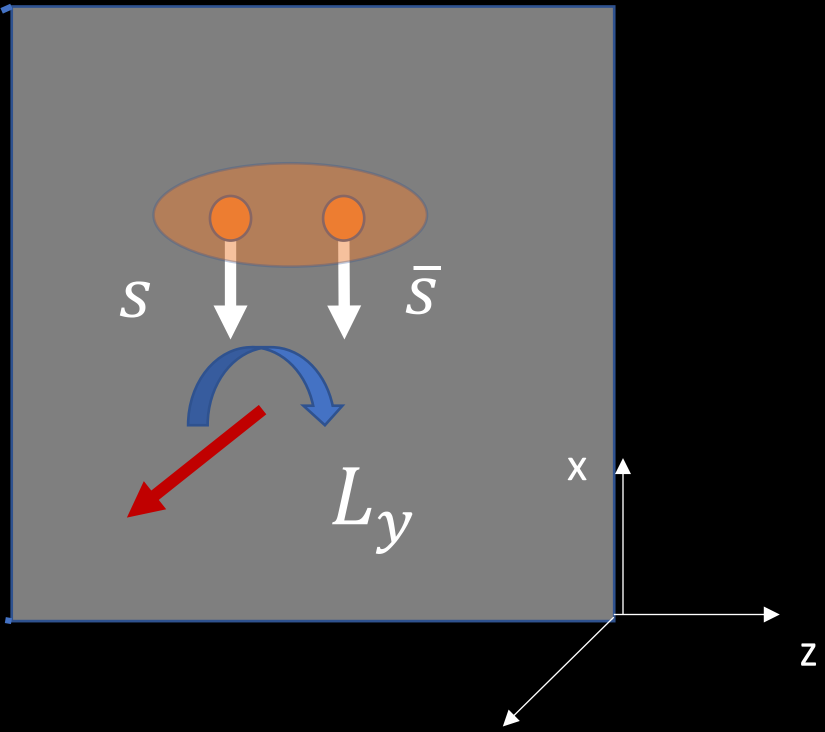
- A clear example is given by the interpretation of the latent space in VAIM



Douglas Adams

Another point where entanglement comes in: measuring the finals states (with G.Goldstein and D. Sivers)





The (ud) diquark is produced with spin S = 0

As the flux tube stretches an ss¯-pair is produced with transverse momentum, k_T

The ss¯- pair have their spins aligned to balance the OAM

Rotational motion generates the breaking of the string

Conclusions and Outlook

- Extracting 3D information from data is an unprecedented challenging problem which is uniquely highly-dimensional with respect to what done in DIS and SIDIS
- it is important to keep developing ML-based approaches and to build a platform with benchmarks for the community to compare results with both epistemic and aleatory uncertainties
- VAIM and similar unsupervised methods in ML allow us to understand and unravel new correlations among data through the concept of latent space
- Entanglement Entropy introduced as a key notion to quantify the information that can be extracted from our system through DVCS experiments.
- Future application: from CFFs to GPDs