Global CFFs fit with Neural Networks: Toward dispersion relation constraints and D-term extraction

Melany Higuera-Angulo

In collaboration with Alexandre Camsonne, Volker Burkert, David Richards, Daniel Lersch, Pierre Chatagnon.

Vector Quarkonia as Pressure Gauges
Jefferson Lab, Newport News, Virginia
March 27th, 2026

Jefferson Lab

Outline

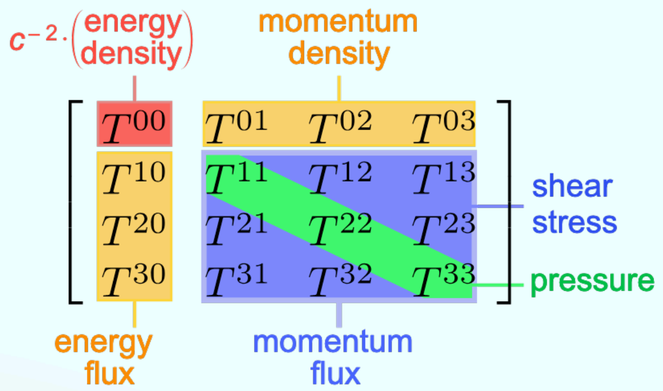
1. Motivation:
   - Mechanical properties of the proton.
   - From observables to Compton form factors.
2. Framework:
   - Gepard + Neural Networks
   - Extraction of CFFs from experimental data using NN.
3. State of the analysis:
   - Model stability: Closure test.
   - Update on global analysis of CFFs.
   - Dispersion relation and D-term extraction roadmap.
4. Deliverables and future directions

Motivation and scope of the project:

- Mechanical properties of the proton.
- From DVCS observables to Compton form factors

Mechanical properties of the proton

The Matrix element of the QCD Energy-Momentum tensor encodes key information of the proton including the mass and spin, as the distributions of energy, angular momentum, and various mechanical properties such as, e.g., internal forces inside the system.



$$\langle p',\vec{s}'|T_a^{\mu\nu}|p,\vec{s}\rangle = \bar{u}(p',\vec{s}')\left[A_a(t)\frac{P^\mu P^\nu}{M_N} + D_a(t)\frac{\Delta^\mu\Delta^\nu - g^{\mu\nu}\Delta^2}{4M_N} + \bar{C}_a(t)M_N\,g^{\mu\nu} + J_a(t)\frac{P^{\{\mu}i\sigma^{\nu\}\lambda}\Delta_\lambda}{M_N} - S_a(t)\frac{P^{[\mu}i\sigma^{\nu]\lambda}\Delta_\lambda}{M_N}\right]u(p,\vec{s})$$

Momentum · Pressure · Total angular momentum · Anisotropy Pressure · Intrinsic angular momentum

Symmetric kinematical variables:

$P = (p'+p)/2$ and $\Delta = p'-p$

Free Dirac spinor: $u(p,\vec{s})$

GFFs: $A_a(t)$, $D_a(t)$, $\bar{C}_a(t)$, and $J_a(t)$

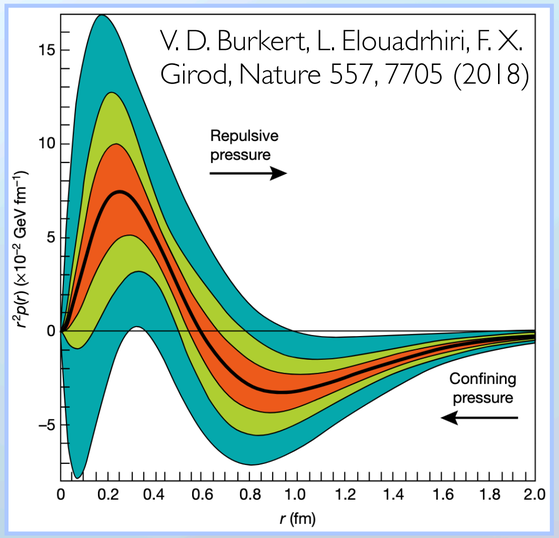
On top of restricting the number of GFFs, Poincaré symmetry imposes additional constraints:

$$\begin{aligned} A(0) &= \sum_q A_q(0) + A_G(0) = 1, \\ J(0) &= \sum_q J_q(0) + J_G(0) = \tfrac{1}{2}, \\ \tfrac{1}{2}\Delta\Sigma &= \sum_q S_q(0), \\ \bar{C}(t) &= \sum_q \bar{C}_q(t) + \bar{C}_G(t) = 0, \end{aligned}$$

D(t) is the lesser-known GFF. Its value is not fixed by spacetime symmetries and must be determined from experiment.

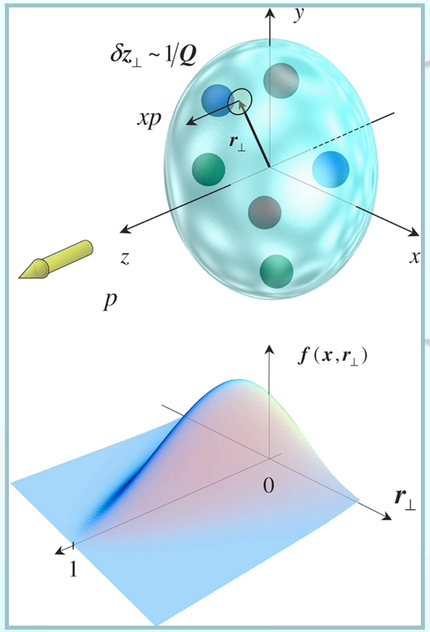
$D(t) = ?$ ⇒ $\Delta(t) \propto D(t) \propto \int d^3\mathbf{r}\; p(r)\frac{j_0(r\sqrt{-t})}{t}$

The D-term is not related to "external properties" of a particle like mass and spin, but to the stress tensor and internal force.



From DVCS observables to Compton form factors

The DVCS amplitude depends on the complex-valued convolution integrals of the GPDs: Compton form factors (CFFs)



We cannot access GPDs directly

$$\int_{-1}^{1} dx\; xH^q(x,\xi,t) = A^q(t) + \xi^2 D^q(t)$$

$$\int_{-1}^{1} dx\; xE^q(x,\xi,t) = B^q(t) - \xi^2 D^q(t)$$

twist 2 GPDs: $\{H, E, \tilde{H}, \tilde{E}\}$

$$\mathrm{Re}\mathcal{H}(\xi,t) + i\,\mathrm{Im}\mathcal{H}(\xi,t) = \sum_q e_q^2 \int_{-1}^{1} dx \left[\frac{1}{\xi - x - i\epsilon} - \frac{1}{\xi + x - i\epsilon}\right] H_q(x,\xi,t)$$

$\xi = x_B/(2 - x_B)$ is the skewness variable
x_B is the Bjorken scaling variable
t is the squared momentum transfer to the nucleon.

- The inverse problem is intrinsically ill-posed.

Thanks to the DVCS LO, it is possible to obtain the fixed-t dispersion relation (DR) where the subtraction constant is related to the so-called D-term.

Subtraction constant

$$\mathrm{Re}\mathcal{H}(\xi,t) = \Delta(t) + \frac{1}{\pi}\,\mathrm{P.V.}\int_0^1 d\xi' \left[\frac{1}{\xi - \xi'} - \frac{1}{\xi + \xi'}\right] \mathrm{Im}\mathcal{H}(\xi',t)$$

Where

$$\Delta(t) = 2\sum_q e_q^2 \int_{-1}^{1} \mathrm{d}z\, \frac{D^q_{\mathrm{term}}(z,t)}{1-z},$$

$$D^q_{\mathrm{term}}(z,t) = (1-z^2) \sum_{\mathrm{odd}\, n} d_n^q(t)\, C_n^{3/2}(z)$$

1. At leading order, we assume the subtraction constant is dominantly coming from quarks.
2. In the asymptotic limit $d_n^q(t)$ vanish for n>1.
3. Dominance of the flavor singlet combination.

$$D^Q(t) = \frac{18}{25}\Delta(t)\,.$$

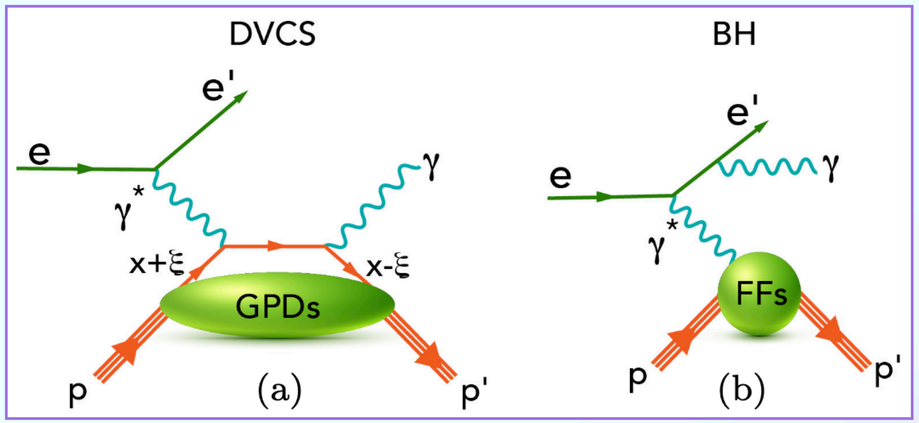
From DVCS observables to Compton form factors

Measurements of DVCS are mostly realized via the process of leptoproduction of a real photon, where also an interference with the Bethe-Heitler radiation also occurs

The BH process gives access to elastic form factors, it does not see the 3D structure of hadrons.

$$\frac{d^5\sigma}{dx_B dQ^2 d|t| d\phi d\phi_S} = \frac{\alpha^3 x_B}{16\pi^2 Q^4 \sqrt{1+\epsilon^2}} \left|\mathcal{T}\right|^2 ,$$

$$|\mathcal{T}|^2 = |\mathcal{T}_{\rm BH} + \mathcal{T}_{\rm DVCS}|^2 = |\mathcal{T}_{\rm BH}|^2 + |\mathcal{T}_{\rm DVCS}|^2 + \mathcal{I} \, .$$



The interference is actually what gives us strong sensitivity to the CFFs. The DVCS amplitude can be decomposed either in helicity amplitudes or, equivalently, in complex valued CFFs.

$$|\mathcal{T}_{\rm DVCS}|^2 = \frac{2\left(2-2y+y^2\right)}{y^2 Q^2 \left(2-x_{\rm B}\right)^2}\left[4\left(1-x_{\rm B}\right)\left(|\mathcal{H}|^2+|\widetilde{\mathcal{H}}|^2\right) - \left(x_{\rm B}^2+\left(2-x_{\rm B}\right)^2\frac{\Delta^2}{4M^2}\right)|\mathcal{E}|^2 \right.$$
$$\left. -x_{\rm B}^2\left(\mathcal{H}\mathcal{E}^*+\mathcal{E}\mathcal{H}^*+\widetilde{\mathcal{H}}\widetilde{\mathcal{E}}^*+\widetilde{\mathcal{E}}\widetilde{\mathcal{H}}^*\right) - x_{\rm B}^2\frac{\Delta^2}{4M^2}|\widetilde{\mathcal{E}}|^2\right], \quad y=\frac{Q^2}{xs}.$$

BMK equations relate the CFFs to measurable quantities such as $\Delta^4\sigma, d^4\sigma, A_{LU}, A_{UL}, A_C, A_{LL}\ldots$

The first sine harmonic of the beam spin asymmetry (BSA), as measured e.g. in Jefferson Lab is defined as:

- Beam spin asymmetry: $A_{LU}^{-,\sin\phi} \equiv \frac{1}{\pi}\int_{-\pi}^{\pi} d\phi \, \sin\phi \, A_{LU}^{-}(\phi)$

The CFFs are the direct bridge between theory and measurable data. →

For truly Global analysis, the CFFs are usually indirectly modeled by GPDs models. However, we don't do that.

Experimental observables

Observables used:

- Beam spin difference cross section:

$$\Delta^4\sigma = \frac{1}{2}\left[\frac{\mathrm{d}^4\sigma(\lambda=+1)}{\mathrm{d}Q^2\mathrm{d}x_\mathrm{B}\mathrm{d}t\mathrm{d}\phi} - \frac{\mathrm{d}^4\sigma(\lambda=-1)}{\mathrm{d}Q^2\mathrm{d}x_\mathrm{B}\mathrm{d}t\mathrm{d}\phi}\right]$$

- Beam spin sum cross section:

$$\mathrm{d}^4\sigma = \frac{1}{2}\left[\frac{\mathrm{d}^4\sigma(\lambda=+1)}{\mathrm{d}Q^2\mathrm{d}x_\mathrm{B}\mathrm{d}t\mathrm{d}\phi} + \frac{\mathrm{d}^4\sigma(\lambda=-1)}{\mathrm{d}Q^2\mathrm{d}x_\mathrm{B}\mathrm{d}t\mathrm{d}\phi}\right]$$

- Beam charge asymmetry:

$$A_\mathrm{C} = \frac{\sigma^+_\mathrm{UU} - \sigma^-_\mathrm{UU}}{\sigma^+_\mathrm{UU} + \sigma^-_\mathrm{UU}} \propto \Re\mathrm{e}\left[F_1\mathcal{H} + \xi\left(F_1+F_2\right)\widetilde{\mathcal{H}} - \frac{\Delta^2}{4M^2}F_2\mathcal{E}\right]\cos\phi$$

- Beam spin asymmetry:

$$A_\mathrm{LU}(\phi) = \frac{\mathrm{d}\sigma^\uparrow(\phi) - \mathrm{d}\sigma^\downarrow(\phi)}{\mathrm{d}\sigma^\uparrow(\phi) + \mathrm{d}\sigma^\downarrow(\phi)} \propto \Im\mathrm{m}\left\{F_1\mathcal{H} + \xi\left(F_1+F_2\right)\widetilde{\mathcal{H}} - \frac{\Delta^2}{4M^2}F_2\mathcal{E}\right\}\sin(\phi)$$

- Target spin asymmetry:

$$A_\mathrm{UL}(\phi) = \frac{\mathrm{d}\sigma^\Rightarrow(\phi) - \mathrm{d}\sigma^\Leftarrow(\phi)}{\mathrm{d}\sigma^\Rightarrow(\phi) + \mathrm{d}\sigma^\Leftarrow(\phi)} \propto \Im\mathrm{m}\left[F_1\widetilde{\mathcal{H}} + \xi\left(F_1+F_2\right)\left(\mathcal{H} + \frac{x_\mathrm{B}}{2}\mathcal{E}\right) - \xi\left(\frac{x_\mathrm{B}}{2}F_1 + \frac{t}{4M^2}F_2\right)\widetilde{\mathcal{E}}\right]\sin(\phi)$$

- Double spin asymmetry:

$$A_\mathrm{LL}(\phi) = \frac{\mathrm{d}\sigma^{\uparrow\Uparrow}(\phi) - \mathrm{d}\sigma^{\downarrow\Uparrow}(\phi) - \mathrm{d}\sigma^{\uparrow\Downarrow}(\phi) + \mathrm{d}\sigma^{\downarrow\Downarrow}(\phi)}{\mathrm{d}\sigma^{\uparrow\Uparrow}(\phi) + \mathrm{d}\sigma^{\downarrow\Uparrow}(\phi) + \mathrm{d}\sigma^{\uparrow\Downarrow}(\phi) + \mathrm{d}\sigma^{\downarrow\Downarrow}(\phi)} \propto \Re\mathrm{e}\left[F_1\widetilde{\mathcal{H}} + \xi\left(F_1+F_2\right)\left(\mathcal{H} + \frac{x_\mathrm{B}}{2}\mathcal{E}\right) - \xi\left(\frac{x_\mathrm{B}}{2}F_1 + \frac{t}{4M^2}F_2\right)\widetilde{\mathcal{E}}\right]\cos\phi + BH$$

Collab	Year	Observable	l.b.e.	x_B	$Q^2[GeV^2]$	$\lvert t\rvert[GeV^2]$
CLAS	2001	$A_{LU}^{\sin(\phi)}$	4.25	0.19	1.25	0.19
CLAS	2006	$A_{UL}^{\sin(\phi)}$	5.7	0.2-0.4	1.82	0.15-0.44
CLAS	2007	A_{LU}	5.77	0.13-0.35	1.1-3	0.1-0.3
CLAS	2009	A_{LU}	4.8	0.12-0.48	1.0-2.8	0.09-1.8
CLAS	2015	$A_{LU}(\phi), A_{UL}(\phi), A_{LL}(\phi)$	5.93	0.18-0.4	1.6-3.2	0.1-0.45
CLAS	2015	$\sigma(\phi), \Delta\sigma(\phi)$	5.75	0.1-0.58	1-4.6	0.09-0.52
CLAS	2018	$\sigma(\phi)$	5.88	0.12-0.5	1.1-4	0.1-1.6
CLAS	2023	A_{LU}	10.2	0.09-0.45	1.3-6	0.1-2.2
CLAS	2023	A_{LU}	10.6	0.09-0.62	1.1-7.2	0.1-2.8
CLAS	2026	$\sigma(\phi)$	11	0.72-0.5	1.1-5.0	0.12-0.91
HERMES	2001	$A_{LU}^{\sin(\phi)}$	27.6	0.11	2.6	0.27
HERMES	2006	$A_{C}^{\cos(\phi)}$	27.6	0.08-0.12	2.0-3.7	0.03-0.42
HERMES	2008	$A_C, A_{UT,I}, A_{UT,DVCS}$	27.6	0.03-0.35	1-10	<0.7
HERMES	2009	$A_C, A_{LU,I}, A_{LU,DVCS}$	27.6	0.05-0.24	1.2-5.75	<0.7
HERMES	2010	A_{UL}, A_{LL}	27.6	0.03-0.35	1-10	<0.7
HERMES	2011	$A_{LT,I}, A_{LT,BH+DVCS}$	27.6	0.03-0.35	1-10	<0.7
HERMES	2012	$A_{LU,I}, A_{LU,DVCS}, A_C$	27.6	0.03-0.35	1-10	<0.7
HALL A	2015	$\sigma(\phi), \Delta\sigma(\phi)$	5.75	0.33-0.40	1.5-2.6	0.17-0.37
HALLA	2017	$\sigma(\phi), \Delta\sigma(\phi)$	3.3-5.5	0.35	1.51-2	0.18-0.36

Lepton-beam-energy (l.b.e.)

We made cuts at t=0.5 l GeV^2, and Q^2<1.5 GeV^2 or -t/Q^2<0.25

If cross-sections are measured, using weighted Fourier integrals can help to cancel the strongly oscillating Bethe-Heitler propagators and interference terms, improving the convergence of harmonic terms.

Framework:

- Gepard + Neural Networks
- Extraction of CFFs from experimental data using NN.

Framework: Gepard + Neural Networks

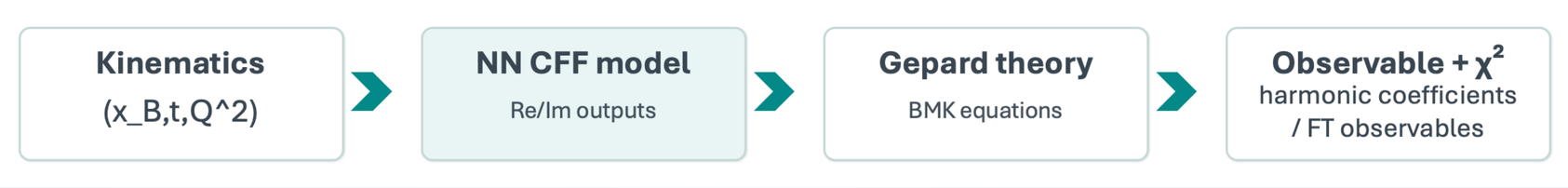
We are building upon existing tools: Gepard

- Python software framework dedicated to the study of GPDs.
- K. Kumerički, D. Müller, K. Passek-Kumerički, Nuclear Physics B, Volume 794, Issues 1–2, 2008

To obtain a more reliable estimate of uncertainties with reduced model dependence, we decided to utilize Artificial Neural Networks to parameterize the CFFs.

- Use State-of-the-art standard ML libraries. PyTorch
- A tool to do a Global fit to extract CFFs and GFFs to ensure the results are driven by data.

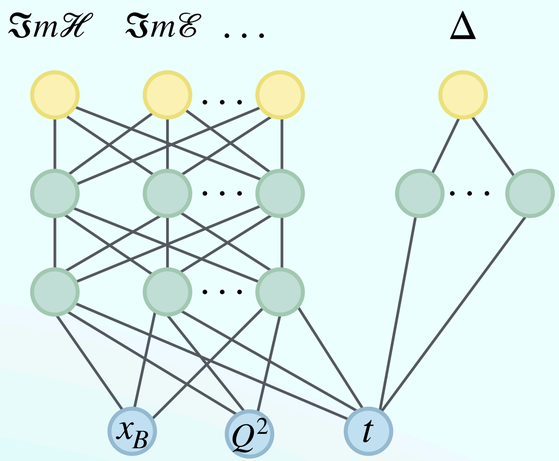
opengpd / gpddatabase: We developed a publicly accessible database format capable of storing all current and future data, as well as pseudodata. Gepard Interface and GitHub public access; C++ & Python APIs arxiv.org/abs/2503.18152

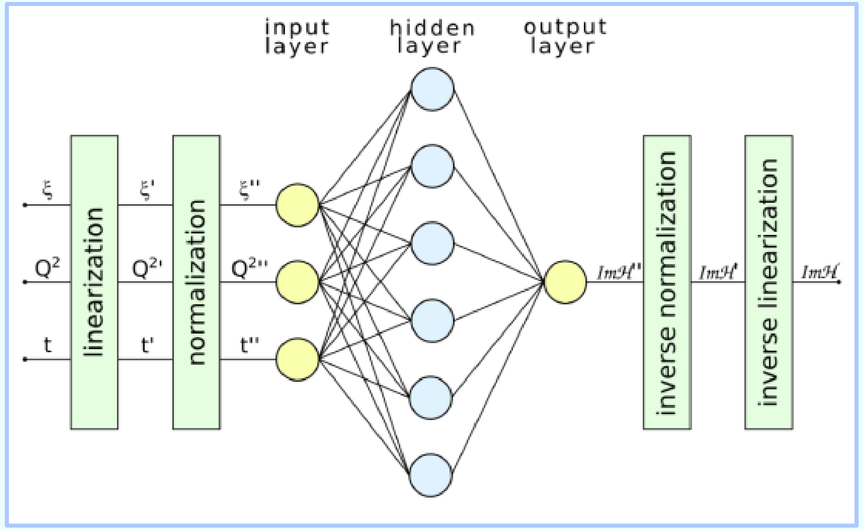


Extraction of CFFs from experimental data using NN

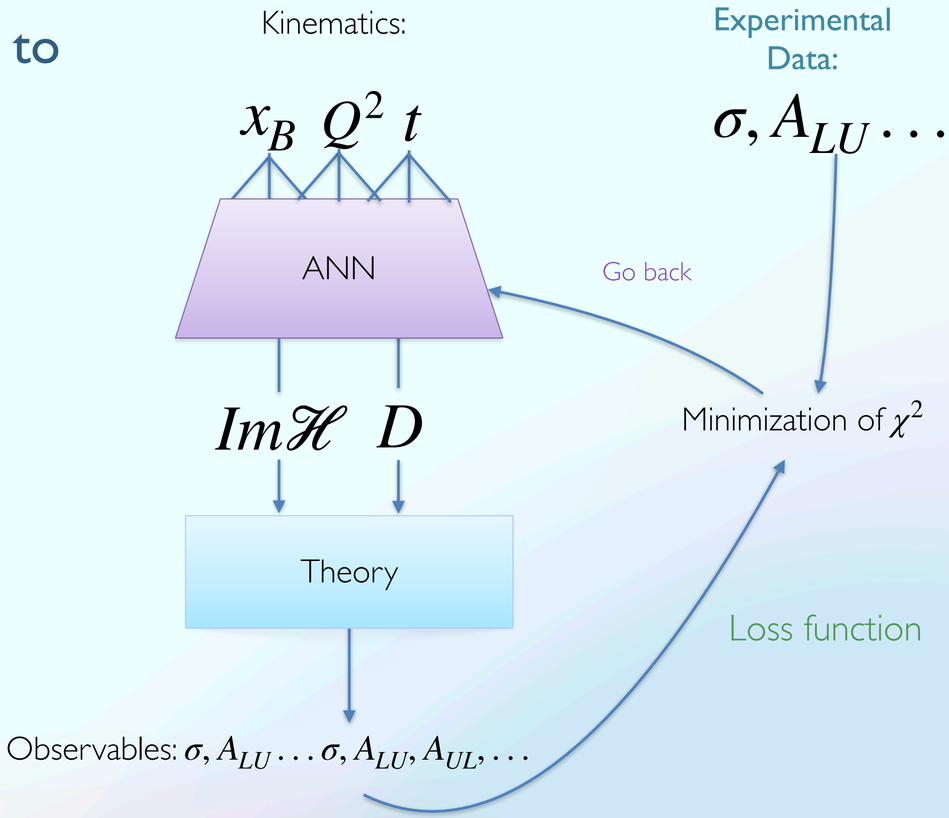
Artificial Neural Networks (ANN)

A machine learning (ML) algorithm that uses artificial intelligence (AI) to learn how to process data by analyzing training examples.





$$\mathrm{Re}\mathcal{H}(\xi,t) = \mathcal{C}_{\mathcal{H}}(t) + \frac{1}{\pi}\,\mathrm{P.V.}\int_0^1 \mathrm{d}\xi' \left[\frac{1}{\xi-\xi'} - \frac{1}{\xi+\xi'}\right] \mathrm{Im}\mathcal{H}(\xi',t)$$



$$\chi^2(\theta) = \frac{1}{N}\sum_i \left(\frac{\mathcal{O}^{\mathrm{Gepard}}(\mathrm{CFF}_\theta;\, kin_i) - \mathcal{O}_i^{\mathrm{data}}}{\sigma_i}\right)^2$$

➡The process is essentially a least-squares fit of a complex, multi-parameter function, implying no theoretical bias.

➡For a truly global coverage, it is necessary to add a Q^2 dependence, which was previously absent.

State of the analysis:

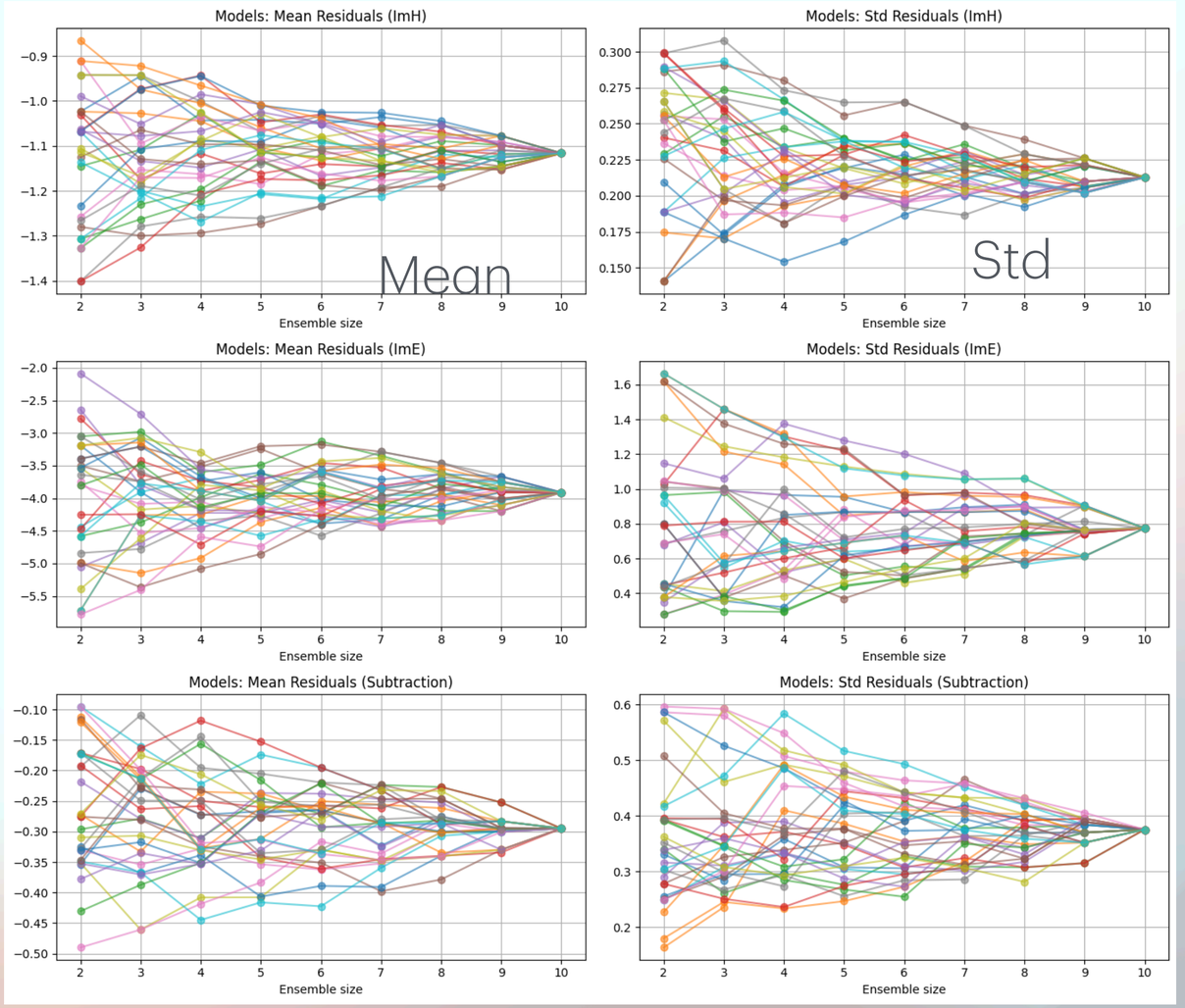
- Model stability and the Closure test.
- Update on global analysis of CFFs.
- Dispersion relation and D-term extraction roadmap.

Ensemble analysis of the Neural Networks

Models

To solve instability in results, we started ensemble analysis by averaging over multiple learned representations.

- To mitigate fluctuations and instability in NN results, we use an ensemble approach, averaging over multiple trained models.
- As the number of models in the ensemble increases, the mean residual prediction stabilizes, reducing noise.
- The standard deviation plots demonstrate that uncertainty decreases as the ensemble size grows.
- We can prune the results by selecting the best nets with a smaller chi2.

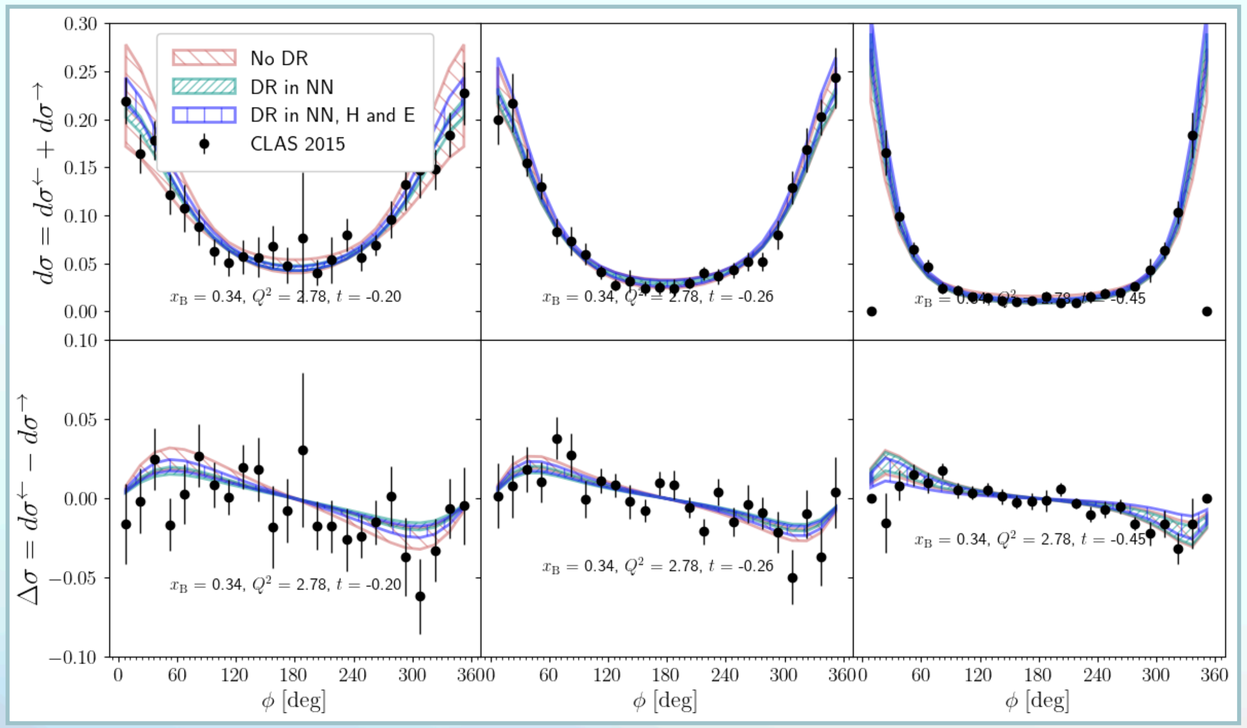


A larger ensemble tends to 'average out' individual model noise, improving the predictive stability results and uncertainty estimation of our model.

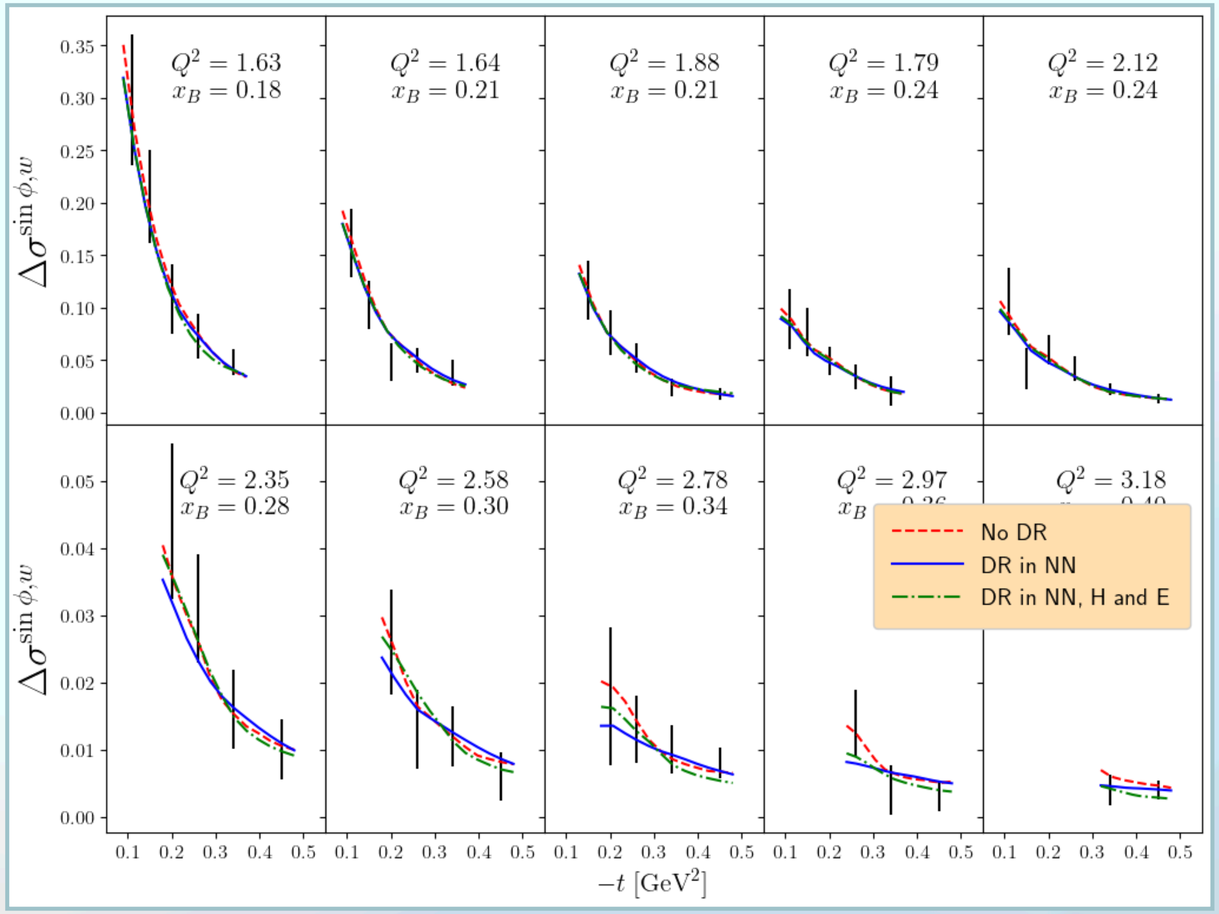
In addition, we can obtain reasonable behavior and reduce uncertainty according to the best hyperparameter selection.

CLAS fit

Example of description of 2015 CLAS data through the NN



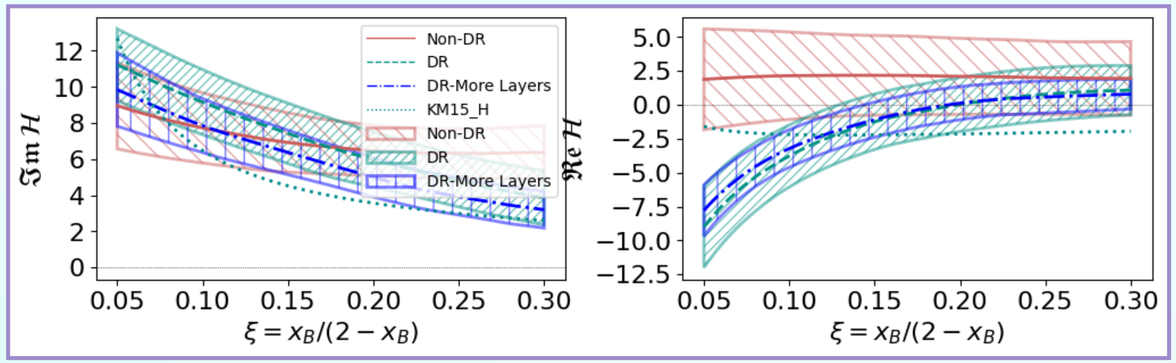
(CLAS Collaboration) Phys. Rev. Lett. 115, 212003,2015

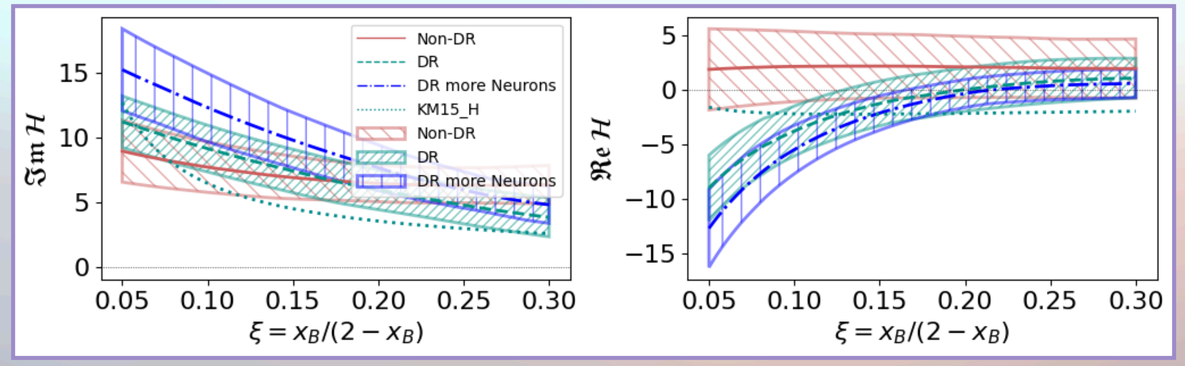


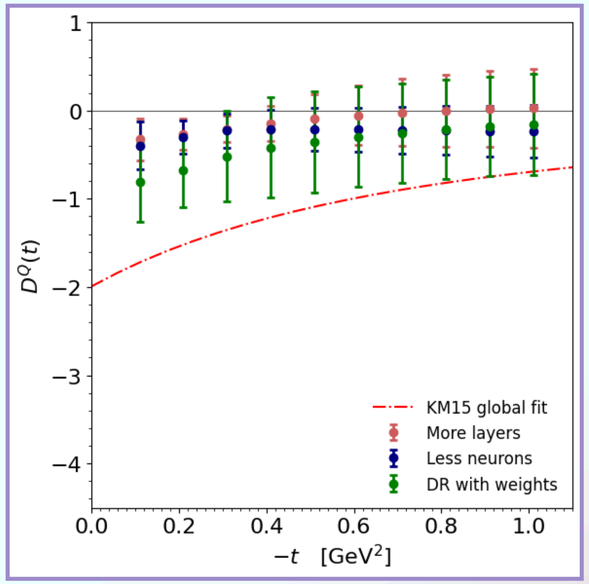
Model stability (architecture dependence)

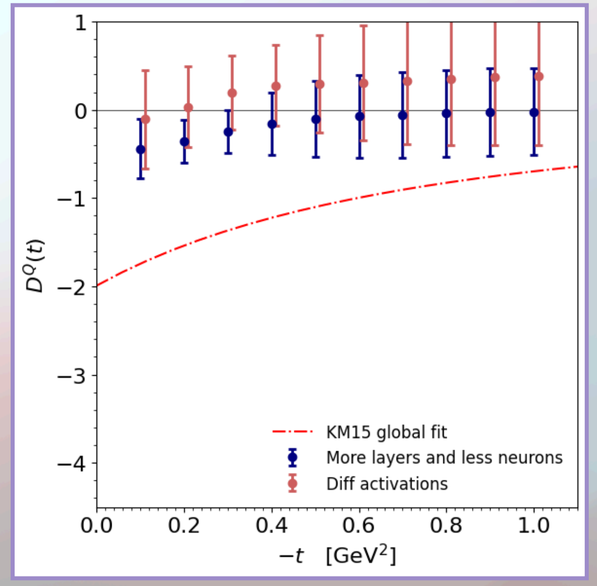
When fitting data:

- Same dataset + same theory → different architectures produce different D(t)
- The fit was found to be very sensitive to changes. Mainly, for ReH and D(t).









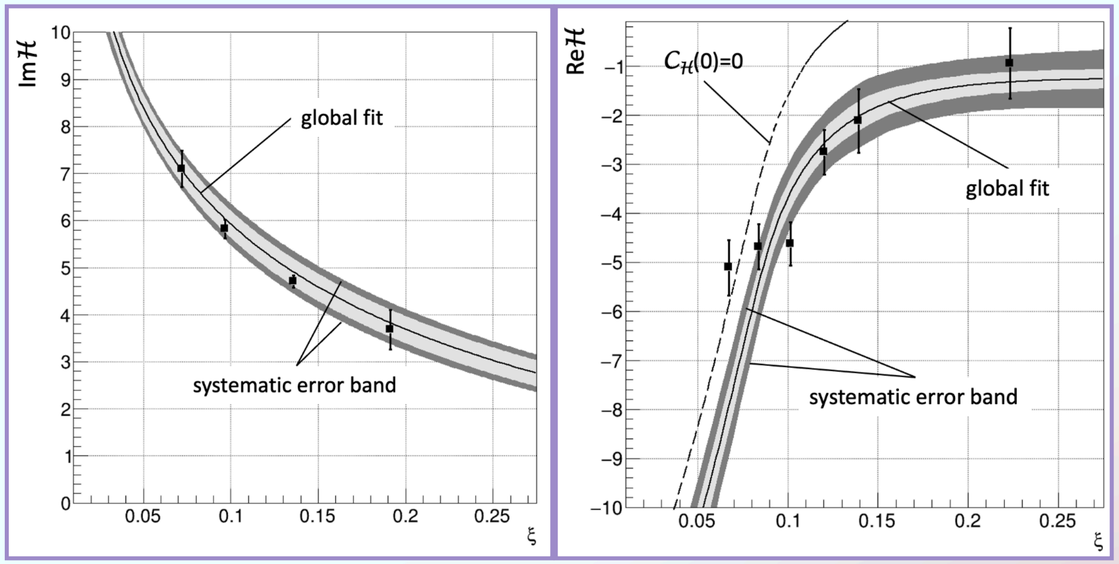
Network architecture analysis on CLAS data

Model stability (architecture dependence)

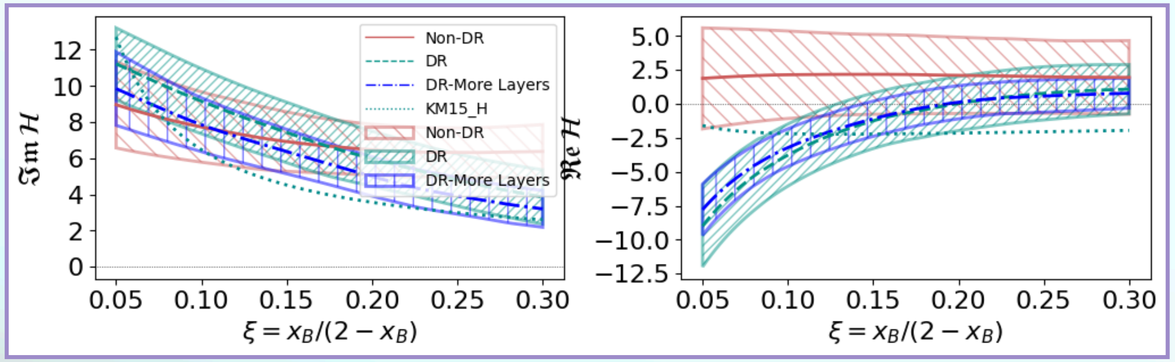
arXiv:2303.08347

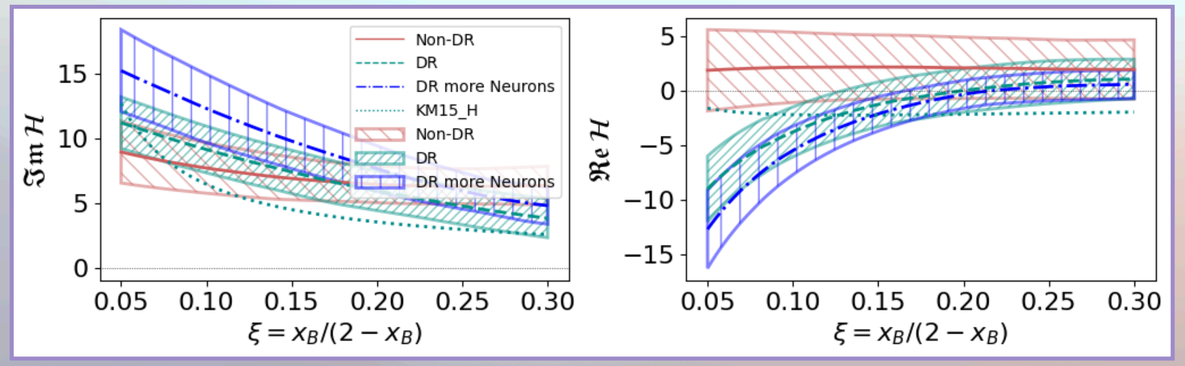
External consistency check using CLAS data

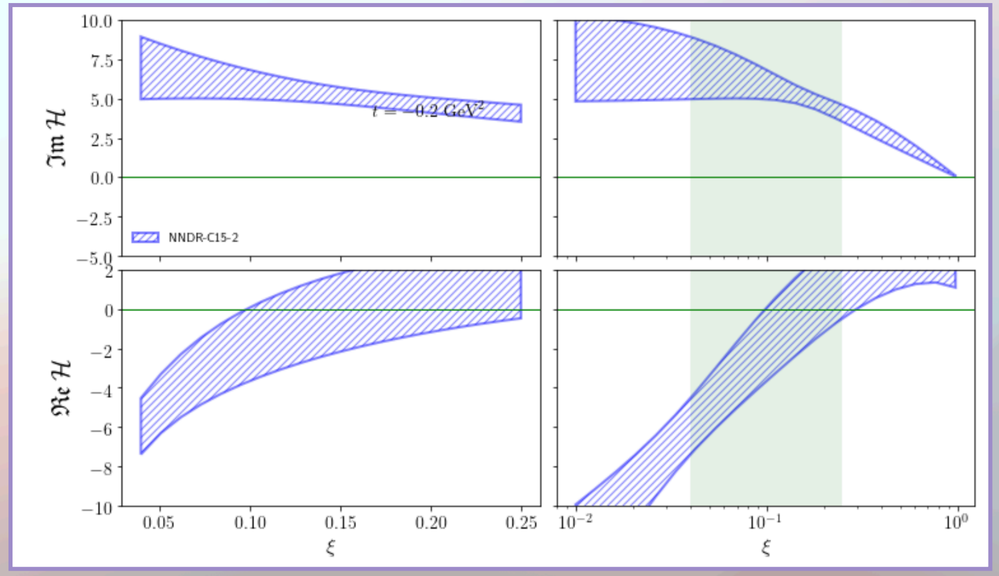
- Architecture choices can shift the bandwidth; we use this as a robustness check.
- Some architectures fall very close, but agreement is not used as the training criterion.



Network architecture analysis on CLAS data







https://github.com/openhep/dterm18/blob/master/dterm.ipynb

Model stability: Closure test

Our objective in the closure test is to generate pseudodata from theoretical models and verify that our neural networks can recover the original CFFs.

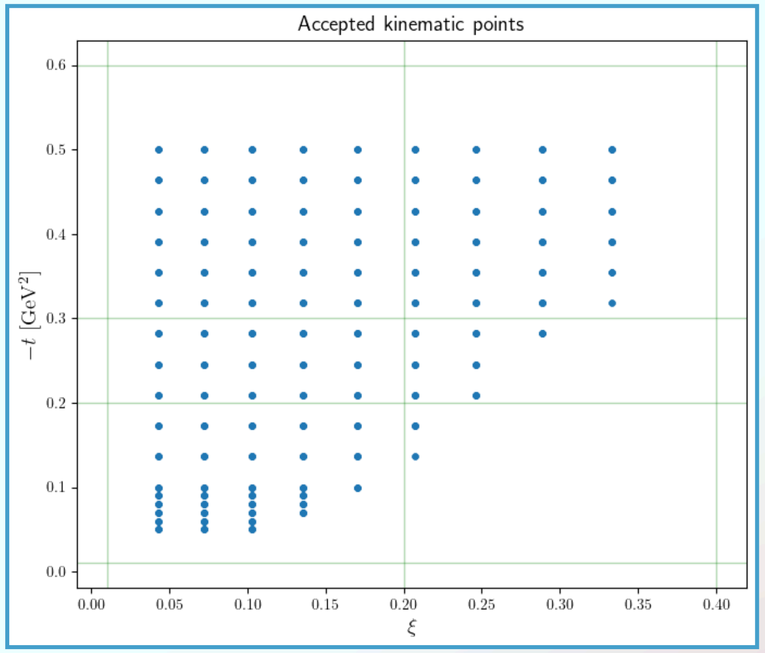
Pseudo-data must lie in the physical DVCS/BH region (BMK propagators must be real)

We fit Fourier coefficients / harmonics → inputs are (xB, Q^2, t).

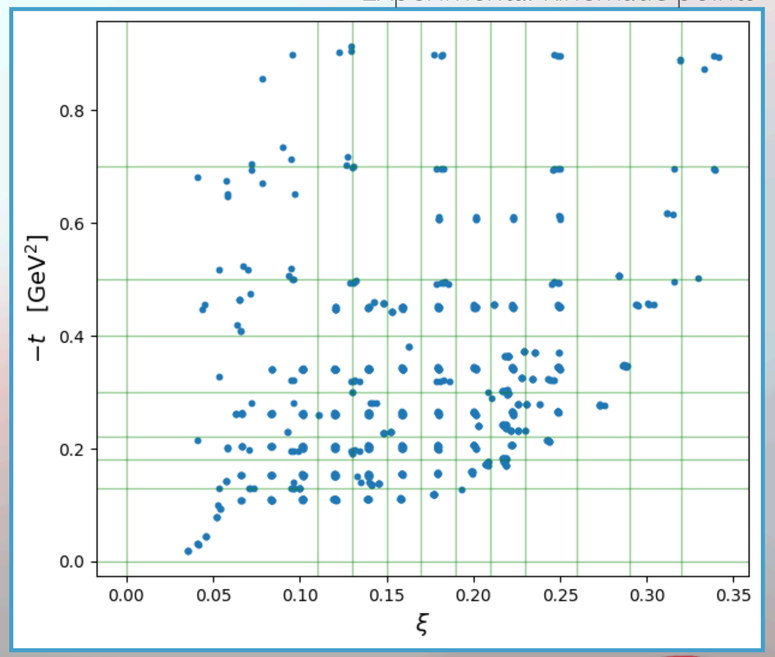
For fixed-target DVCS, not every combination of (xB, Q^2, t) is allowed.

The constraints are:

- $W^2 = M^2 + Q^2(1/x_B - 1) > (M + m_\pi)^2$ (hadronic final state must exist)
- $0 < y < 1$ where $y = \frac{\nu}{E} = \frac{Q^2}{2Mx_BE}$ (lepton energy loss fraction)
- tmin (xB ,Q2) ≤ t ≤ 0 (momentum transfer bounds)



Experimental kinematic points

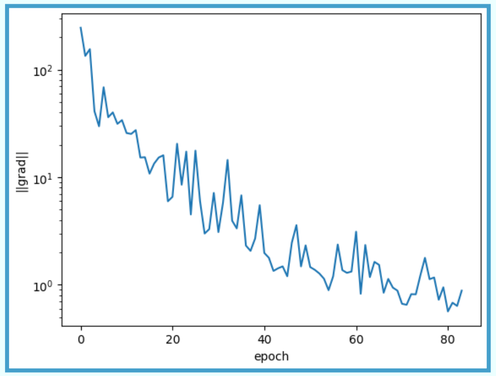


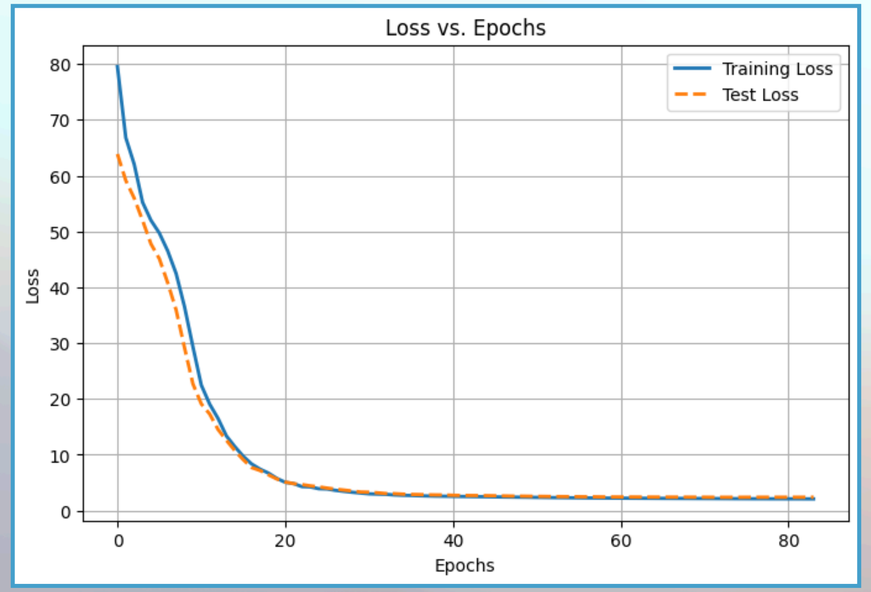
Model stability: Closure test

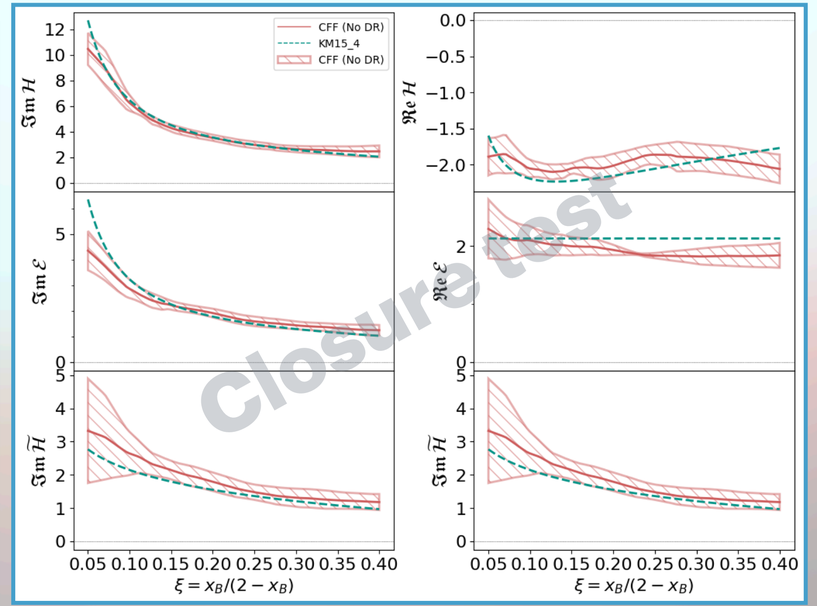
No dispersion relation (non-DR) neural model

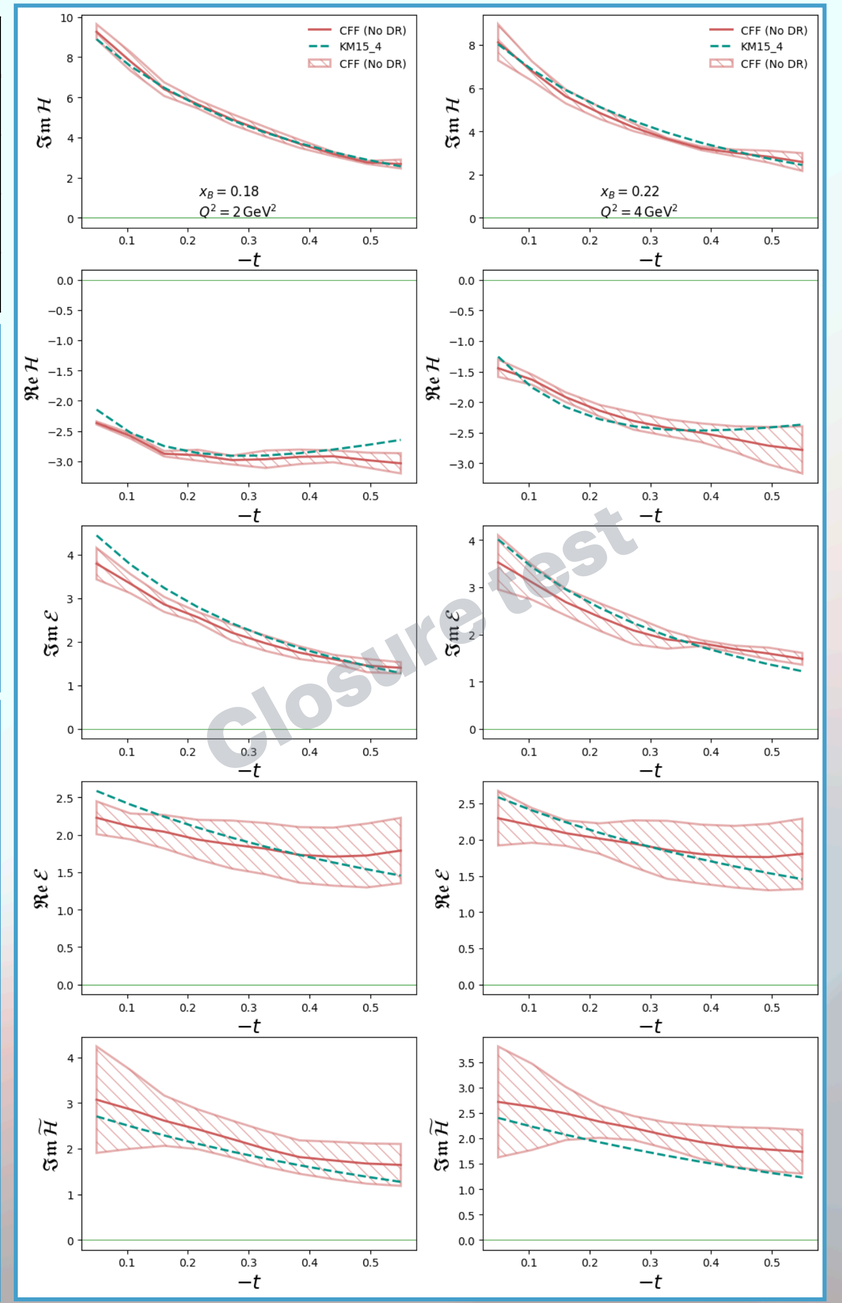
- Pseudo-data generated from KM09
- NN learns CFFs directly → predicts observables with $\chi^2 \sim 1$
- The NN can learn the CFFs that the observables are sensitive to.
- The training pipeline + Gepard theory evaluation + harmonic observable machinery is working.

Obs	χ^2
XLU	1.06
XUU	1.09
AC	1.08
ALU	0.96







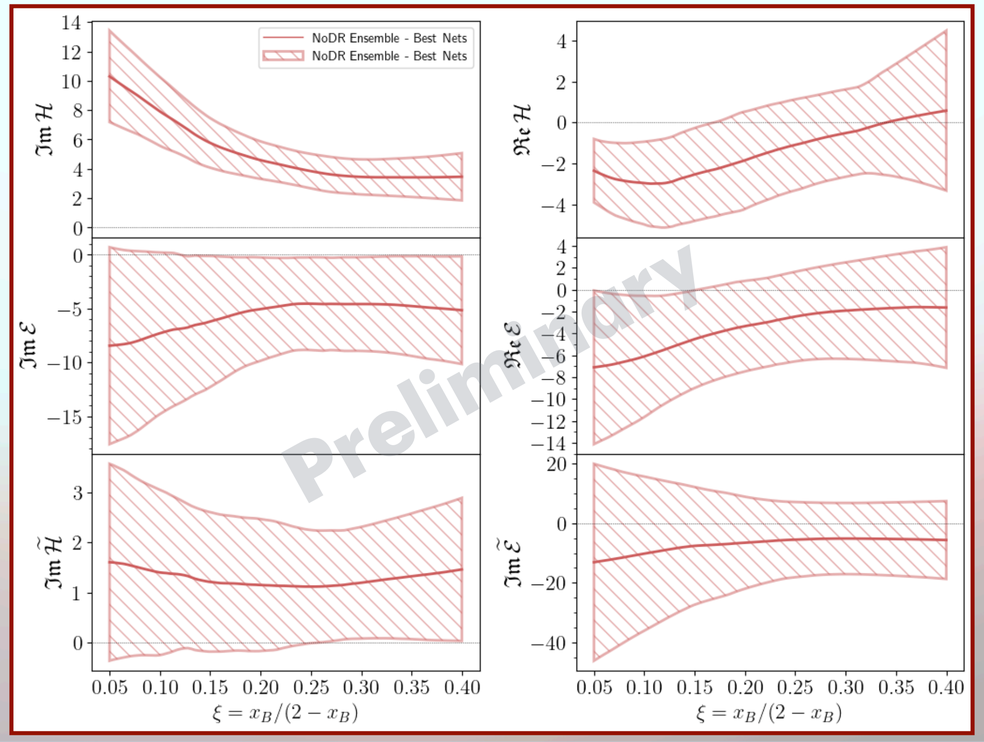


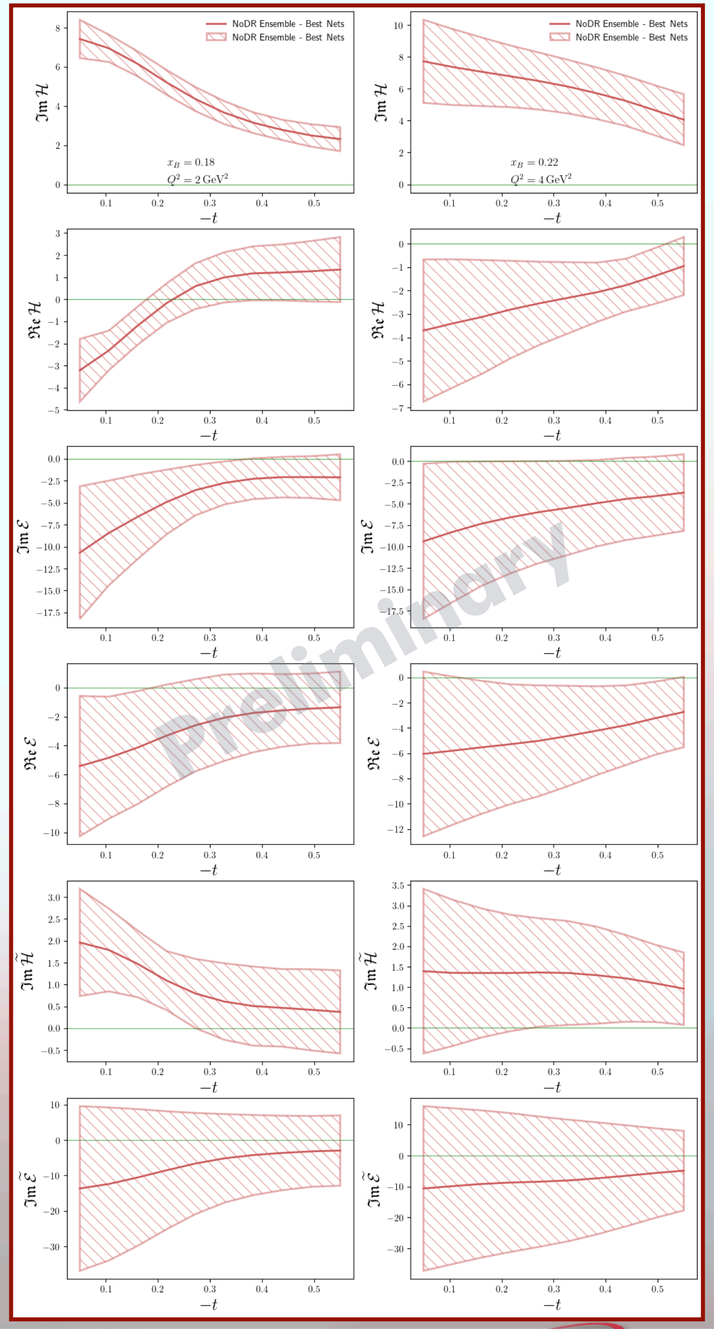
Update on global analysis of CFFs

Dataset: CLAS + Hall A + HERMES (twist-2 safe region cut: $-t/Q^2 < 0.25$)

- Global non-DR fit provides stable ImH and key components. The uncertainties from ensemble spread
- We use an ensemble strategy: multiple nets/seeds → uncertainty band reflects fit flexibility and data constraints.

	observable	collaboration	npts	chi2_ndof
0	ALUI	HERMES	6	4.292
1	XUU	CLAS	161	4.174
2	XLUw	HALLA	17	3.879
3	XLU	CLAS	103	3.192
4	ALU	CLAS	10	3.103
5	BTSA	HERMES	4	2.395
6	XUUw	CLAS	320	2.308
7	ALU	CLAS	48	2.243
8	XLUw	HALLA	30	1.345
9	ALU	CLAS	12	1.250
10	ALU	CLAS	10	1.181
11	BTSA	CLAS	20	1.048
12	XUUw	CLAS	96	0.987
13	XUUw	HALLA	20	0.955
14	AUTI	HERMES	4	0.808
15	XUUw	HALLA	44	0.734
16	XLUw	CLAS	48	0.720
17	AC	HERMES	12	0.579
18	TSA	CLAS	10	0.315
19	TSA	HERMES	4	0.285





Closure test: Dispersion relation (DR) neural model

We implemented a DR-constrained neural model in Gepard

$$\mathrm{Re}\mathcal{H}(\xi,t) = \mathcal{C}_{\mathcal{H}}(t) + \frac{1}{\pi}\,\mathrm{P.V.}\int_0^1 \mathrm{d}\xi' \left[\frac{1}{\xi-\xi'} - \frac{1}{\xi+\xi'}\right] \mathrm{Im}\mathcal{H}(\xi',t)$$

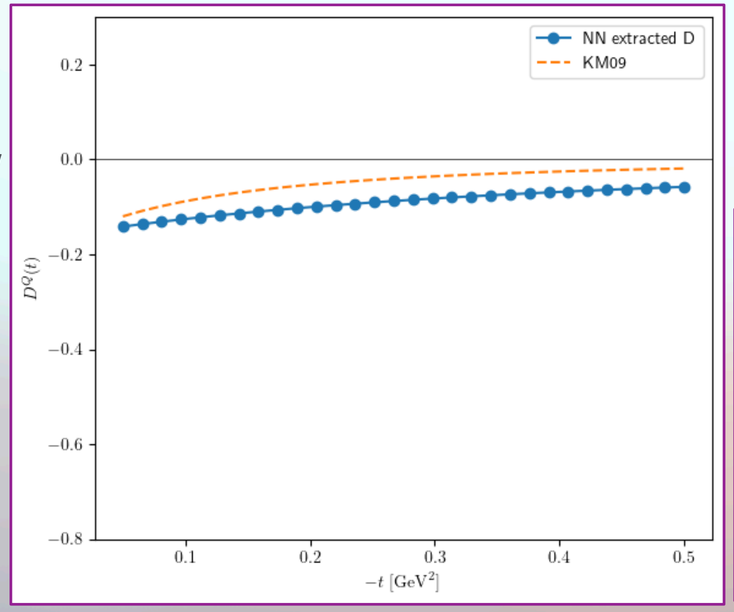
- PV integral + small-ξ correct extrapolation controls ReH stability.
- D-term is still weakly constrained to be left free unless ReH-sensitive observables are ensured.

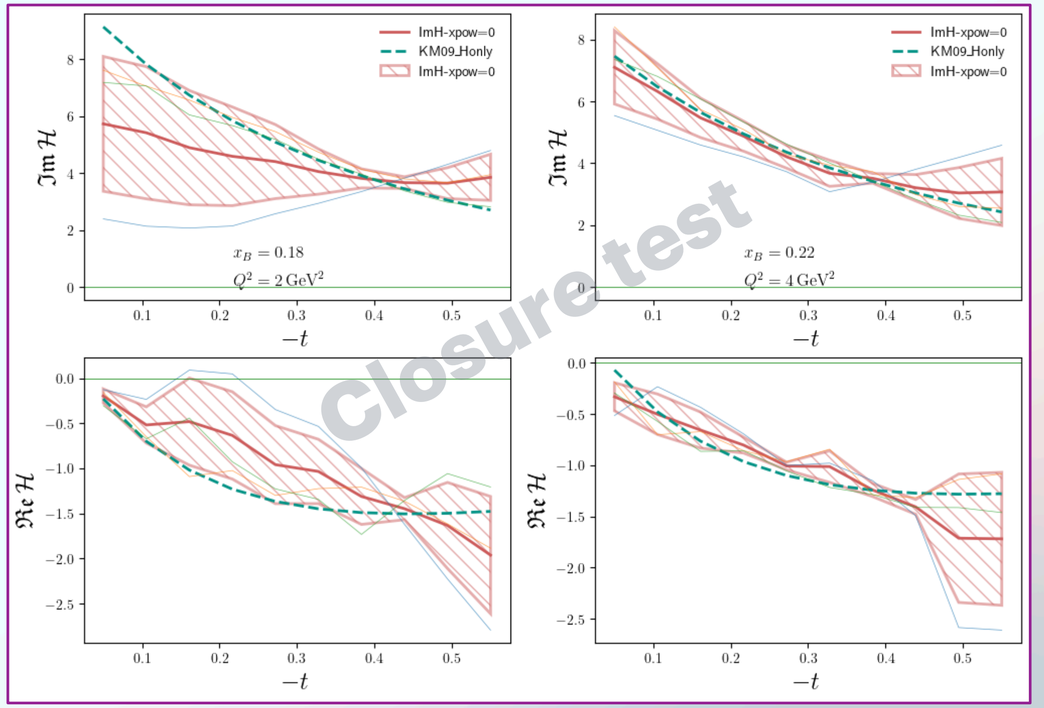
$$D(t) = \frac{D_0}{(1-\frac{t}{M^2})^{\alpha}}$$

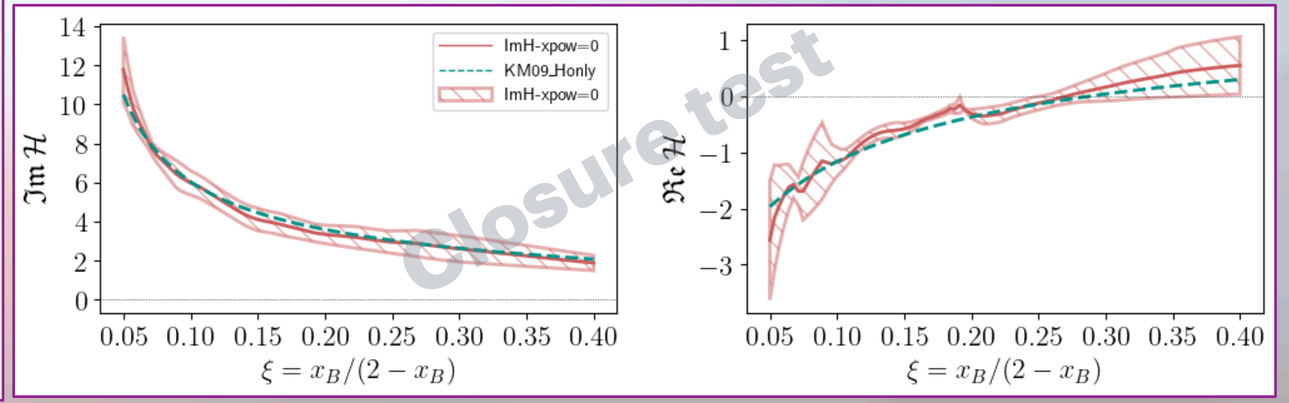
$D_0 = 0.2233 \quad \alpha = 2.1407$

$M^2 = 0.8171$

Obs	χ^2
XLU	5.5
XUU	2.4
ALU	15.95
AUT,DVCS	4.5







Dispersion relation and D-term extraction

We implemented a DR-constrained neural model in Gepard

$$\mathrm{Re}\mathcal{H}(\xi,t) = \mathcal{C}_{\mathcal{H}}(t) + \frac{1}{\pi}\,\mathrm{P.V.}\int_0^1 \mathrm{d}\xi' \left[\frac{1}{\xi-\xi'} - \frac{1}{\xi+\xi'}\right] \mathrm{Im}\mathcal{H}(\xi',t)$$

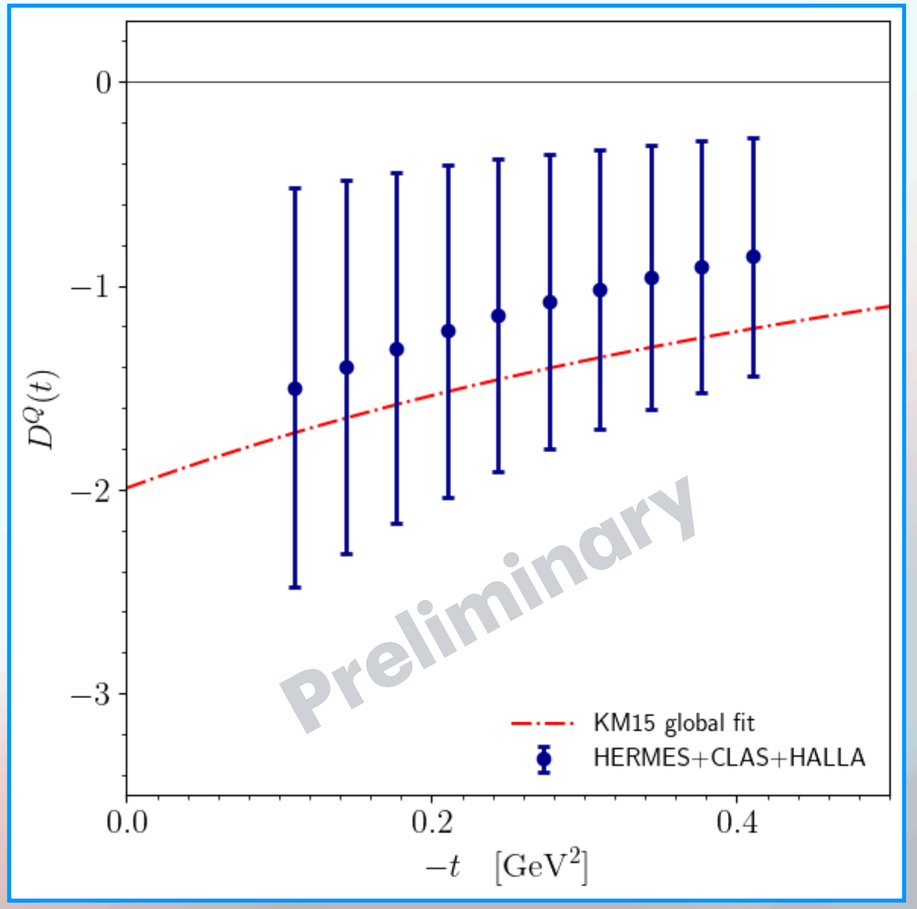
- PV integral + small-ξ correct extrapolation controls ReH stability.
- D-term is still weakly constrained to be left free unless ReH-sensitive observables are ensured.

$$D(t) = \frac{D_0}{(1-\frac{t}{M^2})^{\alpha}}$$

$D_0 = -1.2865 \quad \alpha = 2.2333$

$M^2 = 0.7828$

- The preliminary global fit shows a preference for KM15-like behavior over KM09-like behavior.



Deliverables and future directions

Deliverables and future directions

- Roadmap for a Global analysis: Extract harmonics for each observable, remove outliers and problematic data points, perform preliminary fit adjustments, tune hyperparameters for optimal performance, conduct ensemble analyses from several NN models and make closure tests to the created NN models.

Deliverables:

- Global fit with consistent Q^2 treatment is necessary to treat all data available.
- The DR reconstruction is extremely sensitive to the behavior of ImH outside the region directly constrained by the observables — especially the small-ξ tail and the neighborhood of the principal-value point.
- We now have a validated DR-NN pipeline, and we can study D-term sensitivity in a controlled way.
- The next step is quantifying what combination of CFFs extraction and kinematic coverage makes D(t) identifiable and stable, and what priors are required.

Future directions:

- Global fit including HERA data: Analyze the behavior change due to kinematics.
- Finish global analysis for the DR neural model.
- Create a new module on Gepard for DDVCS.
- Generate DDVCS/EIC pseudodata using EPIC and integrate into the analysis.

Thank you all for your attention

Backup

Database

opengpd / gpddatabase

Currently, there was no unified common database storing all available data which is necessary to reduce the uncertainties.

We developed a publicly accessible database format capable of storing current and future data as well as pseudodata.

- Stores data and pseudodata coming from various elastic and exclusive measurements and lattice-QCD results.
- Interfaces are provided in the main programming languages: Python and C++.
- Interface with GEPARD and Partons.
- Accessible by anyone in the world using Github.

Database for studying Generalised Parton Distributions (GPDs)

This page:
about
installation
data format
basic usage
available data
adding new data

license
acknowledgements
contact

View GitHub repository:
link

View PyPI repository:
link

View reference article:
arXiv inspire

Available datasets

uuid	collaboration	reference	type	pseudo	observables	comment
EqbtDRkv	HallA	https://arxiv.org/pdf/1703.09442.pdf	DVCS	False	CrossSectionDifferenceLU	None
PusMstKs	CLAS	https://arxiv.org/pdf/2211.11274	DVCS	False	ALL	None
vGAKAf7P	CLAS	https://arxiv.org/pdf/1501.07052.pdf	DVCS	False	ALU AUL ALL	None
75ueQoQw	COMPASS	https://arxiv.org/abs/1802.02739	DVCS	False	CrossSectionUUVirtualPhotoProduction TSlope	None
ob8hLTm2	CLAS	https://arxiv.org/pdf/1810.02110	DVCS	False	CrossSectionUU	Differential CrossSection for exclusive DVCS electroproduction
3gYp9R4P	HERMES	https://arxiv.org/pdf/hep-ex/0605108.pdf	DVCS	False	AcCos1Phi	None
bmTzHHvg	HallA	https://arxiv.org/pdf/2109.02076	DVCS	False	CrossSectionUU	Beam-helicity-independent cross section
RQncbKtk	CLAS	https://arxiv.org/pdf/1504.02009v1.pdf	DVCS	False	CrossSectionUU CrossSectionDifferenceLU	None
AtY8o7Ej	HallA	https://arxiv.org/pdf/1504.05453v1.pdf	DVCS	False	CrossSectionUU CrossSectionDifferenceLU	None

In collaboration with P. Sznajder and C. Mezrag

arxiv.org/abs/2503.18152

https://opengpd.github.io/gpddatabase/index.html

Database

opengpd / gpddatabase

Currently, there was no unified common database storing all available data which is necessary to reduce the uncertainties.

We developed a publicly accessible database format capable of storing current and future data as well as pseudodata.

- Get the list of available data files.
- Load a given data file
- Get metadata.
- Upload your own data file for the use of the GPD community.

Database for studying Generalised Parton Distributions (GPDs)

This page:
about
installation
data format
basic usage
available data
adding new data

license
acknowledgements
contact

View GitHub repository:
link

View PyPI repository:
link

View reference article:
arXiv inspire

Available datasets

uuid	collaboration	reference	type	pseudo	observables	comment
EqbtDRkv	HallA	https://arxiv.org/pdf/1703.09442.pdf	DVCS	False	CrossSectionDifferenceLU	None
PusMstKs	CLAS	https://arxiv.org/pdf/2211.11274	DVCS	False	ALL	None
vGAKAf7P	CLAS	https://arxiv.org/pdf/1501.07052.pdf	DVCS	False	ALU AUL ALL	None
75ueQoQw	COMPASS	https://arxiv.org/abs/1802.02739	DVCS	False	CrossSectionUUVirtualPhotoProduction TSlope	None
ob8hLTm2	CLAS	https://arxiv.org/pdf/1810.02110	DVCS	False	CrossSectionUU	Differential CrossSection for exclusive DVCS electroproduction
3gYp9R4P	HERMES	https://arxiv.org/pdf/hep-ex/0605108.pdf	DVCS	False	AcCos1Phi	None
bmTzHHvg	HallA	https://arxiv.org/pdf/2109.02076	DVCS	False	CrossSectionUU	Beam-helicity-independent cross section
RQncbKtk	CLAS	https://arxiv.org/pdf/1504.02009v1.pdf	DVCS	False	CrossSectionUU CrossSectionDifferenceLU	None
AtY8o7Ej	HallA	https://arxiv.org/pdf/1504.05453v1.pdf	DVCS	False	CrossSectionUU CrossSectionDifferenceLU	None

In collaboration with P. Sznajder and C. Mezrag

arxiv.org/abs/2503.18152

https://opengpd.github.io/gpddatabase/index.html

Accessing GFFs

Considering some assumptions like:

1. At LO, gluons don't contribute to DVCS much, so one can assume that the subtraction constant is dominantly coming from quarks, but for the moment, we also neglect strange and heavier quarks.

2. To talk about "quark pressure," one needs to neglect energy-momentum flow from quarks to gluons described by the EMT form factor $\tilde{C}(t)$.

2. In the asymptotic limit $d_n^q(t)$ vanish for n>1. Where $d_1^q(t)$ determine the asymptotic form of the GPDs.

3. Considerate the dominance of the flavor singlet combination.

$$\Delta(t) = 2\sum_q e_q^2 \int_{-1}^{1} \mathrm{d}z \, \frac{D^q_{\rm term}(z,t)}{1-z},$$

$$d_1^u = d_1^d = d_1^Q/2$$

$$D^q_{\rm term}(z,t) = (1-z^2)\sum_{\rm odd\, n} d_n^q(t)\, C_n^{3/2}(z)$$

$$D_q(t) = \frac{4}{5}\, d_1^q(t) = \int_{-1}^{1} \mathrm{d}z\, z\, D^q_{\rm term}(z,t)$$

$$\Delta(t) = 4\left(\frac{4}{9} d_1^u(t) + \frac{1}{9} d_1^d(t)\right)$$

$$D^Q(t) = \frac{18}{25}\Delta(t)\,.$$

To date, a more conservative extraction of the subtraction constant from the currently available experimental data remains compatible with zero within large uncertainties.

By utilizing this DR, along with other phenomenological assumptions, we can extract the CFFs and GFFs from experimental data and enhance the results with lattice calculations!

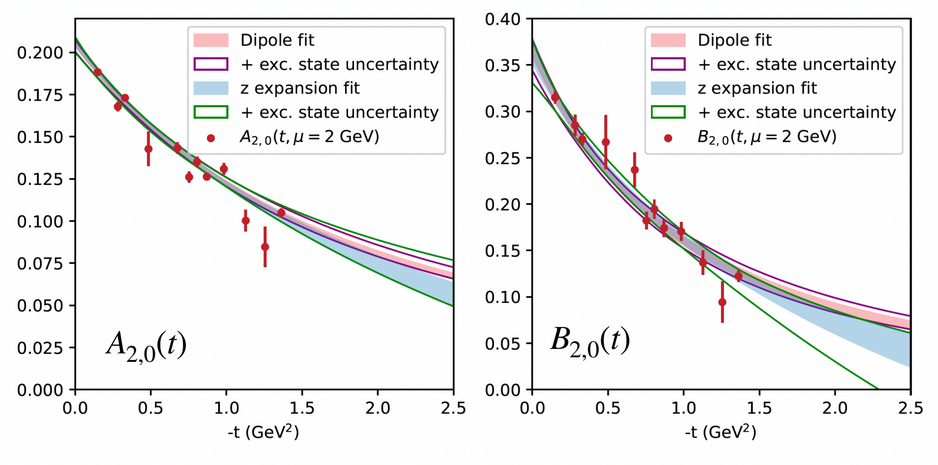
State of the project: Inputs from Lattice QCD

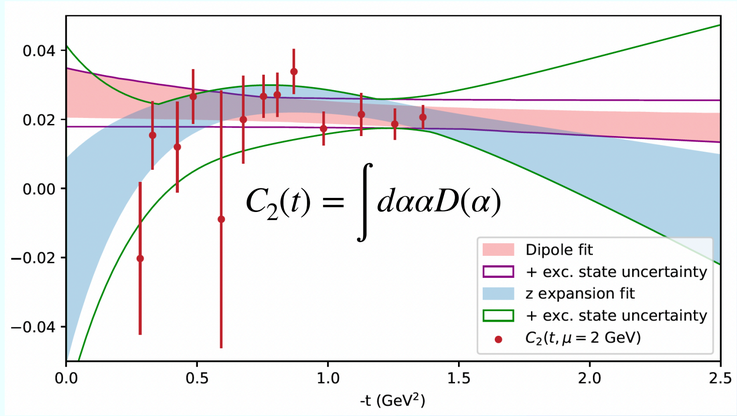
$$D^q_{\text{term}}(z,t) = (1-z^2)\sum_{\text{odd } n} d^q_n(t)\, C^{3/2}_n(z)$$

Lattice QCD (LQCD) constraints can help reduce the large systematic uncertainties in GFF extractions by accounting for higher-order terms in Gegenbauer expansions.

Extraction of the isovector GFFs from Lattice QCD

I.J. High Energy. Phys. 2024, 162 (2024)





- Calculation of the non-singlet "$u - d$" combination of GFF, which cannot be probed through DVCS, has been performed in Lattice across a wide range of t.
- To fully constrain the analysis, isoscalar GFFs (connected contributions) are also needed.

How can we use this data in a Global analysis?

- Lattice calculations can provide direct information about CFF.
- Plan to merge LQCD-derived CFFs into the neural network fitting framework, complementing experimental data and improving constraints in the loss function.

Preliminary fits

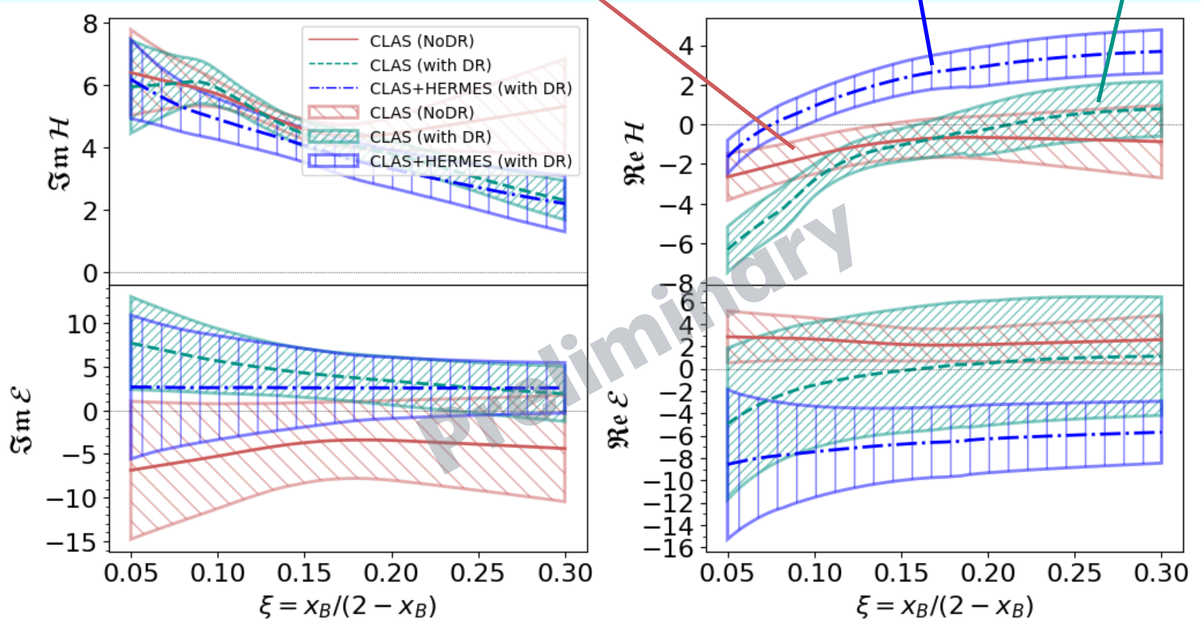
Compton form factors:

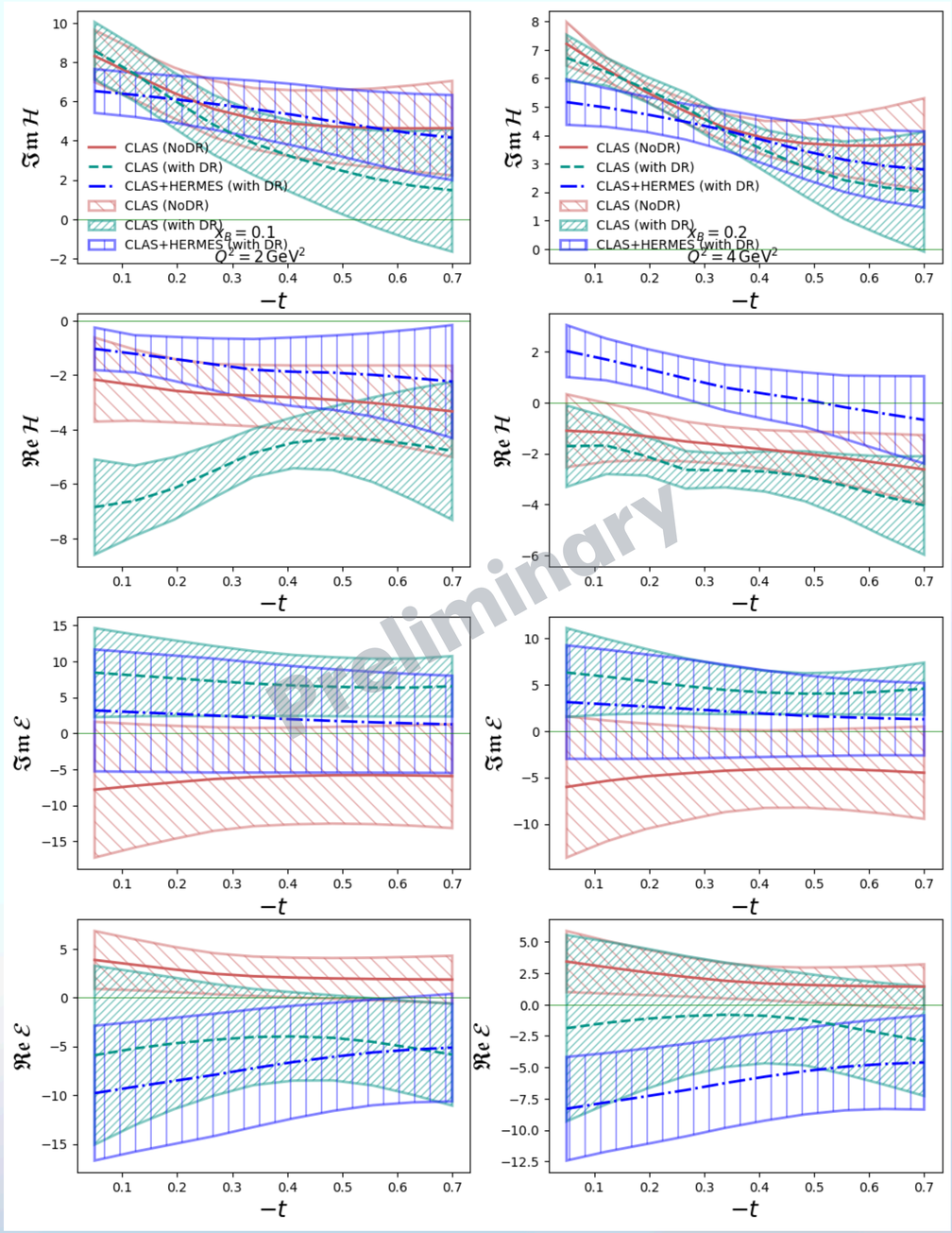
4 CFFs

CLAS with NO DR

CLAS + HERMES with DR

CLAS with DR





CLAS BSA at ~11 GeV data (Phys. Rev. Lett. 130, 211902) raise the range of kinematic restrictions

$\chi^2/dof = 1.34$

Planning to add the upcoming CLAS cross-section data.