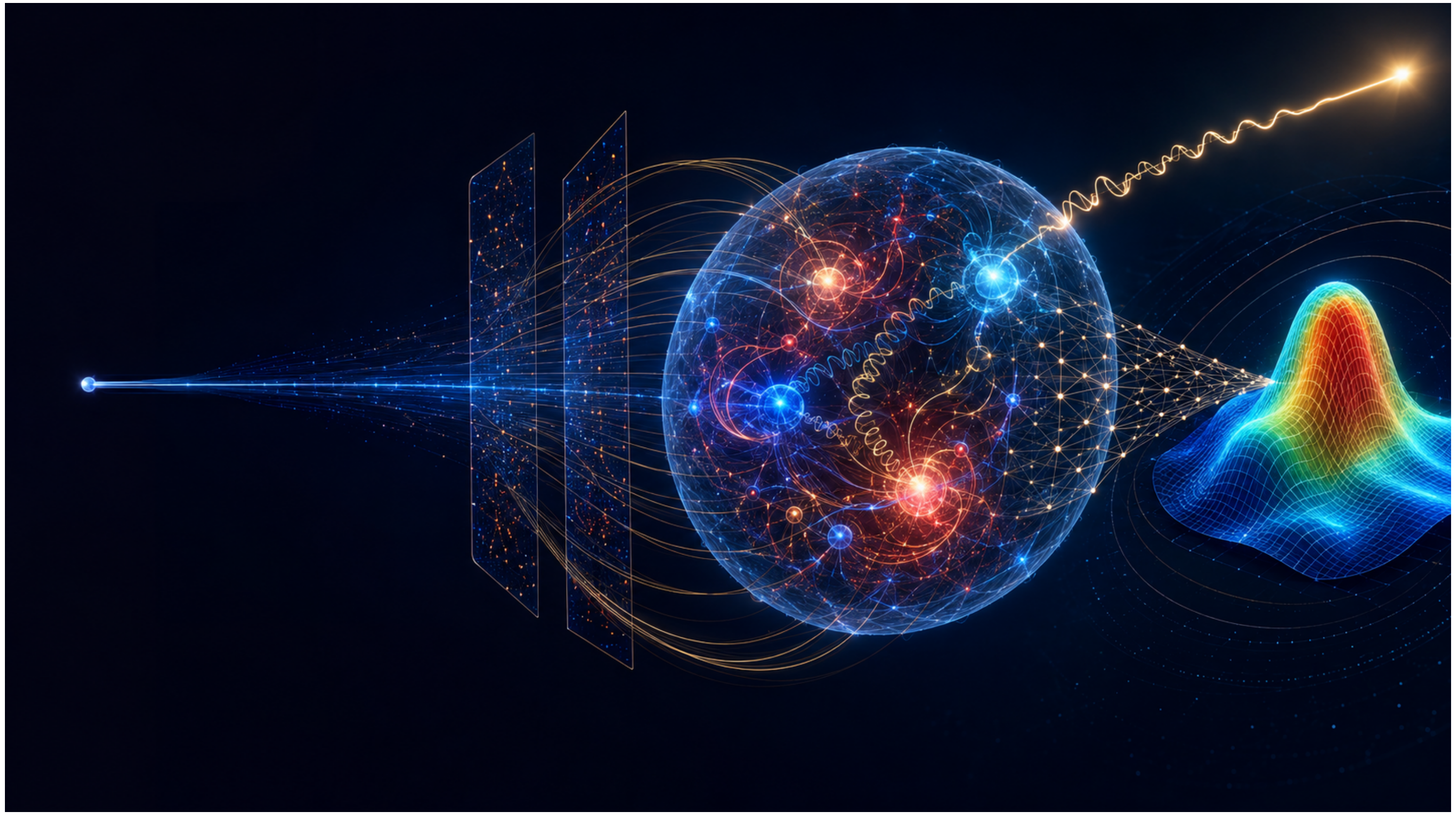
Neural Network Representations of Generalized Parton Distributions

arXiv:2605.06994



Gia-Wei Chern

Towards improved hadron femtography with hard exclusive reactions

July 23 2026, University of Virginia

Neural Network Representation of Generalized Parton Distributions (NNGPD)

Jitao Xu,[1,*] Ho Jang,[2,*] Zaki Panjsheeri,[2,*] Gia-Wei Chern,[2,†] Yaohang Li,[1,‡]
Simonetta Liuti,[2,§] Douglas Q. Adams,[2] Michael Engelhardt,[3] Gary R. Goldstein,[4]
Adil Khawaja,[2] Huey-Wen Lin,[5] Saraswati Pandey,[2] and Kemal Tezgin[6]

[1] *CS Department, Old Dominion University, Norfolk, VA 22904, USA.*
[2] *Department of Physics, University of Virginia, Charlottesville, VA 22904, USA.*
[3] *Department of Physics, New Mexico State University, Las Cruces, NM 88003, USA*
[4] *Department of Phyiscs, Tufts University, Medford, MA 02155*
[5] *Department of Physics and Astronomy, Michigan State University, East Lansing, MI , 48824 USA.*
[6] *Department of Physics, Virginia Tech, Blacksburg, VA 24061, USA.*


Jitao Xu


Ho Jang


Zaki Panjsheeri

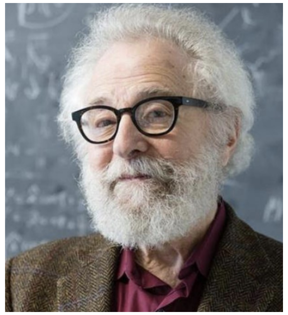

Adil Khawaja


Doug Adams


Saraswati Pandey


Kemal Tezgin


Michael Engelhardt


Gary Goldstein

Huey-Wen Lin


Yaohang Li


Simonetta Liuti

GPD extraction contains two nested inverse problems

The forward theory is known; the information is progressively compressed.

HIDDEN FUNCTION

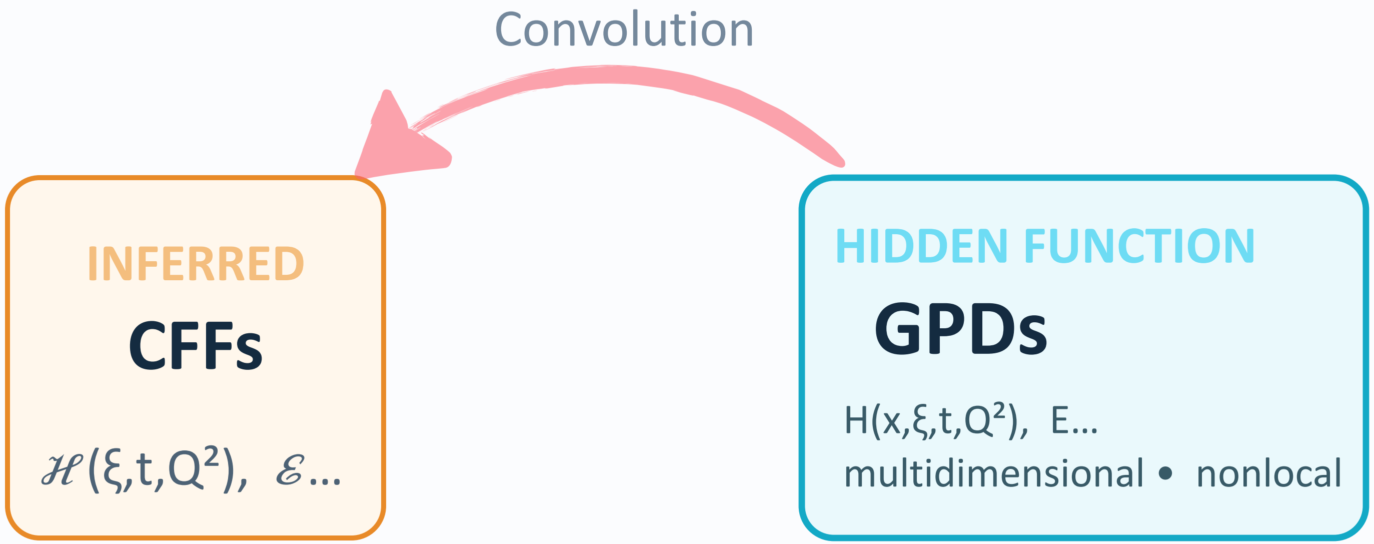
GPDs

$H(x,\xi,t,Q^2)$, E…

multidimensional • nonlocal

GPD extraction contains two nested inverse problems

The forward theory is known; the information is progressively compressed.



THE EXTRACTION PROBLEM

GPD extraction contains two nested inverse problems

The forward theory is known; the information is progressively compressed.

Cross section

Convolution

MEASURED

DVCS observables

cross sections • asymmetries
finite kinematic coverage

INFERRED

CFFs

$\mathscr{H}(\xi,t,Q^2)$, $\mathscr{E}$…

HIDDEN FUNCTION

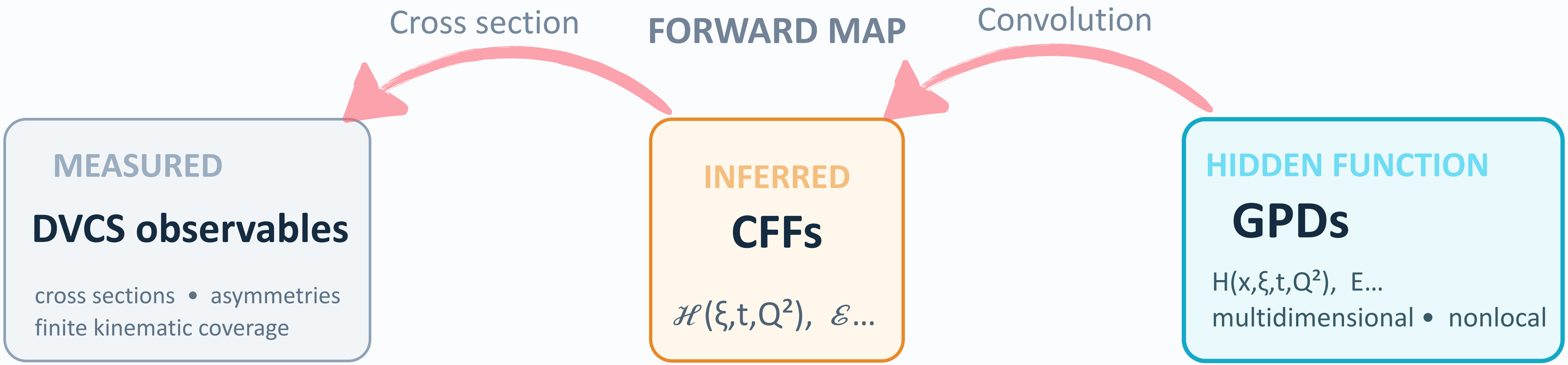
GPDs

$H(x,\xi,t,Q^2)$, E…
multidimensional • nonlocal

THE EXTRACTION PROBLEM

GPD extraction contains two nested inverse problems

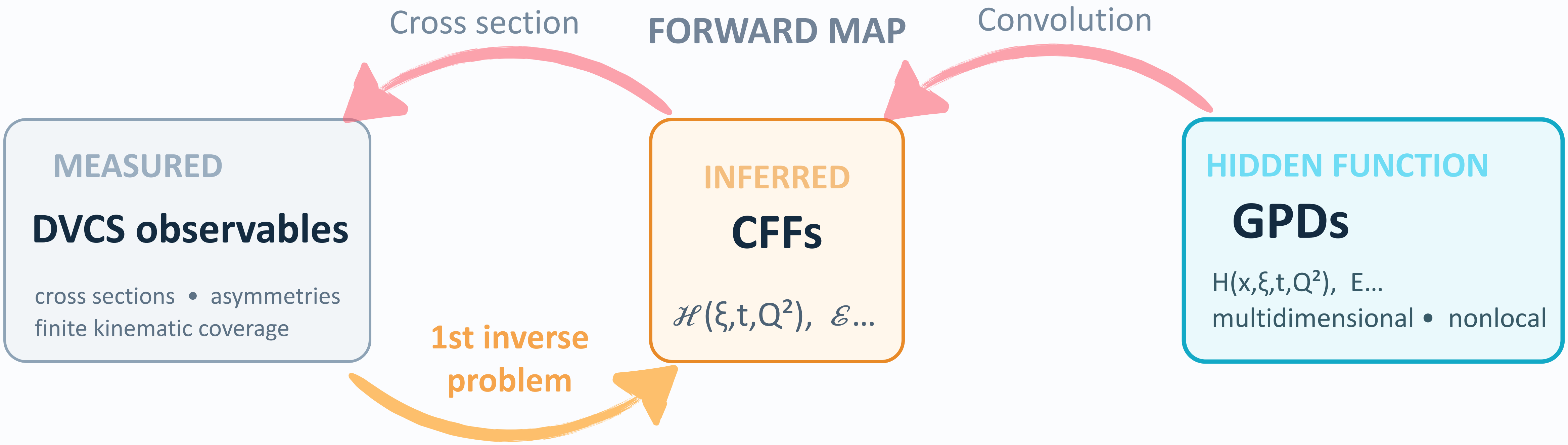
The forward theory is known; the information is progressively compressed.



THE EXTRACTION PROBLEM

GPD extraction contains two nested inverse problems

The forward theory is known; the information is progressively compressed.

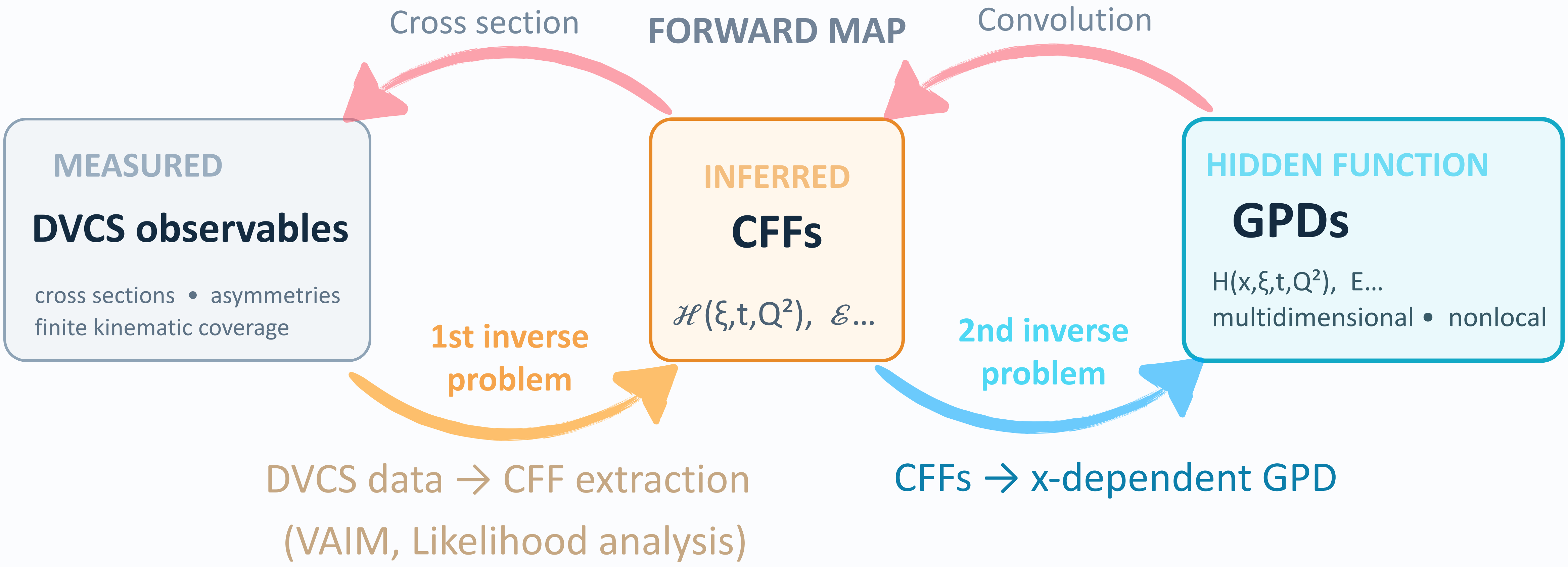


DVCS data → CFF extraction
(VAIM, Likelihood analysis)

THE EXTRACTION PROBLEM

GPD extraction contains two nested inverse problems

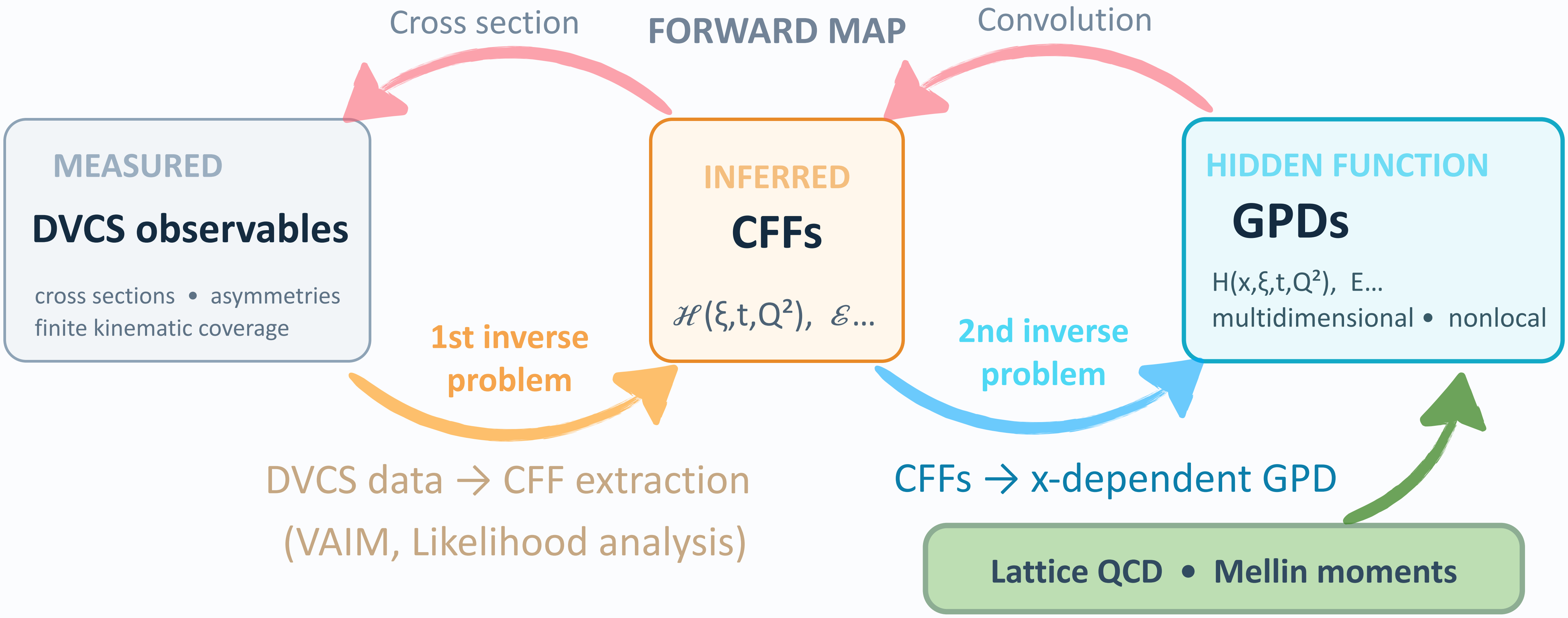
The forward theory is known; the information is progressively compressed.



THE EXTRACTION PROBLEM

GPD extraction contains two nested inverse problems

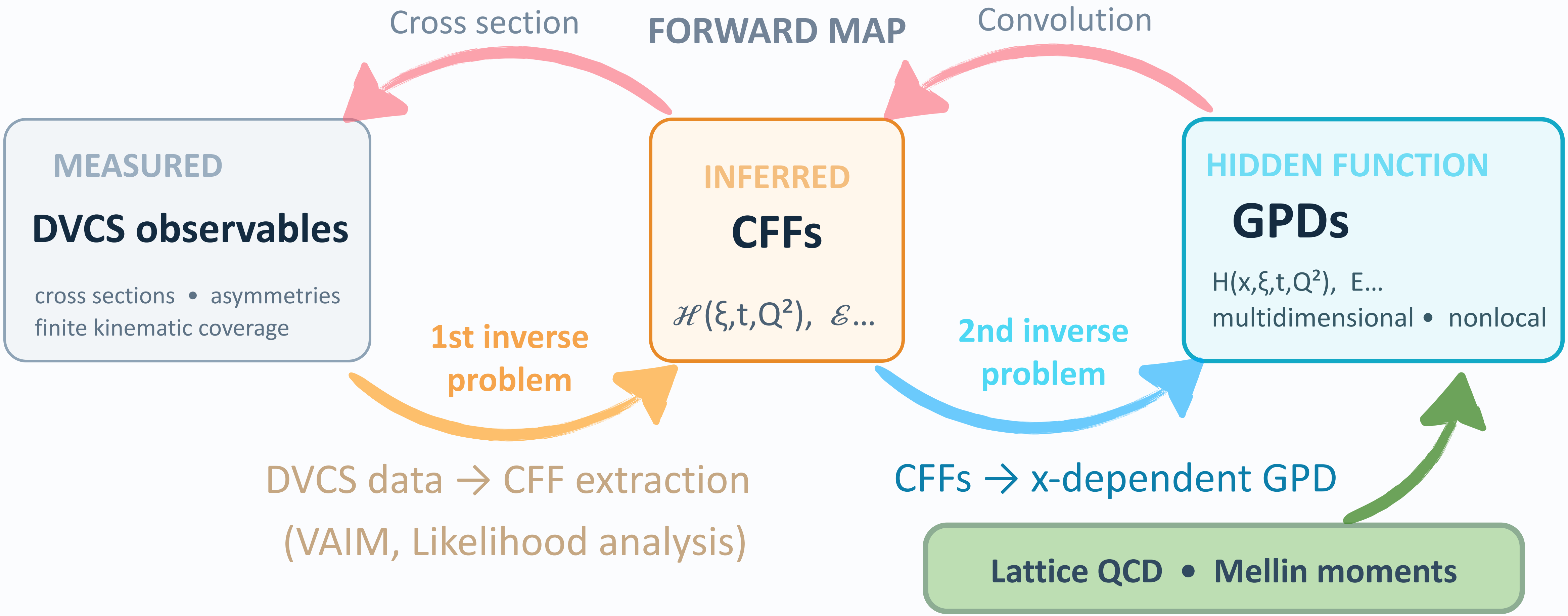
The forward theory is known; the information is progressively compressed.



THE EXTRACTION PROBLEM

GPD extraction contains two nested inverse problems

The forward theory is known; the information is progressively compressed.



NNGPD targets the second inversion: reconstruct the function from complementary integral projections—without pointwise GPD labels.

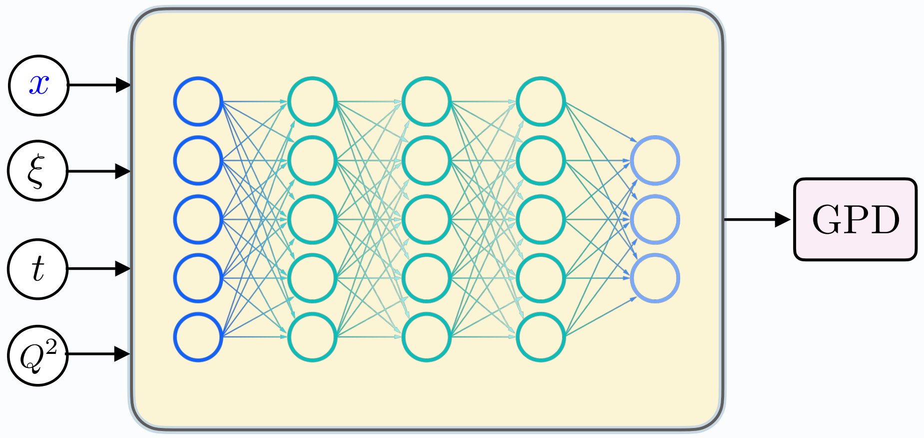
Learn the GPD—not a prescribed functional form

Represent the unknown multidimensional function directly, then require it to satisfy everything physics already tells us.

FLEXIBLE REPRESENTATION

A feed-forward network represents the GPD

$$H(x,\xi,t,Q^2),\ E(x,\xi,t,Q^2),\ \cdots$$



PHYSICS DEFINES ADMISSIBILITY

The learned function must...

1. **Reproduce complex CFFs**
   through the known nonlocal convolution
2. **Match Mellin moments**
   including form-factor and lattice-QCD information
3. **Obey exact symmetries**
   linking the positive- and negative-x regions
4. **Respect endpoint behavior**
   at the kinematic boundaries of the GPD

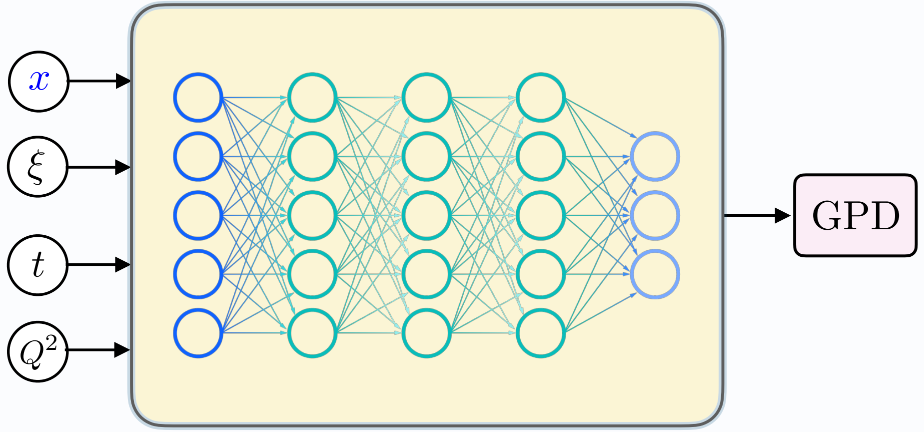
Learn the GPD—not a prescribed functional form

Represent the unknown multidimensional function directly, then require it to satisfy everything physics already tells us.

FLEXIBLE REPRESENTATION

A feed-forward network represents the GPD

$$H(x,\xi,t,Q^2),\ E(x,\xi,t,Q^2),\ \cdots$$



PHYSICS DEFINES ADMISSIBILITY

The learned function must…

1. **Reproduce complex CFFs**
   through the known nonlocal convolution
2. **Match Mellin moments**
   including form-factor and lattice-QCD information
3. **Obey exact symmetries**
   linking the positive- and negative-x regions
4. **Respect endpoint behavior**
   at the kinematic boundaries of the GPD

NOT AN ANSATZ FIT

No fixed x-, ξ-, or t-dependence is assumed.

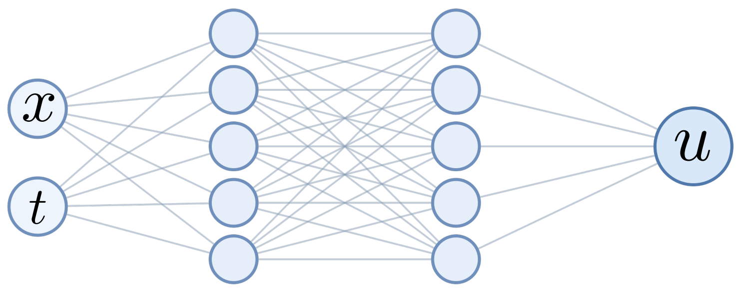
Physics enters through operators and constraints —not through a chosen functional shape.

Physics Informed Neural Networks (PINNs): learning through differential constraints

$$\partial_t u = \mathcal{N}[u]$$

A neural approximation is trained by minimizing differential-equation residuals, together with initial and boundary conditions.

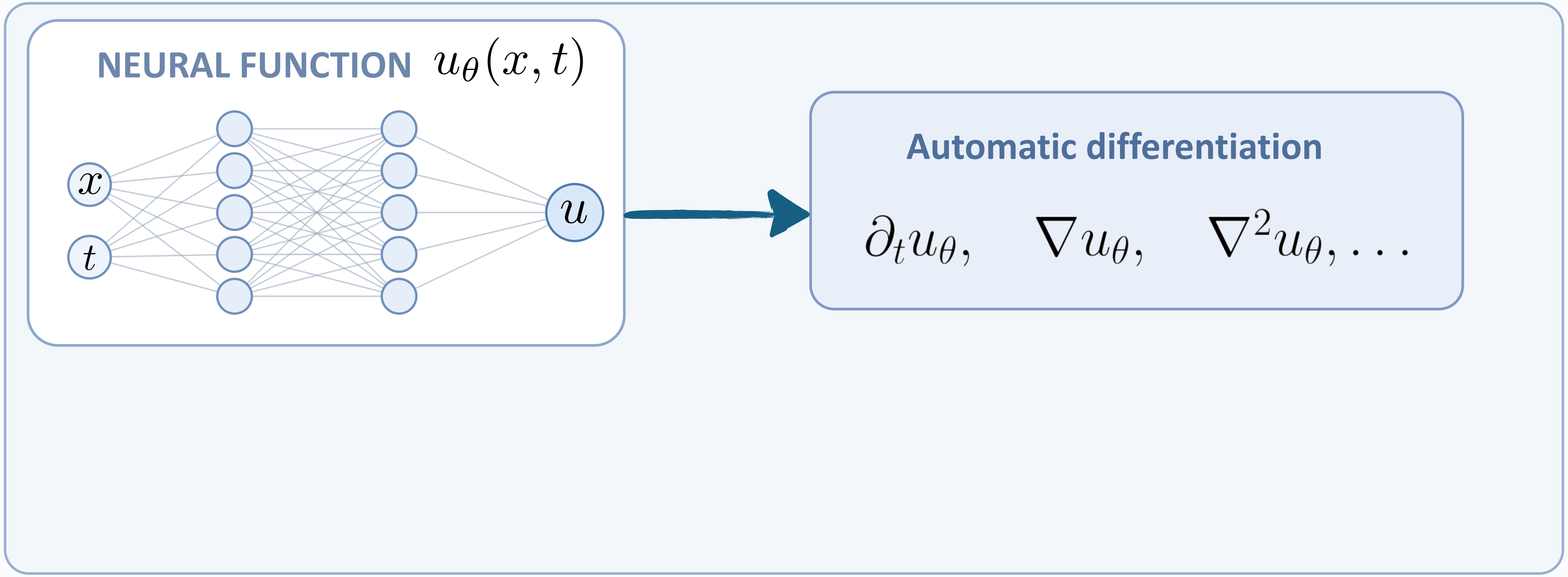
NEURAL FUNCTION $u_\theta(x,t)$



Physics Informed Neural Networks (PINNs): learning through differential constraints

$$\partial_t u = \mathcal{N}[u]$$

A neural approximation is trained by minimizing differential-equation residuals, together with initial and boundary conditions.

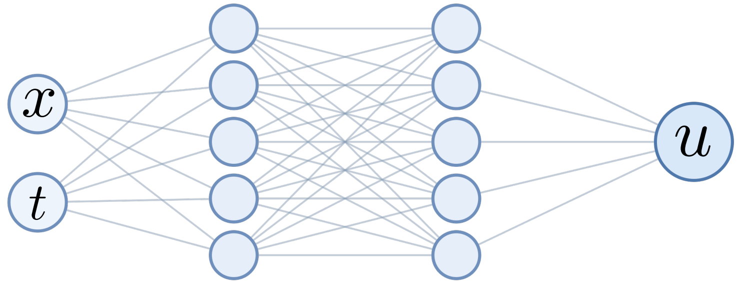


Physics Informed Neural Networks (PINNs): learning through differential constraints

$$\partial_t u = \mathcal{N}[u]$$

A neural approximation is trained by minimizing differential-equation residuals, together with initial and boundary conditions.

NEURAL FUNCTION $u_\theta(x, t)$



Automatic differentiation

$$\partial_t u_\theta, \quad \nabla u_\theta, \quad \nabla^2 u_\theta, \ldots$$

PDE residual $R_\theta = \partial_t u_\theta + \mathcal{N}[u_\theta]$

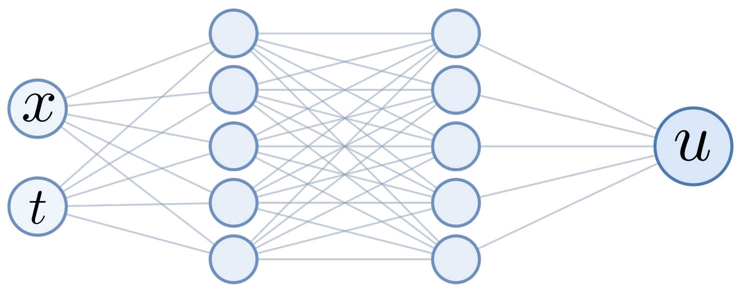
$$\boxed{\mathcal{L}_{\text{PINN}} = ||R_\theta||^2 + \mathcal{L}_{\text{IC}} + \mathcal{L}_{\text{BC}}}$$

Physics Informed Neural Networks (PINNs): learning through differential constraints

$$\partial_t u = \mathcal{N}[u]$$

A neural approximation is trained by minimizing differential-equation residuals, together with initial and boundary conditions.

NEURAL FUNCTION $u_\theta(x,t)$



Automatic differentiation

$$\partial_t u_\theta, \quad \nabla u_\theta, \quad \nabla^2 u_\theta, \ldots$$

PDE residual $R_\theta = \partial_t u_\theta + \mathcal{N}[u_\theta]$

$$\mathcal{L}_{\mathrm{PINN}} = ||R_\theta||^2 + \mathcal{L}_{\mathrm{IC}} + \mathcal{L}_{\mathrm{BC}}$$

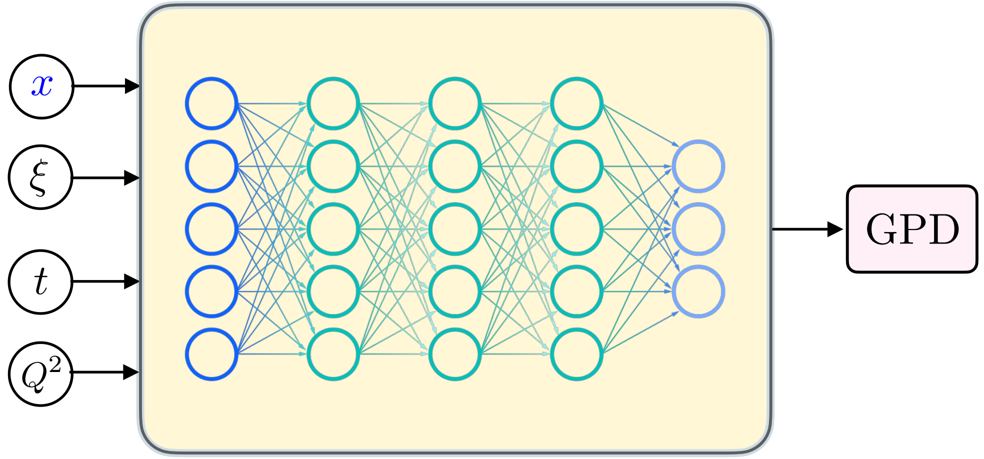
PINN principle

No pointwise labels for the unknown solution are required: the PDE residual, initial conditions, and boundary conditions define the loss.

From differential to integral constraints: NNGPD

NNGPD keeps the PINN strategy, but replaces local differential residuals with nonlocal integral relations to measured observables.

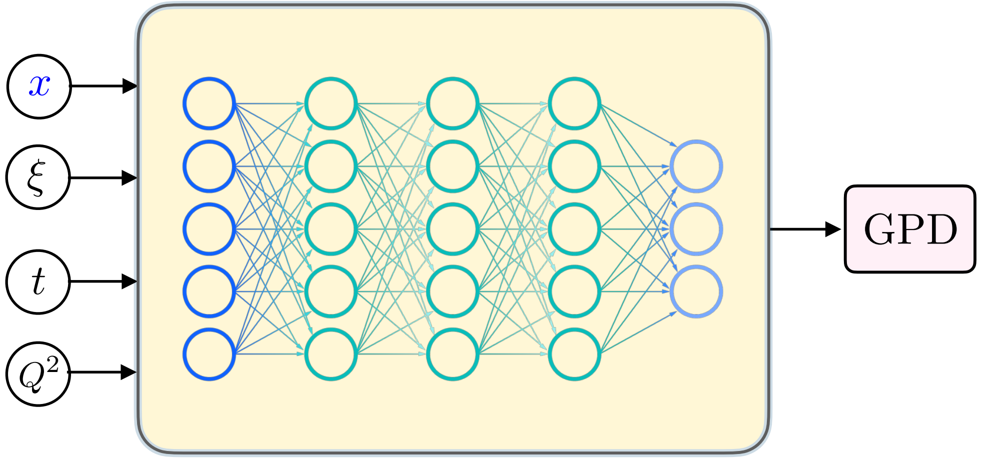
NEURAL FUNCTION $H_\theta(x, \xi, t, Q^2)$



From differential to integral constraints: NNGPD

NNGPD keeps the PINN strategy, but replaces local differential residuals with nonlocal integral relations to measured observables.

NEURAL FUNCTION $H_\theta(x, \xi, t, Q^2)$



Nonlocal integral constraints

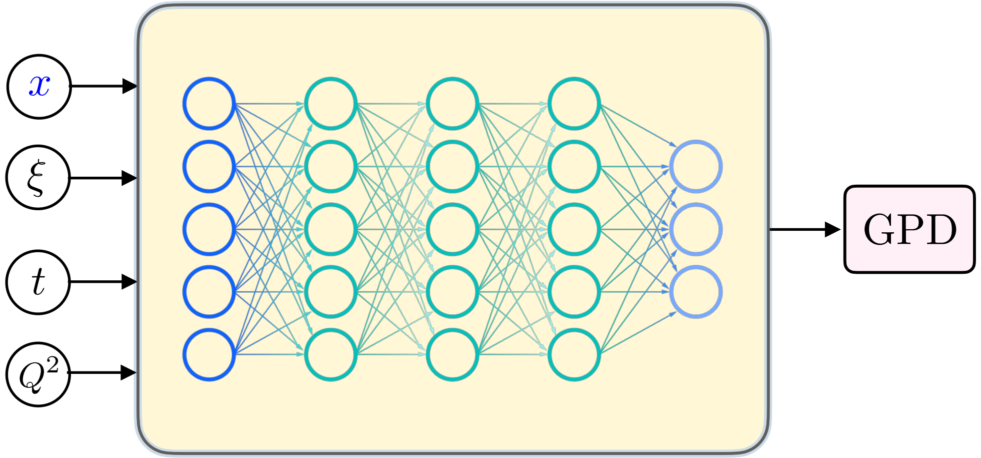
Numerical integration

$$\int dx\, C\, H_\theta, \qquad \int dx\, x^{n-1} H_\theta$$

From differential to integral constraints: NNGPD

NNGPD keeps the PINN strategy, but replaces local differential residuals with nonlocal integral relations to measured observables.

NEURAL FUNCTION $H_\theta(x, \xi, t, Q^2)$



Nonlocal integral constraints

Numerical integration

$$\int dx\, C\, H_\theta, \qquad \int dx\, x^{n-1} H_\theta$$

Observable residuals

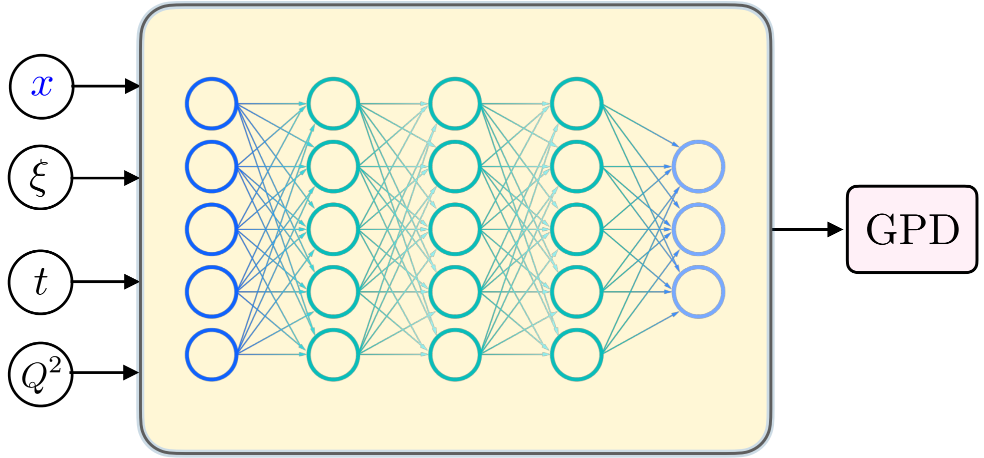
$$\Delta\mathcal{H}_\theta = \mathcal{H}_{\text{data}} - \int dx\, C\, H_\theta$$

$$\mathcal{L}_{\text{NNGPD}} = \mathcal{L}_{\text{CFF}} + \mathcal{L}_{\text{Mellin}} + \mathcal{L}_{\text{sym}} + \mathcal{L}_{\text{BC}}$$

From differential to integral constraints: NNGPD

NNGPD keeps the PINN strategy, but replaces local differential residuals with nonlocal integral relations to measured observables.

NEURAL FUNCTION $H_\theta(x, \xi, t, Q^2)$



Nonlocal integral constraints

Numerical integration

$$\int dx\, C\, H_\theta, \qquad \int dx\, x^{n-1} H_\theta$$

Observable residuals

$$\Delta\mathcal{H}_\theta = \mathcal{H}_{\text{data}} - \int dx\, C\, H_\theta$$

$$\mathcal{L}_{\text{NNGPD}} = \mathcal{L}_{\text{CFF}} + \mathcal{L}_{\text{Mellin}} + \mathcal{L}_{\text{sym}} + \mathcal{L}_{\text{BC}}$$

PINN → NNGPD

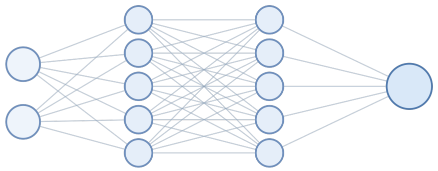
The same physics-informed idea survives: train the unknown function through the projections it must reproduce—not through unavailable pointwise GPD labels.

NNPDF: neural networks as flexible PDF parametrizations

A global fit through the QCD forward map—not supervised learning from pointwise PDF labels.

NN PDFs at Q_0

$$f_a(x, Q_0) = \mathrm{NN}_a(x)$$

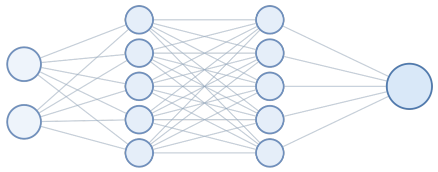


NNPDF: neural networks as flexible PDF parametrizations

A global fit through the QCD forward map—not supervised learning from pointwise PDF labels.

NN PDFs at Q_0

$$f_a(x, Q_0) = \mathrm{NN}_a(x)$$



QCD theory

DGLAP evolution
+ hard coefficients

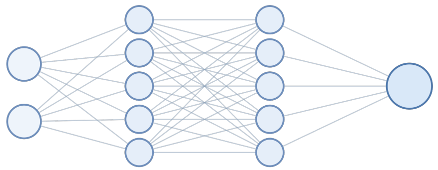
NNPDF: neural networks as flexible PDF parametrizations

A global fit through the QCD forward map—not supervised learning from pointwise PDF labels.

THE FORWARD MAP

NN PDFs at Q_0

$$f_a(x, Q_0) = \mathrm{NN}_a(x)$$



QCD theory

DGLAP evolution
+ hard coefficients

Theory predictions

DIS, DY, jets, W/Z,
top, …

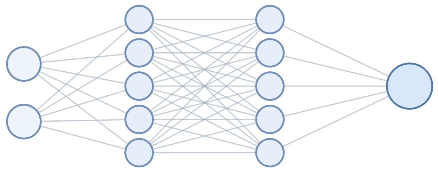
NNPDF: neural networks as flexible PDF parametrizations

A global fit through the QCD forward map—not supervised learning from pointwise PDF labels.

THE FORWARD MAP

NN PDFs at Q_0

$$f_a(x, Q_0) = \mathrm{NN}_a(x)$$



QCD theory

DGLAP evolution
+ hard coefficients

Theory predictions

DIS, DY, jets, W/Z,
top, …

$$T_i[f] = \sum_a \int_0^1 dx\, K_{ia}(x, Q_i, Q_0)\, f_a(x, Q_0)$$

K includes evolution and process-dependent coefficient functions.

COVARIANCE-WEIGHTED TRAINING LOSS

$$\mathcal{L}_{\mathrm{NNPDF}} = \chi^2 = \left(D - T[f]\right)^T C^{-1} \left(D - T[f]\right)$$

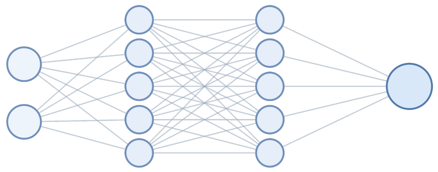
NNPDF: neural networks as flexible PDF parametrizations

A global fit through the QCD forward map—not supervised learning from pointwise PDF labels.

THE FORWARD MAP

NN PDFs at Q_0

$$f_a(x, Q_0) = \mathrm{NN}_a(x)$$



QCD theory

DGLAP evolution
+ hard coefficients

Theory predictions

DIS, DY, jets, W/Z,
top, ...

$$T_i[f] = \sum_a \int_0^1 dx\, K_{ia}(x, Q_i, Q_0)\, f_a(x, Q_0)$$

K includes evolution and process-dependent coefficient functions.

COVARIANCE-WEIGHTED TRAINING LOSS

$$\mathcal{L}_{\mathrm{NNPDF}} = \chi^2 = \big(D - T[f]\big)^T C^{-1} \big(D - T[f]\big)$$

- The network is a flexible functional ansatz; the labels are experimental observables, not PDFs.
- The forward map generally contains integrals—even when implemented as fast matrix contractions.

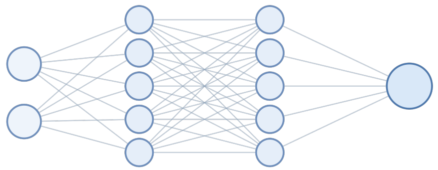
NNPDF: neural networks as flexible PDF parametrizations

A global fit through the QCD forward map—not supervised learning from pointwise PDF labels.

THE FORWARD MAP

NN PDFs at Q_0

$$f_a(x, Q_0) = \mathrm{NN}_a(x)$$



→ **QCD theory**

DGLAP evolution
+ hard coefficients

→ **Theory predictions**

DIS, DY, jets, W/Z, top, ...

$$T_i[f] = \sum_a \int_0^1 dx\, K_{ia}(x, Q_i, Q_0)\, f_a(x, Q_0)$$

K includes evolution and process-dependent coefficient functions.

COVARIANCE-WEIGHTED TRAINING LOSS

$$\mathcal{L}_{\mathrm{NNPDF}} = \chi^2 = \left(D - T[f]\right)^T C^{-1} \left(D - T[f]\right)$$

- The network is a flexible functional ansatz; the labels are experimental observables, not PDFs.
- The forward map generally contains integrals—even when implemented as fast matrix contractions.

WHERE NNGPD GOES FURTHER

A harder second inverse problem

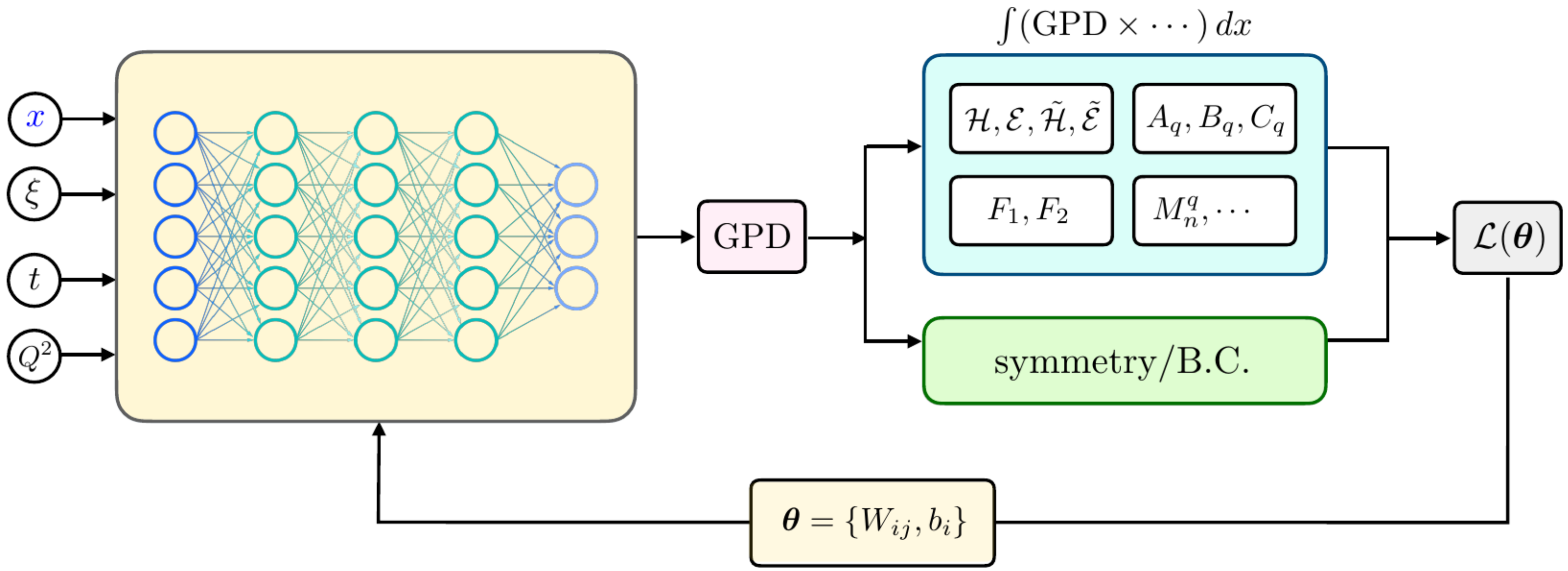
1. PDF data retain a measured x-like scan; DVCS removes GPD x completely under the CFF integral.
2. NNGPD must reconcile complementary projections: complex CFFs, Mellin moments, symmetry, and endpoints.
3. **Uncertainty machinery**

NNPDF: independently fitted Monte Carlo data replicas.
NNGPD here: variational BNN weight posterior.

NNPDF is an important precedent; NNGPD applies the philosophy to a more underdetermined, multidimensional reconstruction.

THE NNGPD FRAMEWORK

Physics acts between the network output and the loss



1 Flexible representation

A neural network maps kinematics to the unknown GPD.

2 Physics operators

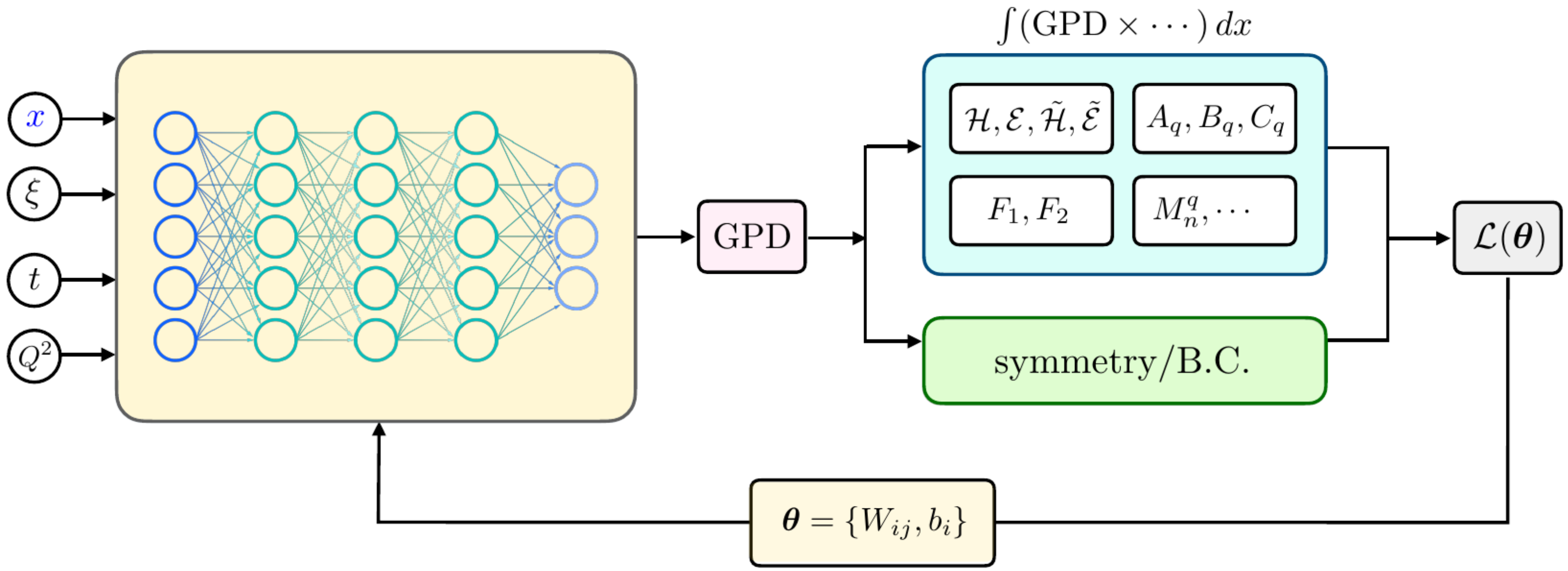
Convolutions, moments, symmetries, and endpoints act on its output.

3 Observable-level training

Their residuals form the loss and update the network parameters.

THE NNGPD FRAMEWORK

Physics acts between the network output and the loss



1 Flexible representation
A neural network maps kinematics to the unknown GPD.

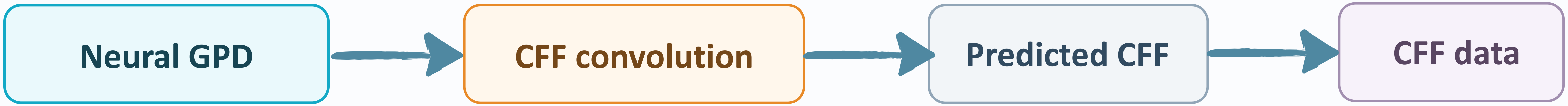
2 Physics operators
Convolutions, moments, symmetries, and endpoints act on its output.

3 Observable-level training
Their residuals form the loss and update the network parameters.

No pointwise GPD values enter the training data.

CFF data constrain the GPD through a nonlocal convolution

The network predicts the function; the known hard kernel maps it to a measurable complex amplitude.



CFF data constrain the GPD through a nonlocal convolution

The network predicts the function; the known hard kernel maps it to a measurable complex amplitude.

Neural GPD → CFF convolution → Predicted CFF → CFF data

FORWARD OPERATOR

$$\widehat{\mathcal{H}}_\theta(\xi, t, Q^2) = \int_{-1}^{1} dx\, C^{(+)}(x, \xi)\, H_\theta(x, \xi, t, Q^2)$$

LO KERNEL

$$C^{(+)}(x, \xi) = \frac{1}{x - \xi + i\epsilon} + \frac{1}{x + \xi - i\epsilon}$$

CONSTRAINT 1 • COMPTON FORM FACTORS

CFF data constrain the GPD through a nonlocal convolution

The network predicts the function; the known hard kernel maps it to a measurable complex amplitude.

Neural GPD → CFF convolution → Predicted CFF → CFF data

FORWARD OPERATOR

$$\widehat{\mathcal{H}}_\theta(\xi,t,Q^2)=\int_{-1}^{1}dx\,C^{(+)}(x,\xi)\,H_\theta(x,\xi,t,Q^2)$$

LO KERNEL

$$C^{(+)}(x,\xi)=\frac{1}{x-\xi+i\epsilon}+\frac{1}{x+\xi-i\epsilon}$$

TRAINING LOSS

$$\mathcal{L}_{\mathrm{CFF}}=\sum_m\left|\mathcal{H}^{\mathrm{data}}(\xi_m,t_m,Q_m^2)-\widehat{\mathcal{H}}_\theta(\xi_m,t_m,Q_m^2)\right|^2$$

CONSTRAINT 1 • COMPTON FORM FACTORS

CFF data constrain the GPD through a nonlocal convolution

The network predicts the function; the known hard kernel maps it to a measurable complex amplitude.

Neural GPD → CFF convolution → Predicted CFF → CFF data

FORWARD OPERATOR

$$\widehat{\mathcal{H}}_\theta(\xi, t, Q^2) = \int_{-1}^{1} dx\, C^{(+)}(x, \xi)\, H_\theta(x, \xi, t, Q^2)$$

LO KERNEL

$$C^{(+)}(x, \xi) = \frac{1}{x - \xi + i\epsilon} + \frac{1}{x + \xi - i\epsilon}$$

TRAINING LOSS

$$\mathcal{L}_{\mathrm{CFF}} = \sum_m \left| \mathcal{H}^{\mathrm{data}}(\xi_m, t_m, Q_m^2) - \widehat{\mathcal{H}}_\theta(\xi_m, t_m, Q_m^2) \right|^2$$

Imaginary part: crossover line

$$\mathrm{Im}\,\mathcal{H}(\xi, t) = -\pi\left[H(\xi, \xi, t) - H(-\xi, \xi, t)\right]$$

Directly samples $x = \pm\xi$.

Real part: principal-value integral

$$\mathrm{Re}\,\mathcal{H}(\xi, t) = \mathrm{P.V.}\int_{-1}^{1} dx \left(\frac{1}{x - \xi} + \frac{1}{x + \xi}\right) H(x, \xi, t)$$

Couples the reconstruction across the full x domain.

A single CFF datum constrains an entire x-dependent slice of the neural GPD.

Mellin moments connect the neural GPD to form factors and lattice QCD

Each power of x provides a different global projection;
together they constrain normalization, shape, and polynomiality.

GLOBAL x-WEIGHTED PROJECTION

$$\widehat{M}^{;q}_{n,\theta}(\xi, t, Q^2) = \int_{-1}^{1} dx\, x^{n-1} H^q_\theta(x, \xi, t, Q^2)$$

Mellin moments connect the neural GPD to form factors and lattice QCD

Each power of x provides a different global projection; together they constrain normalization, shape, and polynomiality.

GLOBAL x-WEIGHTED PROJECTION

$$\widehat{M}^{\,q}_{n,\theta}(\xi, t, Q^2) = \int_{-1}^{1} dx\, x^{n-1} H^q_\theta(x, \xi, t, Q^2)$$

n = 1

Electromagnetic form factors

$$\int_{-1}^{1} dx\, H^q = F_1^q(t), \qquad \int_{-1}^{1} dx\, E^q = F_2^q(t)$$

Experimentally measured F1 and F2 anchor the normalization.

n = 2

Energy–momentum form factors

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

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

$$J^q = \frac{1}{2}\left[A^q(0) + B^q(0)\right]$$

n ≥ 3

Tower of generalized form factors

$$M_n^q(\xi, t) = \sum_{\substack{i=0 \\ i\ \text{even}}}^{n-1} (2\xi)^i A^q_{n,i}(t) + \operatorname{mod}(n-1, 2)(2\xi)^n C_n^q(t)$$

Mellin moments connect the neural GPD to form factors and lattice QCD

Each power of x provides a different global projection;
together they constrain normalization, shape, and polynomiality.

GLOBAL x-WEIGHTED PROJECTION

$$\widehat{M}^{q}_{n,\theta}(\xi,t,Q^2)=\int_{-1}^{1}dx\;x^{n-1}H^q_\theta(x,\xi,t,Q^2)$$

n = 1

Electromagnetic form factors

$$\int_{-1}^{1}dx\;H^q=F_1^q(t),\qquad \int_{-1}^{1}dx\;E^q=F_2^q(t)$$

Experimentally measured F1 and F2 anchor the normalization.

n = 2

Energy–momentum form factors

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

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

$$J^q=\frac{1}{2}\left[A^q(0)+B^q(0)\right]$$

n ≥ 3

Tower of generalized form factors

$$M^q_n(\xi,t)=\sum_{\substack{i=0\\ i\,\mathrm{even}}}^{n-1}(2\xi)^iA^q_{n,i}(t)$$

$$+\,\mathrm{mod}(n-1,2)(2\xi)^nC^q_n(t)$$

MOMENT LOSS

$$\mathcal{L}_{\mathrm{MM}}=\sum_{n,m}\left|M_n^{q,\mathrm{data}}(\xi_m,t_m,Q_m^2)-\widehat{M}^{q}_{n,\theta}(\xi_m,t_m,Q_m^2)\right|^2$$

Vector sector: even n constrain the antisymmetric H+ combination; odd n constrain the symmetric H− combination.

CFFs and Mellin moments illuminate complementary projections of the same hidden function.

CONSTRAINT 3 • EXACT STRUCTURE

Symmetry and endpoints restrict the admissible function space

These are exact pointwise properties—not additional fitted observables.

CHARGE CONJUGATION	$H^{\bar{q}}(x,\xi,t) = -H^{q}(-x,\xi,t),$	$H^{q\pm} = H^{q} \pm H^{\bar{q}}$

Symmetry and endpoints restrict the admissible function space

These are exact pointwise properties—not additional fitted observables.

CHARGE CONJUGATION $H^{\bar{q}}(x,\xi,t) = -H^{q}(-x,\xi,t), \qquad H^{q\pm} = H^{q} \pm H^{\bar{q}}$

SYMMETRIC COMBINATION

Mirror symmetry across x = 0

$$H^{q-}(-x,\xi,t) = H^{q-}(x,\xi,t)$$

The negative-x antiquark region is fixed by the positive-x prediction.

ANTISYMMETRIC COMBINATION

Sign reversal across x = 0

$$H^{q+}(-x,\xi,t) = -H^{q+}(x,\xi,t)$$

Continuity therefore enforces a node at x = 0.

Symmetry and endpoints restrict the admissible function space

These are exact pointwise properties—not additional fitted observables.

CHARGE CONJUGATION

$$H^{\bar{q}}(x,\xi,t) = -H^{q}(-x,\xi,t), \qquad H^{q\pm} = H^{q} \pm H^{\bar{q}}$$

SYMMETRIC COMBINATION

Mirror symmetry across x = 0

$$H^{q-}(-x,\xi,t) = H^{q-}(x,\xi,t)$$

The negative-x antiquark region is fixed by the positive-x prediction.

ANTISYMMETRIC COMBINATION

Sign reversal across x = 0

$$H^{q+}(-x,\xi,t) = -H^{q+}(x,\xi,t)$$

Continuity therefore enforces a node at x = 0.

KINEMATIC ENDPOINTS

$$H^{q\pm}(x=\pm 1,\xi,t) = 0$$

Both combinations vanish as the active parton carries all longitudinal momentum.

CONSTRAINT 3 • EXACT STRUCTURE

Symmetry and endpoints restrict the admissible function space

These are exact pointwise properties—not additional fitted observables.

CHARGE CONJUGATION

$$H^{\bar{q}}(x,\xi,t) = -H^{q}(-x,\xi,t), \qquad H^{q\pm} = H^{q} \pm H^{\bar{q}}$$

SYMMETRIC COMBINATION

Mirror symmetry across x = 0

$$H^{q-}(-x,\xi,t) = H^{q-}(x,\xi,t)$$

The negative-x antiquark region is fixed by the positive-x prediction.

ANTISYMMETRIC COMBINATION

Sign reversal across x = 0

$$H^{q+}(-x,\xi,t) = -H^{q+}(x,\xi,t)$$

Continuity therefore enforces a node at x = 0.

KINEMATIC ENDPOINTS

$$H^{q\pm}(x=\pm 1,\xi,t) = 0$$

Both combinations vanish as the active parton carries all longitudinal momentum.

TWO EQUIVALENT IMPLEMENTATION ROUTES

Architectural enforcement

Learn only x ≥ 0; reconstruct x < 0 using the exact symmetry relation.

Penalty enforcement

$$\mathcal{L}_{\mathrm{sym}} = \sum_m \left|\widehat{H}_\theta^{q+}(0,\xi_m,t_m,Q_m^2)\right|^2$$

$$\mathcal{L}_{\mathrm{end}} = \sum_m \left|\widehat{H}_\theta^{q\pm}(1,\xi_m,t_m,Q_m^2)\right|^2$$

Exact constraints remove unphysical solutions before the integral data determine the remaining shape.

The loss is sector-specific—and deliberately reweighted

The physics terms are the same projections introduced earlier; the numerical hierarchy is different in H^+ and H^-.

CFF-VISIBLE ANTISYMMETRIC SECTOR H^+

$$\mathcal{L}^{+}_{\text{total}} = w_{CFF}\mathcal{L}_{CFF} + w^{+}_{MM}\mathcal{L}^{+}_{MM} + w^{+}_{EP}\mathcal{L}^{+}_{EP} + w_{KL}\mathcal{L}_{KL}$$

EP combines $H^+(0)=0$ with the $x=1$ endpoint condition.

CFF-BLIND SYMMETRIC SECTOR H^-

$$\mathcal{L}^{-}_{\text{total}} = w^{-}_{MM}\mathcal{L}^{-}_{MM} + w^{-}_{EP}\mathcal{L}^{-}_{EP} + w_{\text{smooth}}\mathcal{L}_{\text{smooth}} + w_{\text{diversity}}\mathcal{L}_{\text{div}} + w_{KL}\mathcal{L}_{KL}$$

No CFF term: smoothness and ensemble diversity stabilize the underconstrained sector.

The loss is sector-specific—and deliberately reweighted

The physics terms are the same projections introduced earlier; the numerical hierarchy is different in H^+ and H^-.

CFF-VISIBLE ANTISYMMETRIC SECTOR H^+

$$\mathcal{L}^+_{\text{total}} = w_{CFF}\mathcal{L}_{CFF} + w^+_{MM}\mathcal{L}^+_{MM} + w^+_{EP}\mathcal{L}^+_{EP} + w_{KL}\mathcal{L}_{KL}$$

EP combines $H^+(0)=0$ with the $x=1$ endpoint condition.

CFF-BLIND SYMMETRIC SECTOR H^-

$$\mathcal{L}^-_{\text{total}} = w^-_{MM}\mathcal{L}^-_{MM} + w^-_{EP}\mathcal{L}^-_{EP} + w_{\text{smooth}}\mathcal{L}_{\text{smooth}} + w_{\text{diversity}}\mathcal{L}_{\text{div}} + w_{KL}\mathcal{L}_{KL}$$

No CFF term: smoothness and ensemble diversity stabilize the underconstrained sector.

Smoothness penalty

$$\mathcal{L}_{\text{smooth}} = \sum_m \int dx \left| \frac{\partial \hat{H}^-_q}{\partial x} \right|^2$$

Suppresses uncontrolled local oscillations; implemented as a first-difference penalty on the x-grid.

BNN-diversity penalty

$$\mathcal{L}_{\text{div}} = -\mathbb{E}\left[\log \mathrm{Var}(\hat{H}^-_q)\right]$$

Penalizes collapse of the posterior spread, retaining an ensemble of functions compatible with the moments.

The loss is sector-specific—and deliberately reweighted

The physics terms are the same projections introduced earlier; the numerical hierarchy is different in H^+ and H^-.

CFF-VISIBLE ANTISYMMETRIC SECTOR H^+

$$\mathcal{L}^+_{\text{total}} = w_{CFF}\mathcal{L}_{CFF} + w^+_{MM}\mathcal{L}^+_{MM} + w^+_{EP}\mathcal{L}^+_{EP} + w_{KL}\mathcal{L}_{KL}$$

EP combines $H^+(0)=0$ with the x=1 endpoint condition.

CFF-BLIND SYMMETRIC SECTOR H^-

$$\mathcal{L}^-_{\text{total}} = w^-_{MM}\mathcal{L}^-_{MM} + w^-_{EP}\mathcal{L}^-_{EP} + w_{\text{smooth}}\mathcal{L}_{\text{smooth}} + w_{\text{diversity}}\mathcal{L}_{\text{div}} + w_{KL}\mathcal{L}_{KL}$$

No CFF term: smoothness and ensemble diversity stabilize the underconstrained sector.

Smoothness penalty

$$\mathcal{L}_{\text{smooth}} = \sum_m \int dx \left| \frac{\partial \hat{H}^-_q}{\partial x} \right|^2$$

Suppresses uncontrolled local oscillations; implemented as a first-difference penalty on the x-grid.

BNN-diversity penalty

$$\mathcal{L}_{\text{div}} = -\mathbb{E}\Big[\log \mathrm{Var}(\hat{H}^-_q)\Big]$$

Penalizes collapse of the posterior spread, retaining an ensemble of functions compatible with the moments.

Fixed coefficients

H^+ TRAINING

$$\begin{aligned} w_{\text{MM}} &= 10, \\ w^{\text{Re}}_{\text{CFF}} = w^{\text{Im}}_{\text{CFF}} &= 0.1, \\ w_{\text{end-point}} &= 10, \\ w_{\text{KL}} &= 0.001. \end{aligned}$$

H^- TRAINING

$$\begin{aligned} w_{\text{MM}_1} &= 10, & w_{\text{MM}_3} &= 160, \\ w_{\text{MM}_5} &= 730, & w_{\text{MM}_7} &= 2500, \\ w_{\text{end-point}} &= 1, & w_{\text{smooth}} &= 0.5, \\ w_{\text{diversity}} &= 5, & w_{\text{KL}} &= 0.001. \end{aligned}$$

Same values in every run

The weights compensate disparate raw scales—they keep small moment losses from being ignored by large CFF losses.

PROOF OF CONCEPT • CLOSURE TEST

Can integral projections recover a hidden GPD?

A controlled closure test separates the information used for training from the pointwise truth used for evaluation.

TRAINING PHASE

1 HIDDEN TRUTH

UVA2 GPD model

Known only to the data generator

2 FORWARD PHYSICS

Generate projections

CFF convolutions and Mellin moments

3 TRAINING TARGETS

Synthetic observables

Real/imaginary CFFs
Moments through n = 7

4 Train NNGPD

Loss uses only CFFs, moments, symmetry, and endpoints—never pointwise GPD values.

EVALUATION PHASE

5 RECONSTRUCTION

NNGPD prediction

Pointwise function with Bayesian uncertainty

Unblind and compare

Does the reconstructed GPD recover the hidden UVA2 shape across kinematics?

unblind

No information leakage

UVA2 pointwise values never enter the optimization.

Closure criterion

Reproduce both the imposed observables and the withheld underlying function.

The closure test asks whether the inverse reconstruction works before real-data complications are introduced.

CLOSURE TEST • TRAINING SETUP

One fixed model across a broad kinematic domain

The closure test probes the constraint structure—not architecture or hyperparameter tuning.

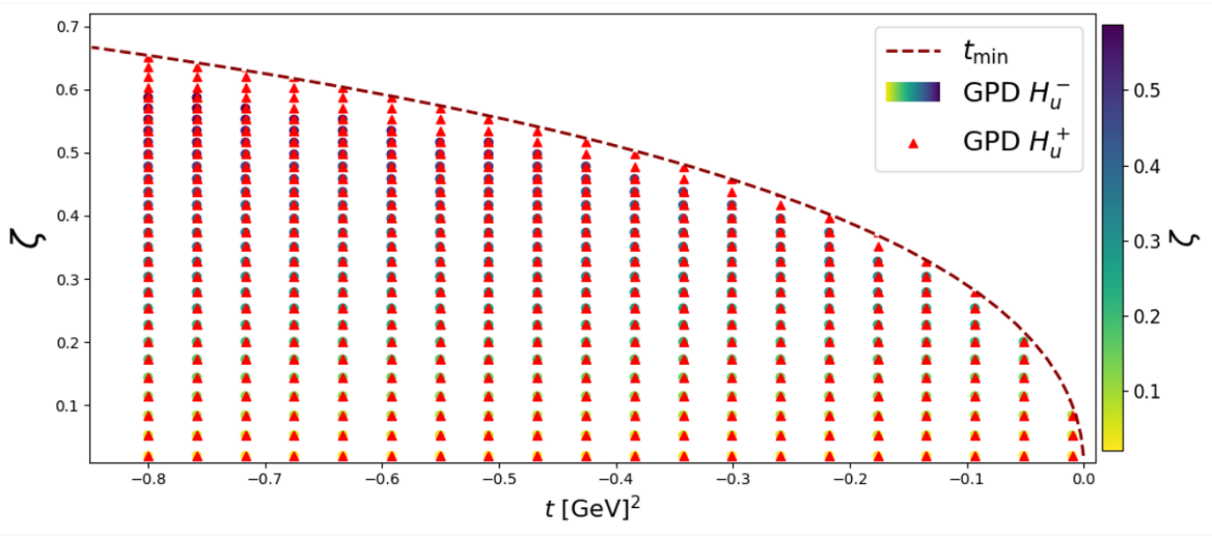
9,455
kinematic configurations

4 → 1
inputs (x, xi, t, Q2) → one GPD

Q^2: 4–10 GeV^2
xi: 0.01–0.50 • −t: 0.01–0.80 GeV2

Integral targets
complex CFFs • moments (n ≤ 7)

Kinematic coverage at fixed Q2



Model and uncertainty

ARCHITECTURE

Fully connected network; hidden width 100; SiLU activations; layer normalization and dropout (0.1).

SECTOR-SPECIFIC TARGETS

Antisymmetric sector: even moments (2, 4, 6) plus real and imaginary CFFs.
Symmetric sector: odd moments (1, 3, 5, 7).

BAYESIAN NETWORK

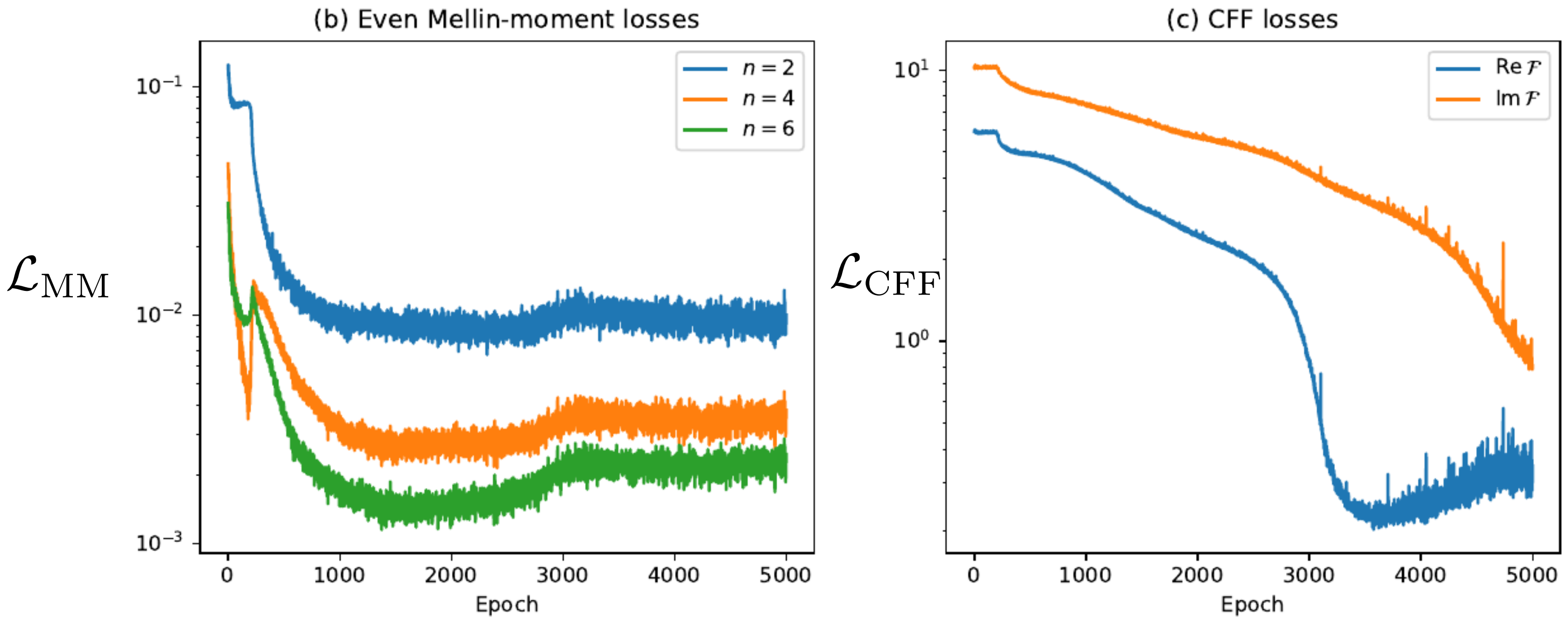
Gaussian weight posterior; repeated samples give the mean reconstruction and epistemic uncertainty.

Fixed across all runs: architecture and regularization. Varied only: random initialization and sampled Bayesian weights.

RESULT 1 • OBSERVABLE-LEVEL CLOSURE

NNGPD simultaneously fits moments and complex CFFs

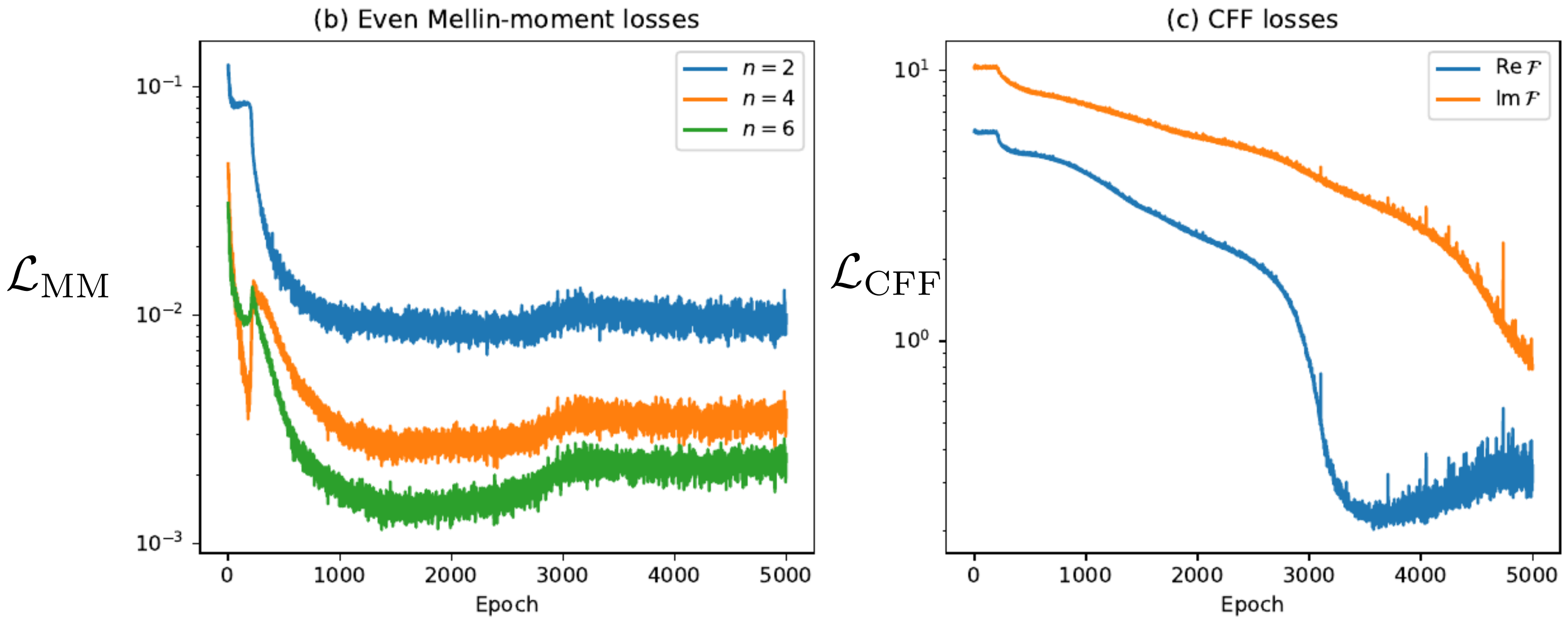
All imposed integral residuals decrease together—before asking whether the hidden pointwise function is recovered.



RESULT 1 • OBSERVABLE-LEVEL CLOSURE

NNGPD simultaneously fits moments and complex CFFs

All imposed integral residuals decrease together—before asking whether the hidden pointwise function is recovered.



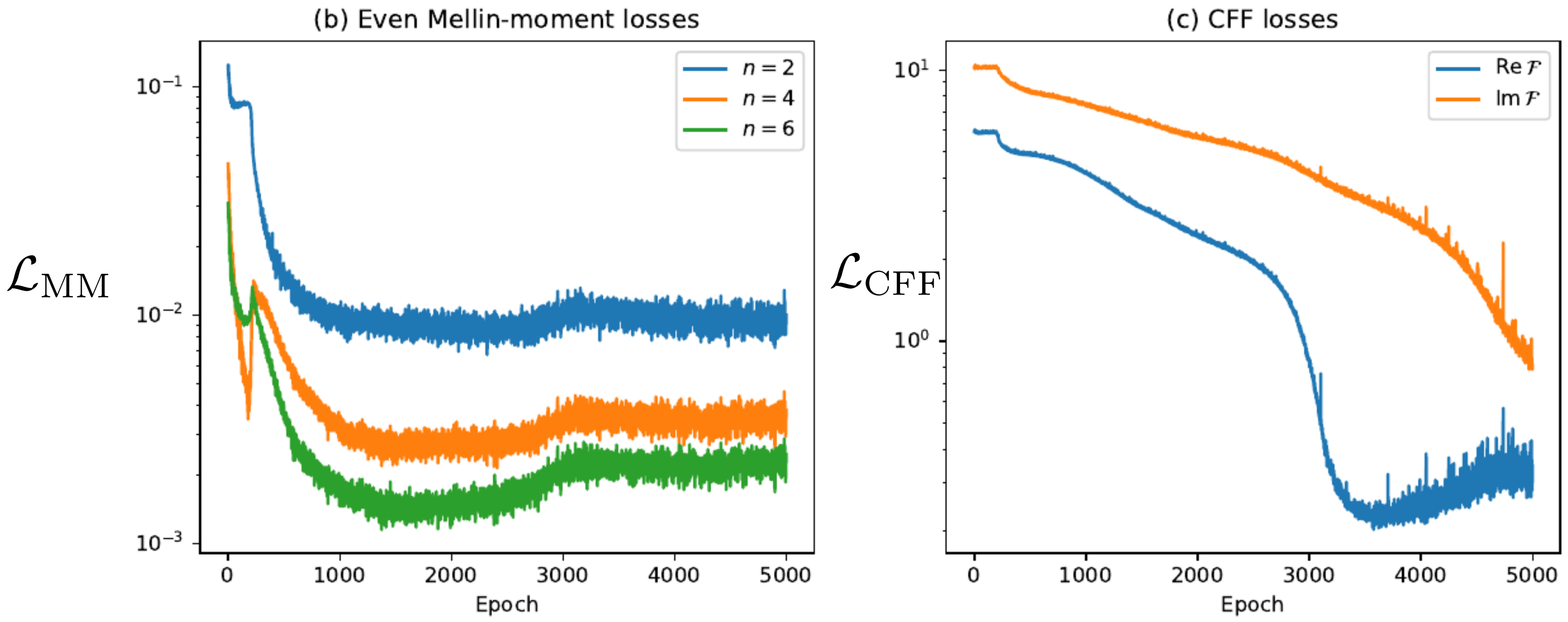
Mellin hierarchy

The n = 2, 4, and 6 residuals all fall and stabilize without sacrificing one order for another.

RESULT 1 • OBSERVABLE-LEVEL CLOSURE

NNGPD simultaneously fits moments and complex CFFs

All imposed integral residuals decrease together—before asking whether the hidden pointwise function is recovered.



Mellin hierarchy

The n = 2, 4, and 6 residuals all fall and stabilize without sacrificing one order for another.

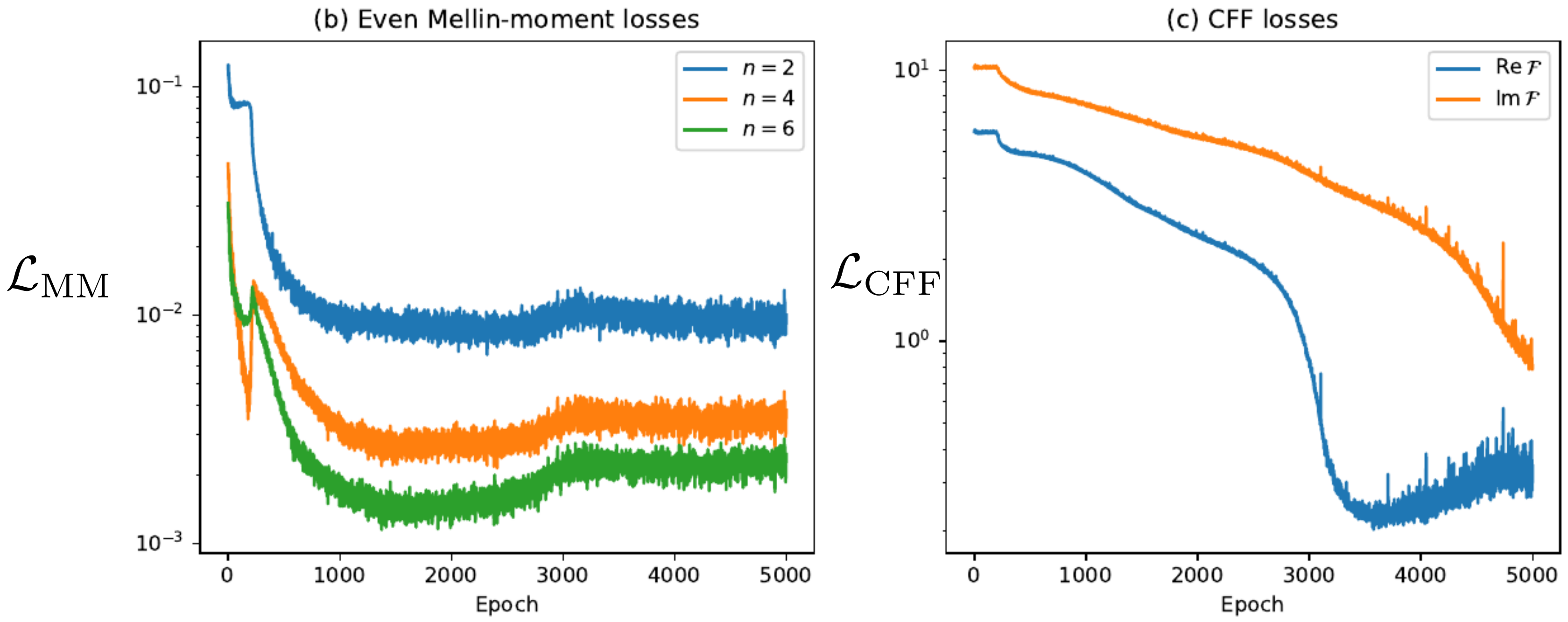
Complex CFF

Real and imaginary components decrease concurrently, preserving both analytic pieces of the amplitude.

RESULT 1 • OBSERVABLE-LEVEL CLOSURE

NNGPD simultaneously fits moments and complex CFFs

All imposed integral residuals decrease together—before asking whether the hidden pointwise function is recovered.



Mellin hierarchy

The n = 2, 4, and 6 residuals all fall and stabilize without sacrificing one order for another.

Complex CFF

Real and imaginary components decrease concurrently, preserving both analytic pieces of the amplitude.

What this establishes

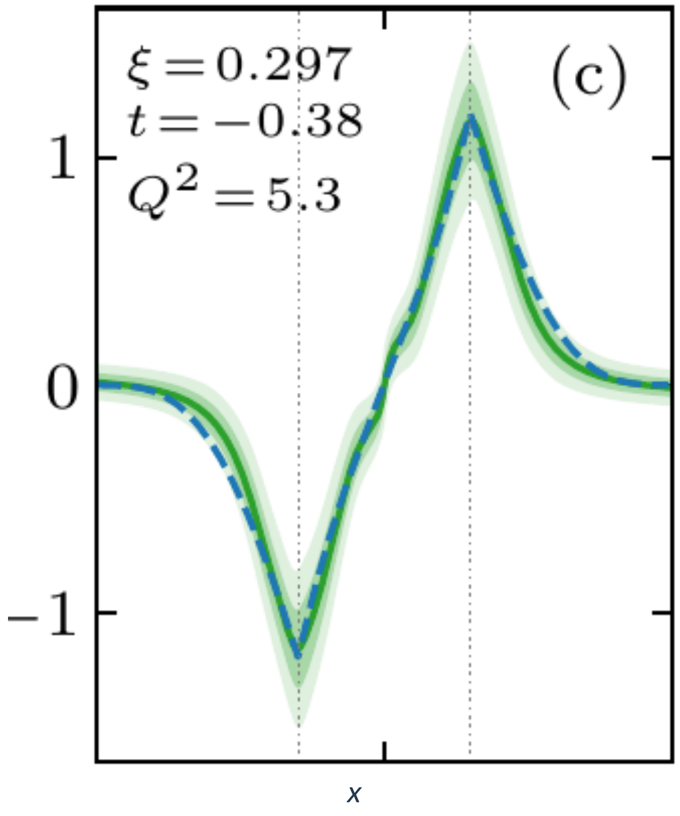
The learned function is compatible with the supplied projections; pointwise closure is the stronger next test.

Observable-level agreement is necessary—but not sufficient—to solve the inverse problem.

RESULT 2 • POINTWISE CLOSURE

Integral constraints recover the hidden x-dependence

A representative antisymmetric GPD slice reconstructed from even moments and complex CFFs—without pointwise labels.



Adapted from manuscript Fig. 7(c)

DGLAP | ERBL | DGLAP

How to read the closure test

NNGPD Bayesian ensemble mean

Withheld UVA2 ground truth

Posterior uncertainty bands (1 sigma and 2 sigma)

1 The NNGPD mean follows the hidden curve across both DGLAP and ERBL regions.

2 Sharp variations near the kinematic boundaries x = ±xi are reproduced, with small local deviations.

3 The withheld truth remains within the 2-sigma posterior band throughout this slice.

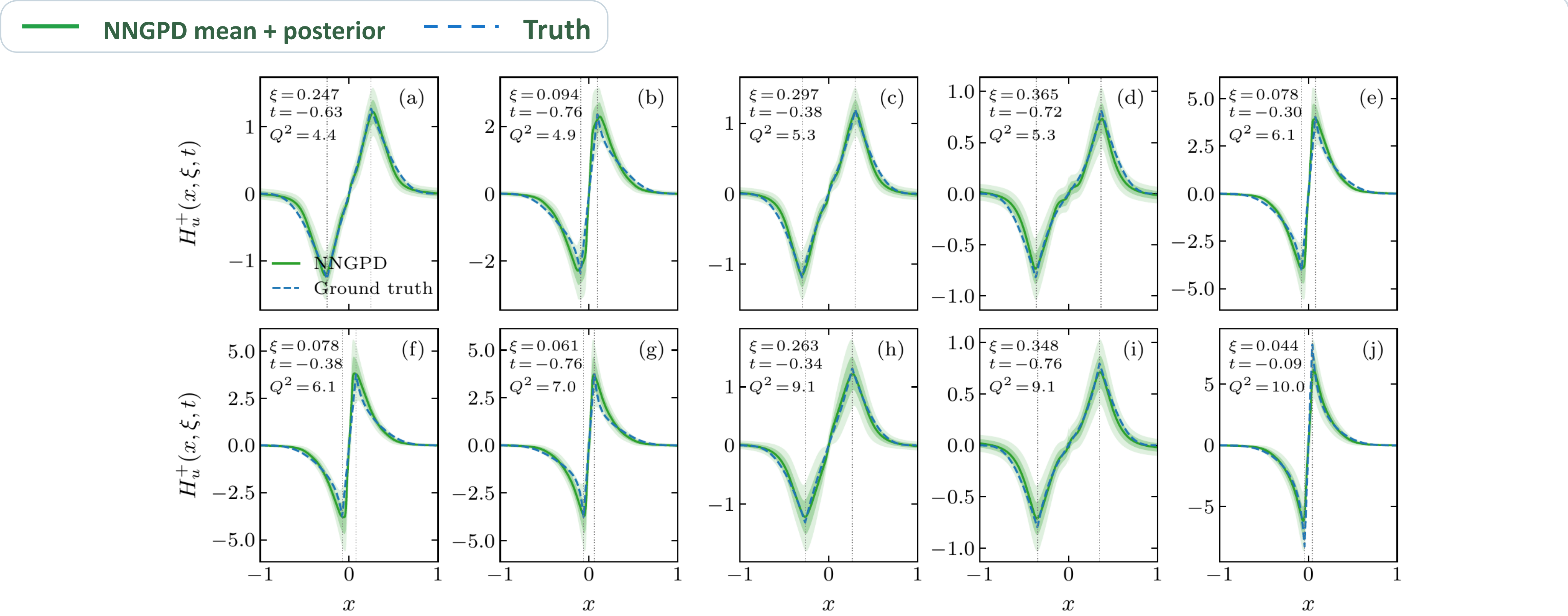
The blue curve was never shown to the network.

This is genuine pointwise closure from nonlocal information.

RESULT 3 • KINEMATIC DEPENDENCE

One representation, many kinematics

Across ten representative configurations, the learned H^{+}_{u} changes shape and scale while retaining pointwise closure.



The closure is global—not tied to one favored kinematic point.

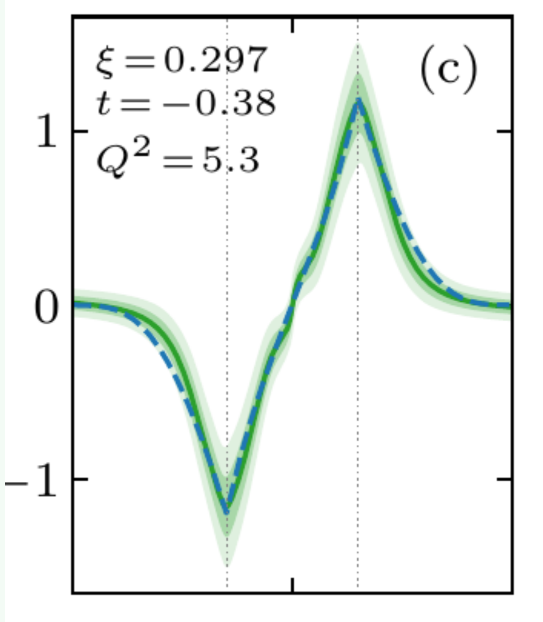
Residual differences remain localized near x = ±ξ, where the target varies most rapidly; the truth remains covered by the posterior bands.

What does the CFF projection buy us?

A channel-level ablation: compare the sector seen by CFFs with the sector constrained by Mellin moments alone.

H^{+}_{u} — antisymmetric channel

Re CFF + Im CFF + even moments n = 2, 4, 6



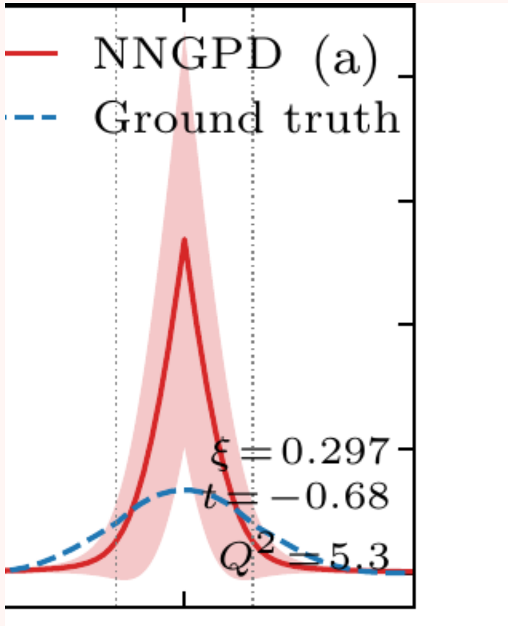
narrow

posterior bands truth closely tracked

sharp x = ±ξ structure

H^{-}_{u} — symmetric channel

CFF-blind + odd moments n = 1, 3, 5, 7 only



~10× wider

posterior bands overall shape retained

local detail unresolved

CFFs and Mellin moments are complementary projections—not interchangeable constraints.

CFFs anchor detailed x-dependence in H^{+}; moments provide indispensable access to symmetry sectors and integral structure.

TAKE-HOME MESSAGE

NNGPD reconstructs GPDs from complementary projections

A unified, uncertainty-aware solution of the CFF-to-GPD inverse problem—without pointwise GPD labels.

1 THE SECOND INVERSE PROBLEM

From CFFs to GPDs

$$H(x, \xi, t, Q^2) \to \mathcal{H}(\xi, t, Q^2)$$

The experimentally inaccessible x-dependence disappears under the CFF convolution. Reconstructing it is more underdetermined than the PDF problem reviewed in NNPDF.

No measured GPD x-axis

2 PHYSICS-CONSTRAINED LEARNING

Complementary constraints

- complex CFFs
- Mellin moments
- symmetry
- endpoints

One BNN must satisfy every operator simultaneously. Their complementarity—not any one integral loss—is the central idea.

The posterior quantifies unresolved freedom

3 WHAT THE CLOSURE TEST SHOWS

The hidden x-shape is recovered

✓ CFFs and moments close together

✓ Pointwise x-shape is recovered across kinematics

✓ H^+ versus H^- bands expose resolving power

Pointwise GPD truth is never used in training

NEXT:

- **Replace pseudo-observables with experimental CFFs and lattice-QCD moments.**
- **Incorporate constraints such as polynomiality into NN representations**