Accessing x-dependent GPDs from generative AI

Zhite Yu
(Jefferson Lab, Theory Center)

- GPD extraction provides a perfect testbed for AI
- Can GPD be uniquely extracted in principle?
- DVCS, evolution, new observables, polynomiality

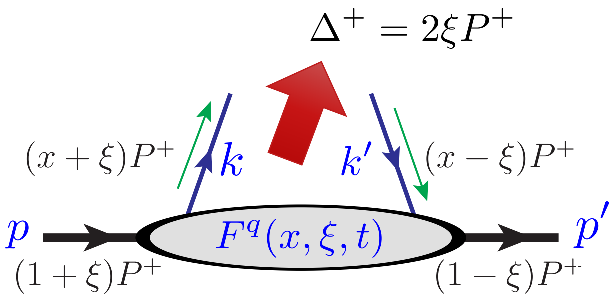
Sep/19/2025
Argonne National Laboratory

In collaboration with:
Wally Melnitochouk
Jianwei Qiu
Nobuo Sato
Marco Zaccheddu

JHEP 08 (2022) 103
PRD 107 (2023) 014007
PRL 131 (2023) 161902
PRD 109 (2024) 074023
PRD 111 (2025) 094014
paper in preparation

GPD and hadron structure

❑ Generalized parton distribution (GPD)

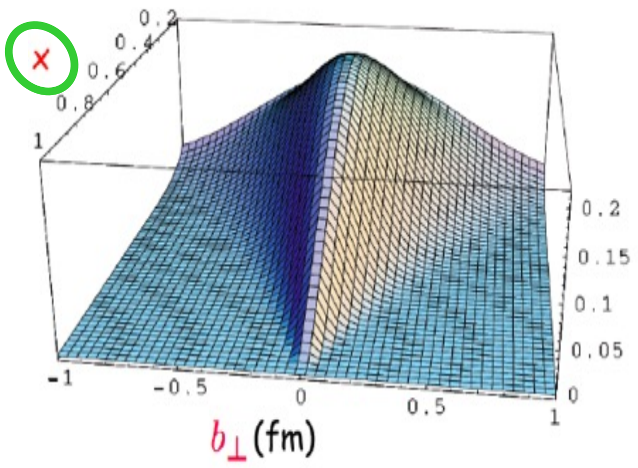


$$F^q(x,\xi,t) = \int \frac{\mathrm{d}z^-}{4\pi} e^{-ixP^+z^-} \langle p' | \bar{q}(z^-/2)\gamma^+ q(-z^-/2) | p \rangle$$
$$= \frac{1}{2P^+}\left[H^q(x,\xi,t)\,\bar{u}(p')\gamma^+ u(p) - E^q(x,\xi,t)\,\bar{u}(p')\frac{i\sigma^{+\alpha}\Delta_\alpha}{2m} u(p)\right]$$

❑ Tomography

$$f_i(x,\boldsymbol{b}_T) = \int \mathrm{d}^2\boldsymbol{\Delta}_T\, e^{i\boldsymbol{\Delta}_T\cdot\boldsymbol{b}_T} F_i(x,0,-\Delta_T^2)$$

3D image



M. Burkardt, 2000, 2003

❑ Emergent hadron properties

$$\int_{-1}^{1} dx\, x\, F_i(x,\xi,t) \propto \langle p' | T_i^{++}(0) | p \rangle \qquad i = q, g$$

$$\int_{-1}^{1} dx\, x\, H_i(x,\xi,t) = A_i(t) + \xi^2 D_i(t)$$
$$\int_{-1}^{1} dx\, x\, E_i(x,\xi,t) = B_i(t) - \xi^2 D_i(t)$$

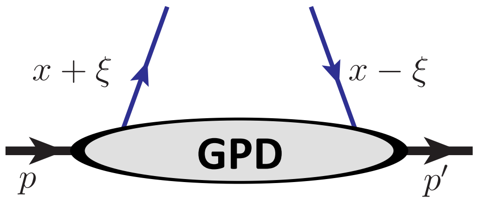
$$J_i = \lim_{t\to 0}\int_{-1}^{1} dx\, x\,[H_i(x,\xi,t) + E_i(x,\xi,t)]$$

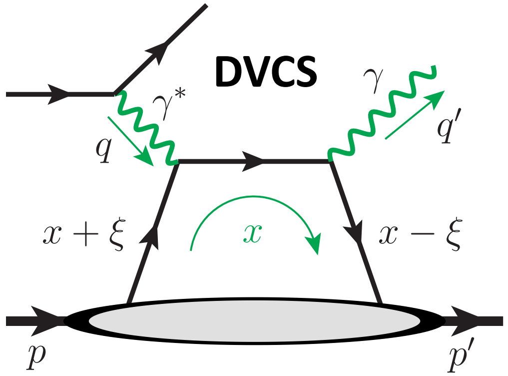
X.-D. Ji, 1997

Need full x-dependence at a substantial range of (t,ξ)!

x-dependence problem for GPD

- **Amplitude nature: exclusive processes**



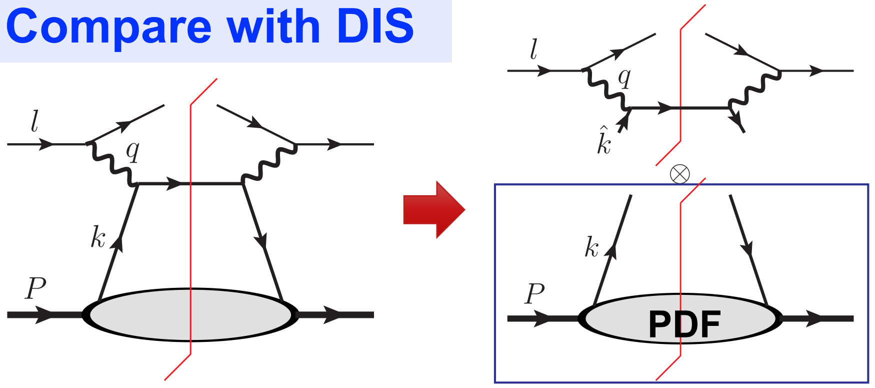


***x* ~ loop momentum**

$$\mathcal{M}(\xi,t;Q)=\int_{-1}^{1}dx\,F(x,\xi,t;\mu)\,C(x,\xi;Q,\mu)$$

never pin down to some x

Compare with DIS

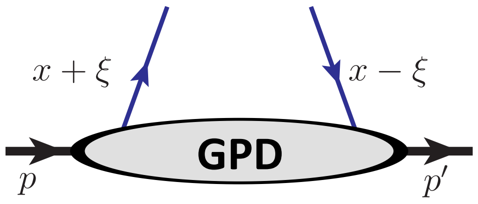


***cross section*: cut diagram**

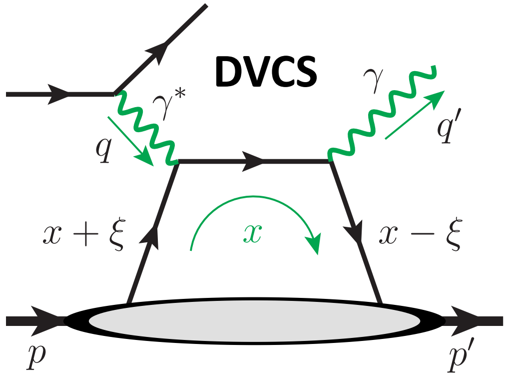
$$\sigma_{\rm DIS}\simeq\int_{x_B}^{1}{\rm d}x\,f(x)\,\hat{\sigma}(x/x_B)$$

x-dependence problem for GPD

❑ Amplitude nature: exclusive processes



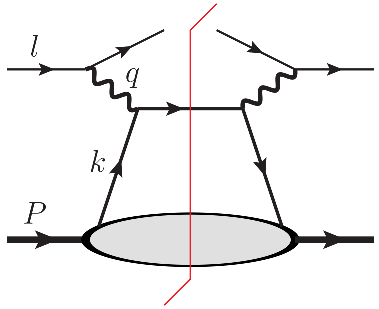
x ~ loop momentum

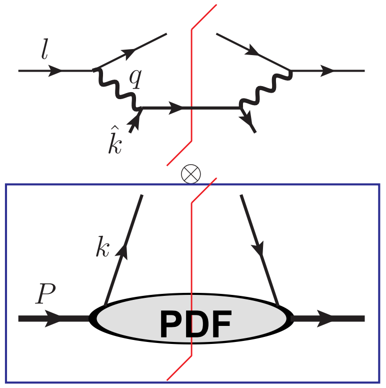


$$\mathcal{M}(\xi, t; Q) = \int_{-1}^{1} dx\, F(x, \xi, t; \mu)\, C(x, \xi; Q, \mu)$$

never pin down to some x

Compare with DIS





cross section: **cut diagram**

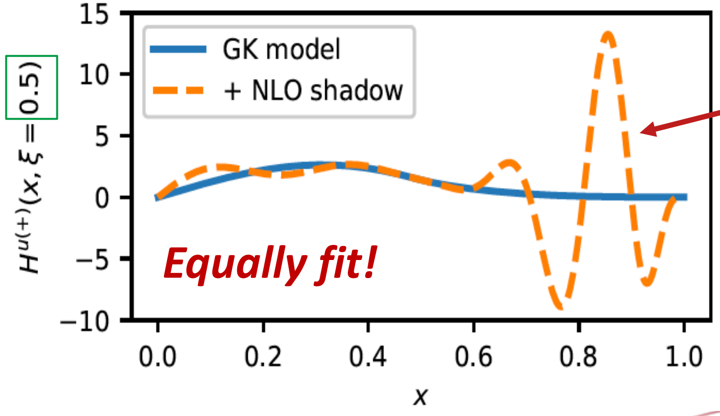
$$\sigma_{\rm DIS} \simeq \int_{x_B}^{1} \mathrm{d}x\, f(x)\, \hat{\sigma}(x/x_B)$$

❑ Sensitivity to x: comes from $C(x, \xi; Q, \mu)$

$$C^{\rm S}(x, \xi; Q, \mu) = T(Q/\mu) \cdot G(x, \xi) \propto \frac{1}{x - \xi + i\epsilon}$$

$$\Rightarrow \mathcal{M}(\xi, t; Q) \propto \int_{-1}^{1} dx\, \frac{F^+(x, \xi, t)}{x - \xi + i\epsilon}$$

"Scaling integral"



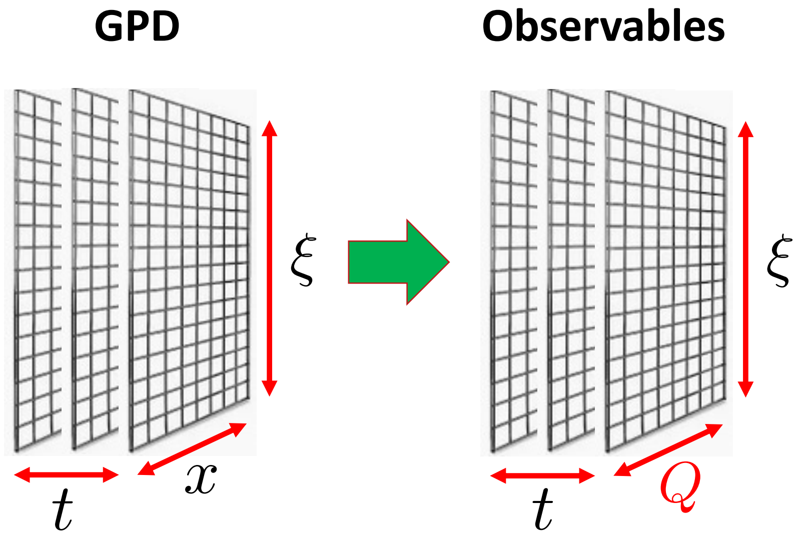
Shadow GPD

[Bertone et al. PRD `21]

Pixelated construction of GPD

Shadow GPDs make parametric method *biased*

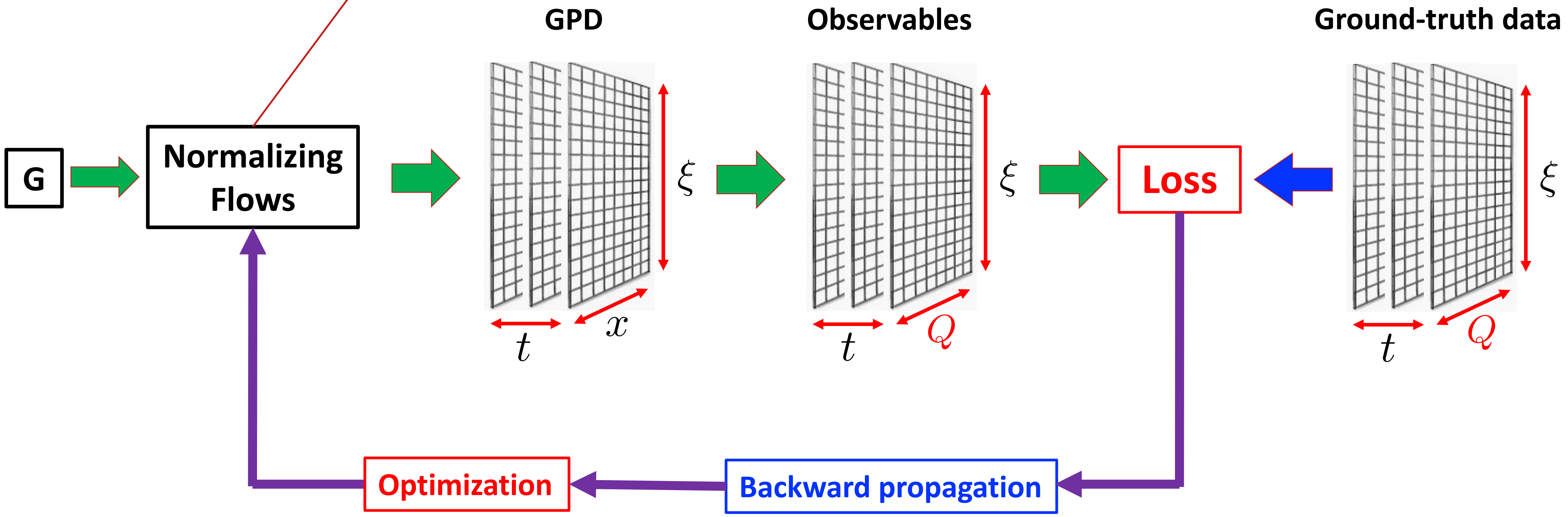
→ **Construct GPDs from most flexible pixelation method**



Normalizing flows

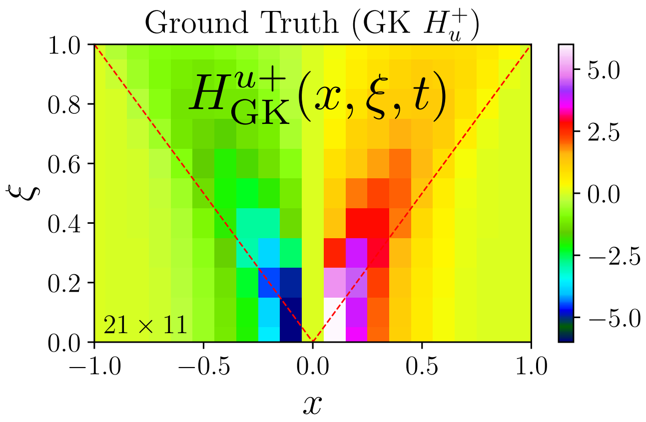
Normalizing flows = A set of *differentiable and invertible* changes of variable.

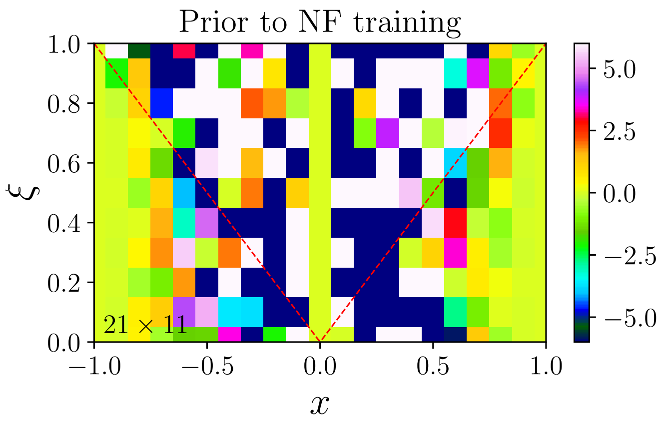
Gaussian distribution $N_D(\vec{x}) \xrightarrow{\vec{y} = f(\vec{x})} P_D(\vec{y}) = \left| \frac{\partial^D f(\vec{x})}{\partial \vec{x}^D} \right|^{-1} N_D(\vec{x}) \Big|_{\vec{x} = f^{-1}(\vec{y})}$

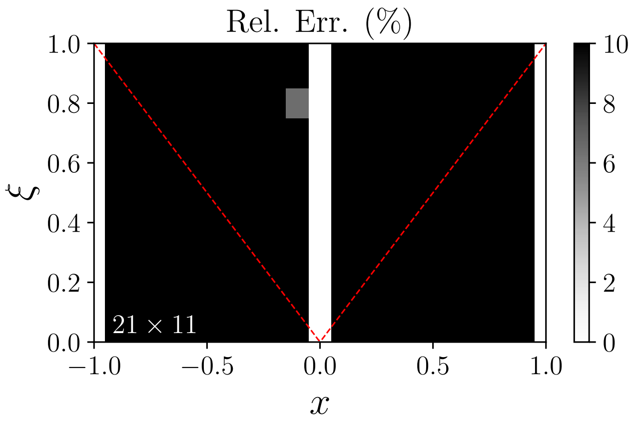


Jefferson Lab

Reconstruct with only scaling DVCS integral







- **Neural network generator**

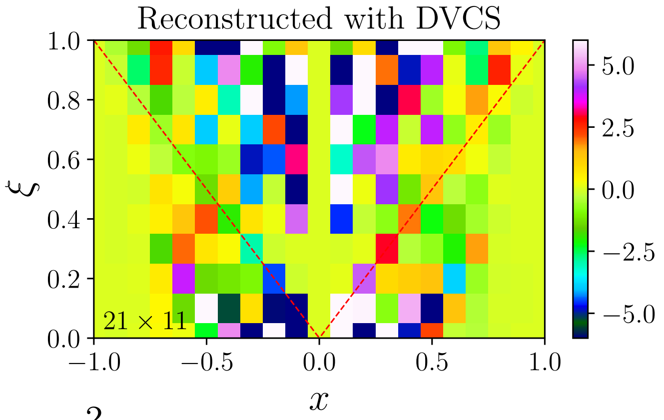
$$H_{\mathrm{NF}}^{u}(x,\xi,t) = H_{\mathrm{GK}}^{u}(x,\xi,t) * \epsilon(x,\xi,t)$$

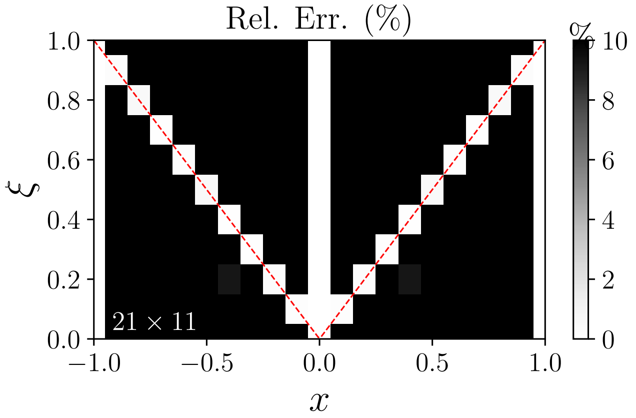
$\epsilon(x,\xi,t)$ is generated by NF

Training with DVCS moment. $r = 0.01$

- **Observable: DVCS integral**

$$\mathcal{M}^{[H]}(\xi,t) = \int_{-1}^{1} dx \, \frac{H^{+}(x,\xi,t)}{x-\xi+i\epsilon}$$





- **Optimize with a loss function**

$$\chi^2(H_{\mathrm{NF}}, H_{\mathrm{GK}}) = \sum_{\xi,t} \left| \frac{\mathcal{M}^{[H_{\mathrm{NF}}]}(\xi,t) - \mathcal{M}^{[H_{\mathrm{GK}}]}(\xi,t)}{r \cdot \mathcal{M}^{[H_{\mathrm{GK}}]}(\xi,t)} \right|^2$$

Sensitivity only on the ridge.

How to understand the result? [Neglecting complication of NF for now]

$$\mathcal{M}^{[H]}(\xi,t) = \int_{-1}^{1} dx\, \frac{H^{+}(x,\xi,t)}{x-\xi+i\epsilon} = \boxed{P\int_{-1}^{1} dx\, H(x,\xi,t)\left[\frac{2x}{x^2-\xi^2}\right]} - i\pi\left[H(\xi,\xi,t) - H(-\xi,\xi,t)\right]$$

Each GPD pixel is independent.

➢ Optimization process

Calculate observable and get loss

Backward propagation

$$\frac{\partial\,\chi^2}{\partial H_{\rm NF}(x,\xi,t)} \propto \left[\mathcal{M}^{[H_{\rm NF}]}(\xi,t) - \mathcal{M}^{[H_{\rm GK}]}(\xi,t)\right]\cdot\left[\frac{2x}{x^2-\xi^2}\right]$$

(Real part only)

Optimization (gradient descent)

$$H_{\rm NF}(x,\xi,t) \to H_{\rm NF}(x,\xi,t) - \mathrm{lr}\cdot\frac{\partial\chi^2}{\partial H_{\rm NF}(x,\xi,t)}$$

How to understand the result?

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

Each GPD pixel is independent.

➢ Optimization process

Calculate observable and get loss

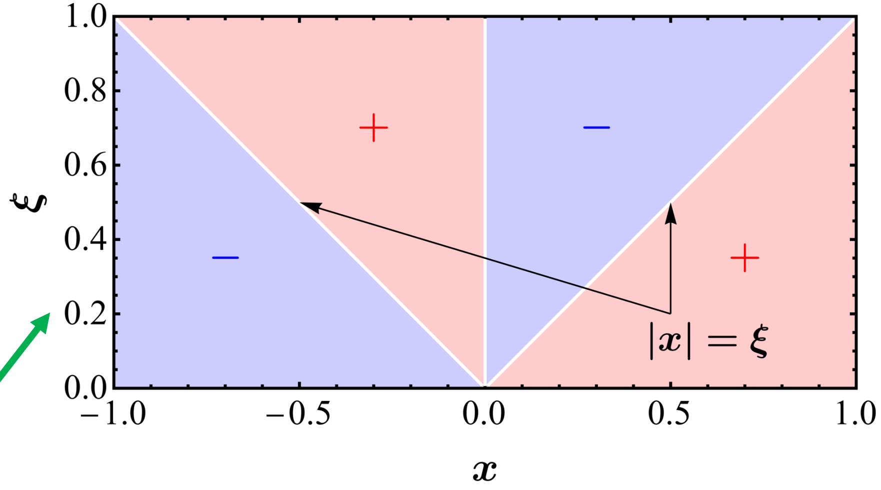
Backward propagation

$$\frac{\partial\chi^2}{\partial H_{\mathrm{NF}}(x,\xi,t)} \propto \left[\mathcal{M}^{[H_{\mathrm{NF}}]}(\xi,t) - \mathcal{M}^{[H_{\mathrm{GK}}]}(\xi,t)\right]\cdot\left[\frac{2x}{x^2-\xi^2}\right]$$

(Real part only)

Optimization (gradient descent)

$$H_{\mathrm{NF}}(x,\xi,t) \to H_{\mathrm{NF}}(x,\xi,t) - \mathrm{lr}\cdot\frac{\partial\chi^2}{\partial H_{\mathrm{NF}}(x,\xi,t)}$$



Tunning of each pixel is *deterministic*!

Determined by the sign of

$$\mathcal{M}^{[H_{\mathrm{NF}}]}(\xi,t) - \mathcal{M}^{[H_{\mathrm{GK}}]}(\xi,t)$$

in the *initial* input.

Keeps tunning until reaching a "solution"

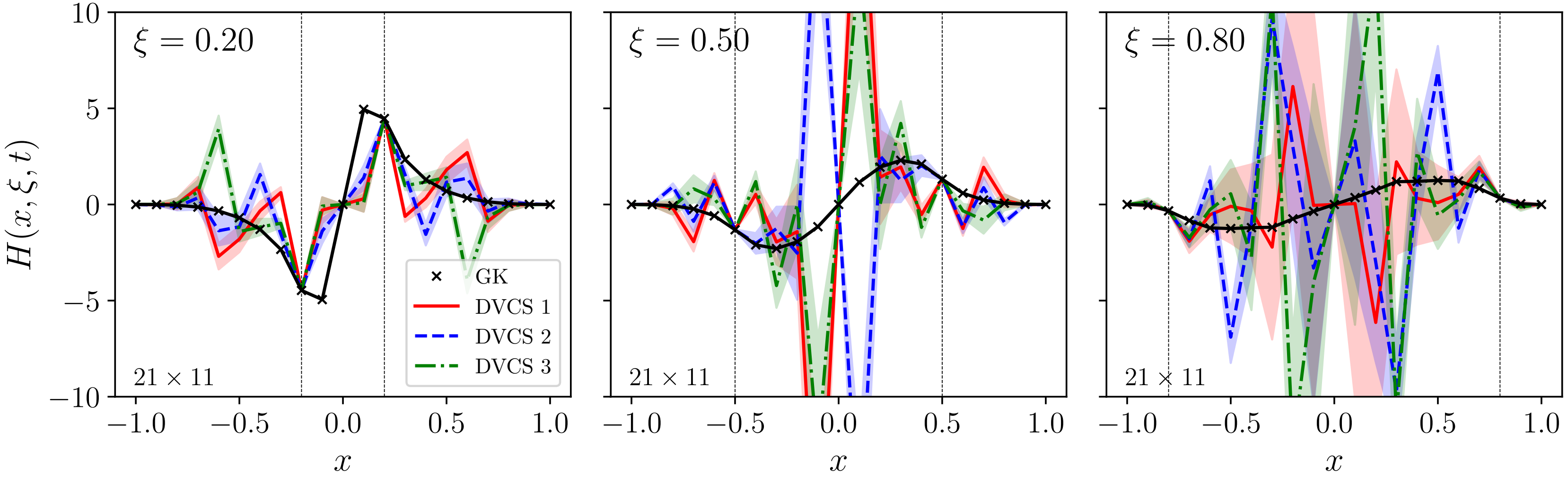
→ shadow GPD!

Jefferson Lab

Shadow GPDs from pixelation approach

DVCS integral

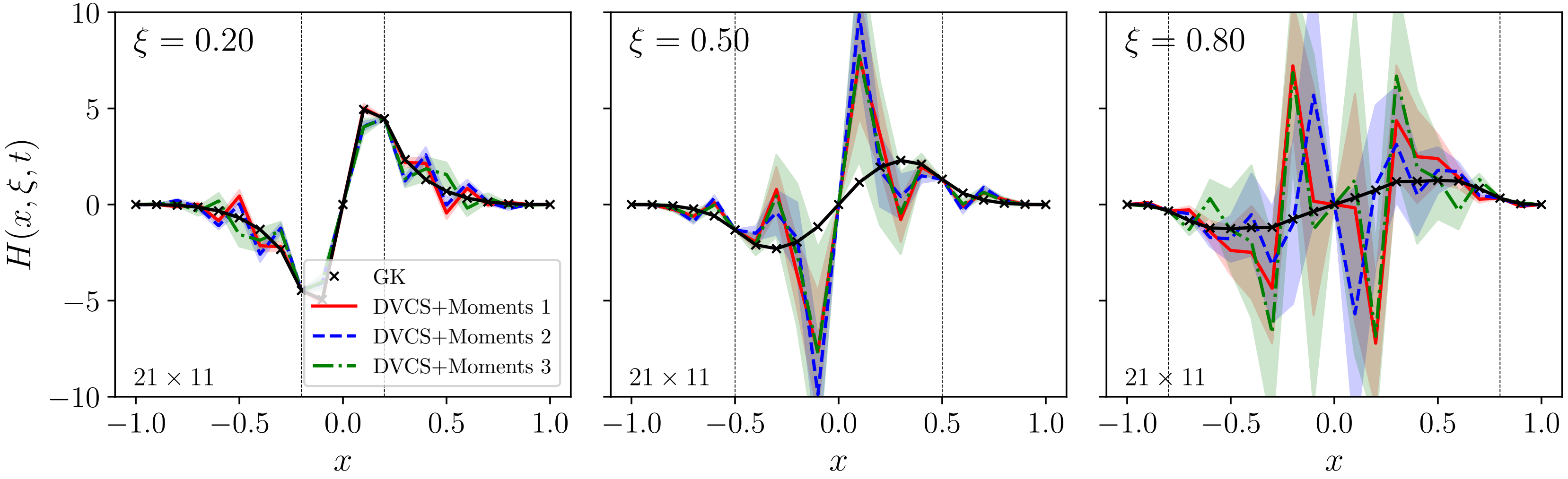
$$\mathcal{M}^{\mathrm{DVCS}}(\xi,t)=\int_{-1}^{1}dx\,\frac{H^{+}(x,\xi,t)}{x-\xi+i\epsilon}$$



DVCS integral
+ first two x-moments

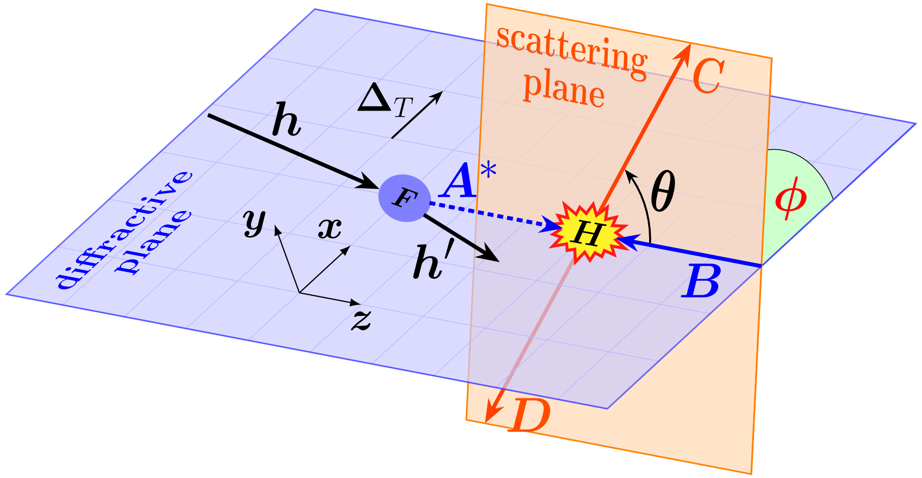
$$H_n^{+}(\xi,t)=\int_{-1}^{1}dx\,x^{n-1}H^{+}(x,\xi,t)$$

$$n=2,4$$



Jefferson Lab

Types of x-sensitivity



❑ x-sensitivity ⇔ 2 → 2 hard scattering

Kinematics:

1. $\hat{s} = 2\,\xi\, s\,/\,(1+\xi)$ ⇐ ξ
2. θ or $q_T = (\sqrt{\hat{s}}/2)\sin\theta$ ⇔ x
3. ϕ ⇐ (A^*B) spin states

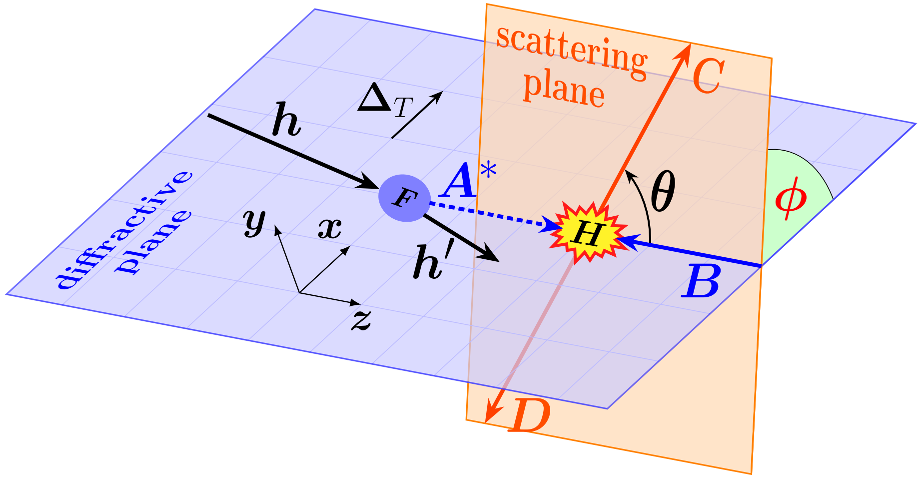
$$\mathcal{M}(Q,\phi) \simeq \sum_A e^{i(\lambda_A-\lambda_B)\phi}\cdot\int_{-1}^{1} dx\, F_A(x)\, C_A(x;Q) \qquad (Q=\theta \text{ or } q_T)$$

[suppressing t and ξ dependence]

Jefferson Lab

Types of x-sensitivity

[Qiu, Yu, PRD 107 (2023) 014007; Qiu, Sato, Yu, PRD 111 (2025) 094014]



❑ x-sensitivity ⇔ 2 → 2 hard scattering

Kinematics:

1. $\hat{s} = 2\,\xi\,s\,/\,(1+\xi)$ ⟵ ξ
2. θ or $q_T = (\sqrt{\hat{s}}/2)\sin\theta$ ⟷ x
3. ϕ ⟵ (A^*B) spin states

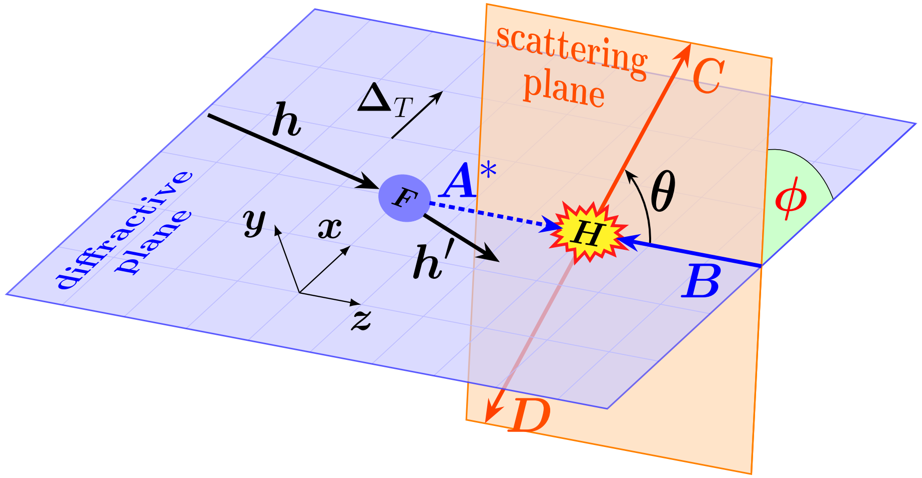
$$\mathcal{M}(Q,\phi) \simeq \sum_A e^{i(\lambda_A-\lambda_B)\phi} \cdot \int_{-1}^{1} dx\, F_A(x)\, C_A(x;Q) \qquad (Q = \theta \text{ or } q_T)$$

[suppressing t and ξ dependence]

➢ **Scaling kernel** $\boxed{C(x;Q) = G(x)\cdot T(Q)}$ ⟹ $F_G = \int_{-1}^{1} dx\, G(x)\, F(x,\xi,t)$ **Independent of Q.**

⟹ **Inversion problem: *shadow GPD*** $S_G = \int_{-1}^{1} dx\, G(x)\, S(x,\xi) = 0$ [Bertone et al. PRD `21]

Types of x-sensitivity



❑ x-sensitivity ⇔ 2 → 2 hard scattering

Kinematics:

1. $\hat{s} = 2\,\xi\, s\,/\,(1+\xi)$ ⟸ ξ
2. θ or $q_T = (\sqrt{\hat{s}}/2)\,\sin\theta$ ⟺ x
3. ϕ ⟸ (A^*B) spin states

$$\mathcal{M}(Q,\phi) \simeq \sum_A e^{i(\lambda_A-\lambda_B)\phi}\cdot\int_{-1}^{1} dx\, F_A(x)\, C_A(x;Q) \qquad (Q=\theta \text{ or } q_T)$$

[suppressing t and ξ dependence]

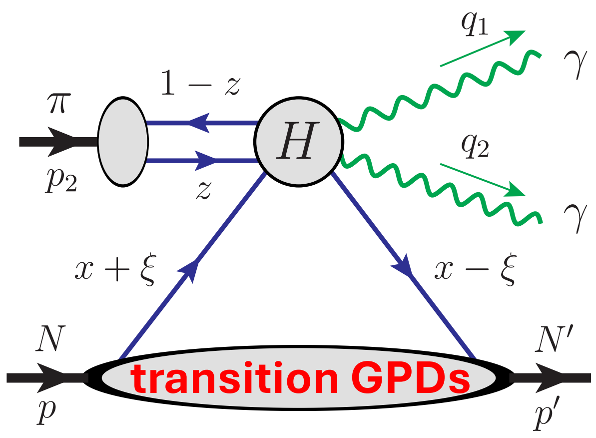
➢ **Scaling kernel** $\boxed{C(x;Q)=G(x)\cdot T(Q)}$ ⟹ $F_G = \int_{-1}^{1} dx\, G(x)\, F(x,\xi,t)$ **Independent of Q.**

⟹ **Inversion problem: *shadow GPD*** $S_G = \int_{-1}^{1} dx\, G(x)\, S(x,\xi) = 0$ [Bertone et al. PRD `21]

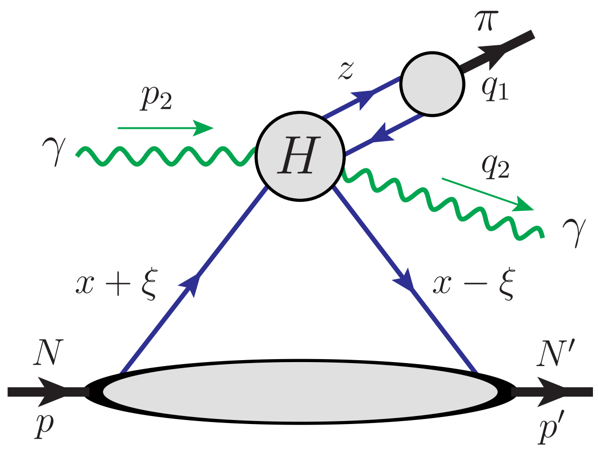
➢ **Non-scaling kernel** $\boxed{C(x;Q)\neq G(x)\cdot T(Q)}$ ⟹ $d\sigma/dQ \sim \left|C(x;Q)\otimes_x F(x,\xi,t)\right|^2$

Jefferson Lab

Two processes with non-scaling hard coefficients



J-PARC, AMBER



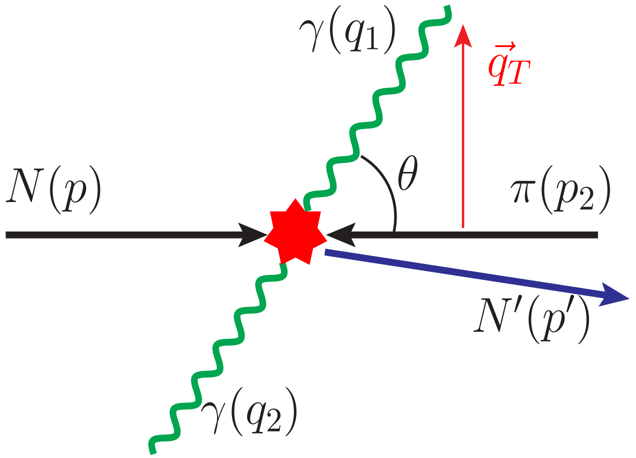
JLab Hall D

Qiu & Yu, JHEP 08 (2022) 103
Qiu & Yu, PRD 109 (2024) 074023

G. Duplancic et al., JHEP 11 (2018) 179
G. Duplancic et al., JHEP 03 (2023) 241
G. Duplancic et al., PRD 107 (2023), 094023
Qiu & Yu, PRD 107 (2023), 014007
Qiu & Yu, PRL 131 (2023), 161902

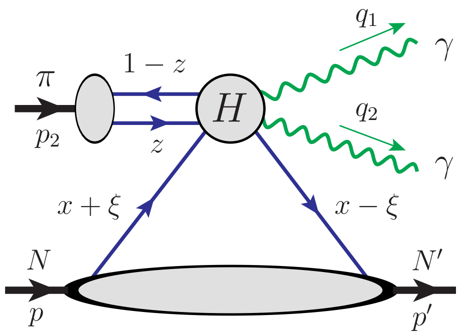
Enhanced x-sensitivity: (1) diphoton mesoproduction

[Qiu & Yu, JHEP 08 (2022) 103;
PRD 109 (2024) 074023]



In addition to scaling integral

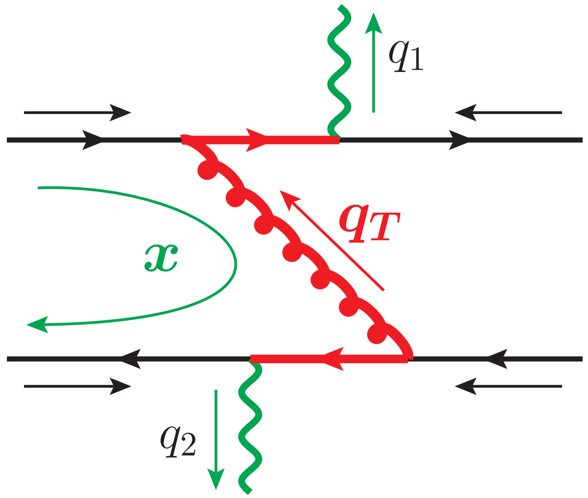
$$F_0(\xi,t) = \int_{-1}^{1} \frac{dx\, F(x,\xi,t)}{x-\xi+i\epsilon}$$

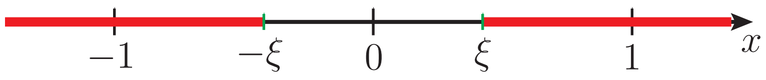


$\mathcal{M}$ **also contains *non-scaling* integral**

$$I(t,\xi;z,\theta) = \int_{-1}^{1} \frac{dx\, F(x,\xi,t)}{x-\rho(z;\theta)+i\epsilon\,\mathrm{sgn}\left[\cos^2(\theta/2)-z\right]}$$

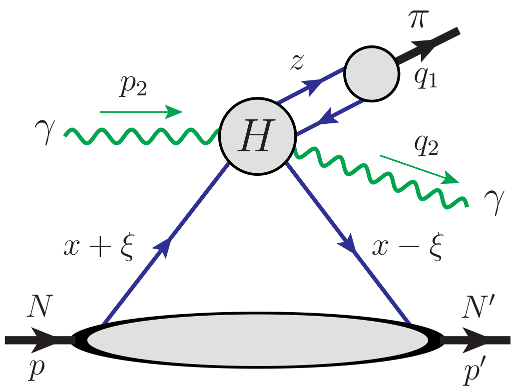
$$\rho(z;\theta) = \xi \cdot \left[\frac{1-z+\tan^2(\theta/2)\,z}{1-z-\tan^2(\theta/2)\,z}\right] \in (-\infty,-\xi]\cup[\xi,\infty)$$

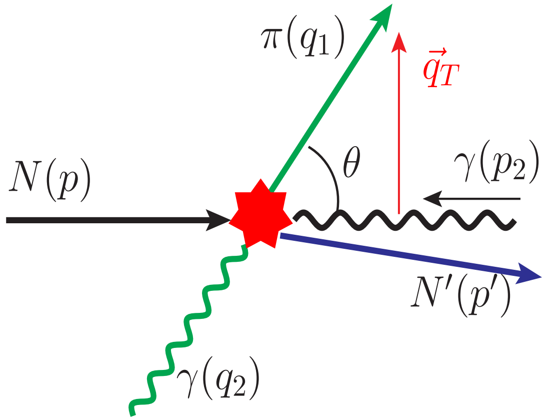




Enhanced x-sensitivity: (2) γ-π pair photoproduction

[Qiu & Yu, PRL 131 (2023) 161902]

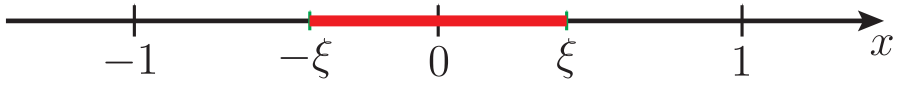




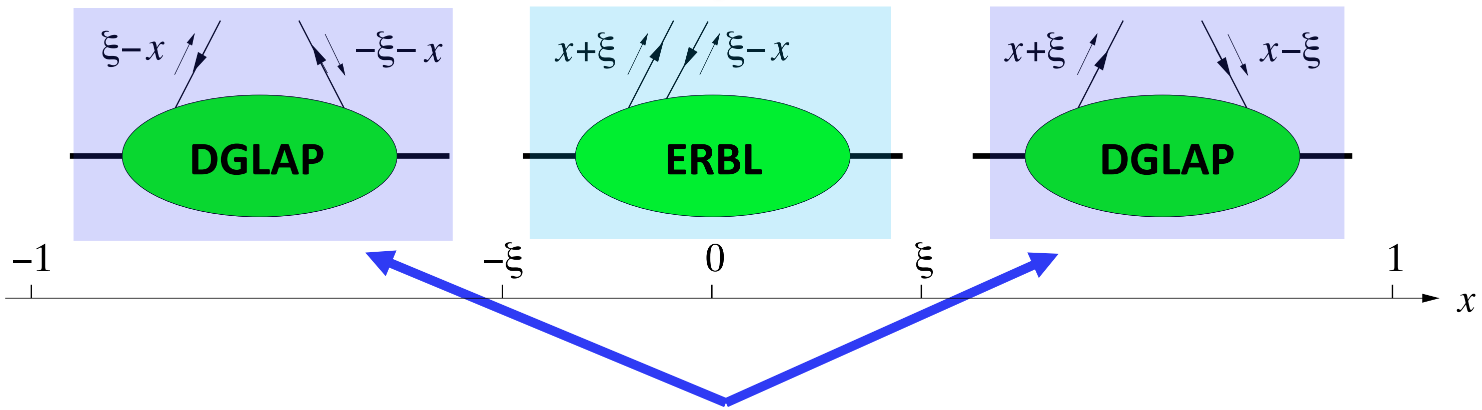
$\mathcal{M}$ also contains the ***non-scaling*** integral

$$I'(t,\xi;z,\theta) = \int_{-1}^{1} \frac{dx\, F(x,\xi,t)}{x - \rho'(z;\theta) + i\epsilon}$$

$$\rho'(z;\theta) = \xi \cdot \left[\frac{\cos^2(\theta/2)\,(1-z) - z}{\cos^2(\theta/2)\,(1-z) + z} \right] \in [-\xi, \xi]$$



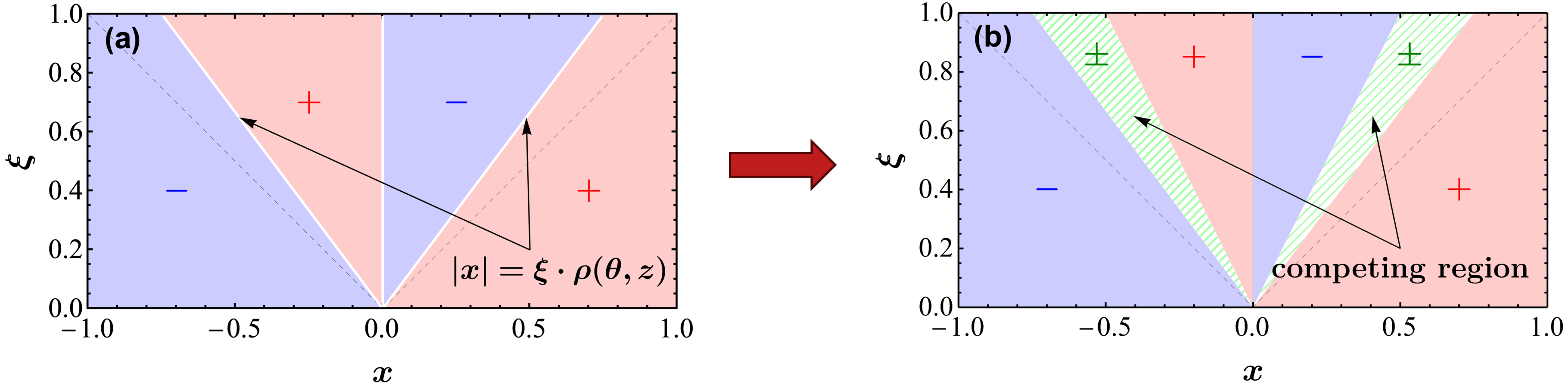
Complementary sensitivity



$N\,\pi \to N'\,\gamma\,\gamma$

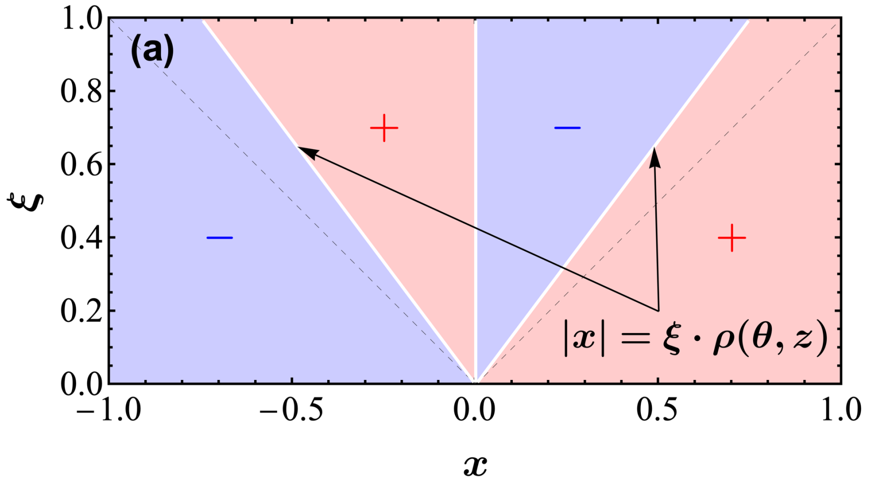
Improve GPD extraction with non-scaling integrals

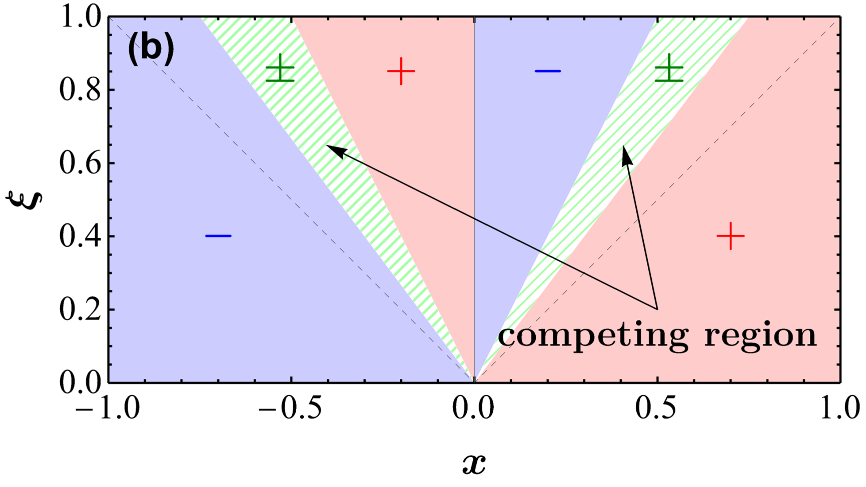
Non-scaling integral $\mathcal{M}^{[H]}(\xi, t, \theta) = \int_{-1}^{1} dx \, \frac{H^{+}(x,\xi,t)}{x - x_p(\xi,\theta) + i\epsilon}$ $\Rightarrow$ $\frac{\partial \chi^2}{\partial H(x,\xi,t)} \propto \frac{2x}{x^2 - x_p^2}$

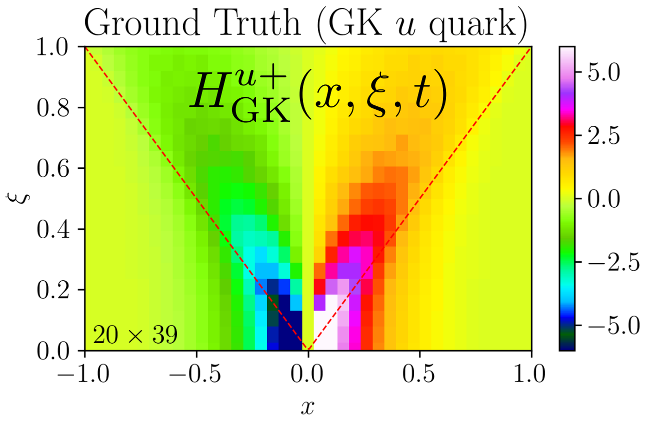


Improve GPD extraction with non-scaling integrals

Non-scaling integral $\mathcal{M}^{[H]}(\xi, t, \theta) = \int_{-1}^{1} dx \, \frac{H^+(x,\xi,t)}{x - x_p(\xi,\theta) + i\epsilon}$ → $\frac{\partial \chi^2}{\partial H(x,\xi,t)} \propto \frac{2x}{x^2 - x_p^2}$







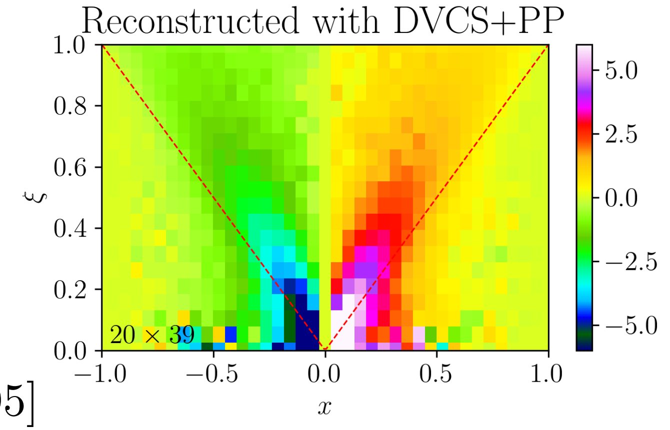
NN

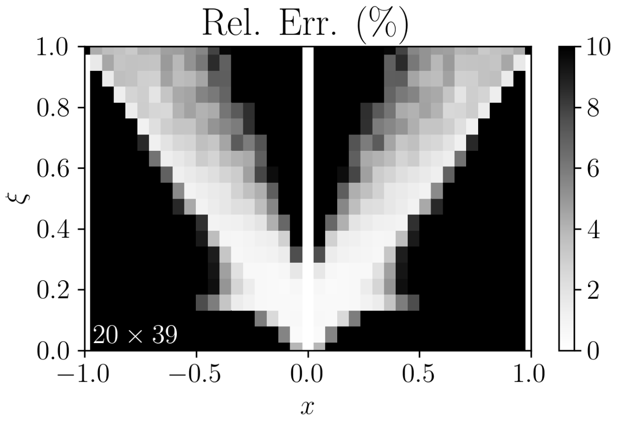
DVCS +

Photoproduction (PP)

$\cos\theta \in [-0.95, 0.95]$

$r = 0.01$



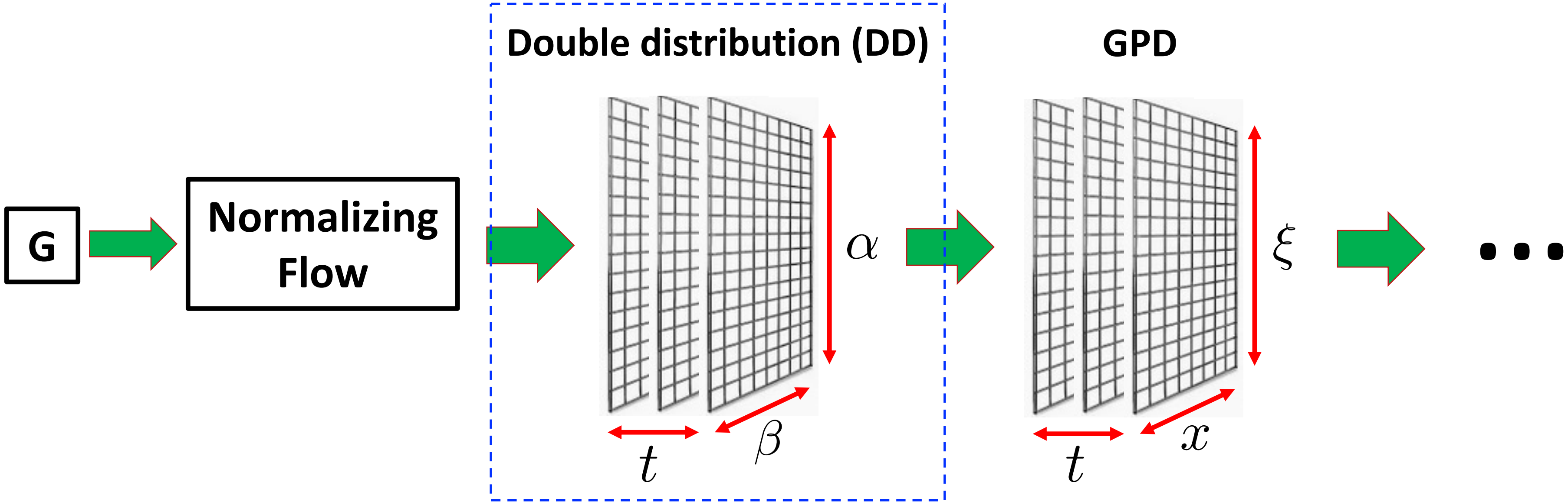


Pixelated construction of GPD with polynomiality [neglect the D-term]

GPDs satisfy polynomiality due to Lorentz covariance $\int_{-1}^{1} dx\, x^n\, H^q(x,\xi,t) = \sum_{i=0,2,\cdots}^{n} (2\xi)^i\, A^q_{n+1,i}(t)$

$\boldsymbol{H(x,\xi,t)}$ **at different** $\boldsymbol{\xi}$**'s are *not independent*!** → **Construct GPDs from double distribution (DD)**

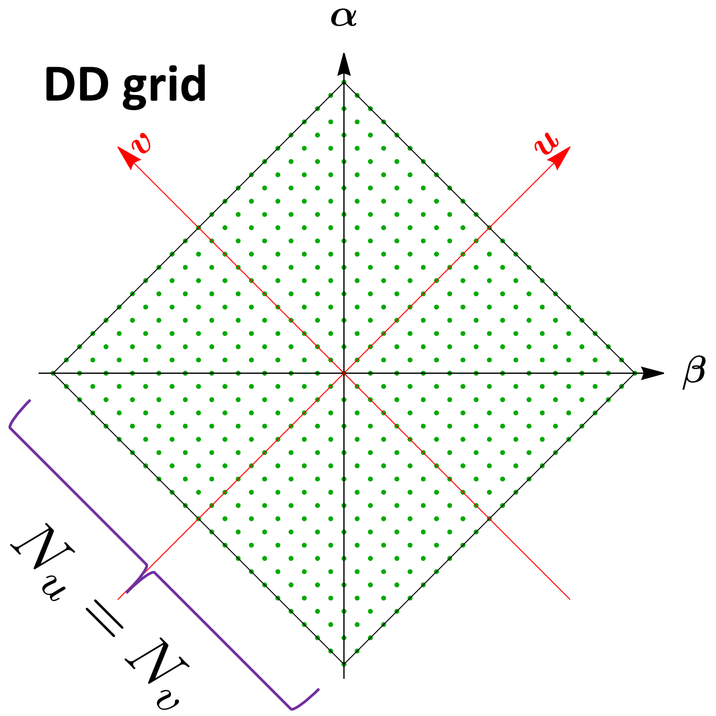
$$H^q(x,\xi,t) = \int_{-1}^{1} d\beta\, d\alpha\, \theta(1-|\alpha|-|\beta|)\, \delta(x-\beta-\xi\alpha)\, f^q(\beta,\alpha,t)$$ **Linear transformation!**

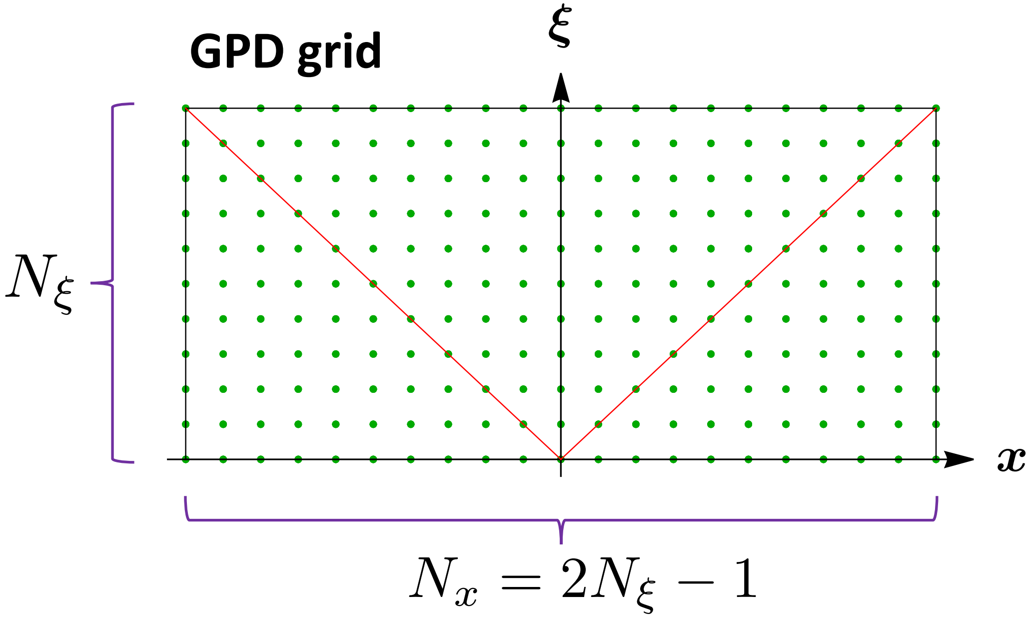


$$f_{\mathrm{NF}}(\beta,\alpha,t) = f_{\mathrm{GK}}(\beta,\alpha,t) * \epsilon(\beta,\alpha,t)$$

Choice of relative DD and GPD sizes

Conversion from DD to GPD is non-invertible! $$H^q(x,\xi,t) = \int_{-1}^{1} d\beta\, d\alpha\, \theta(1-|\alpha|-|\beta|)\,\delta(x-\beta-\xi\alpha)\, f^q(\beta,\alpha,t)$$



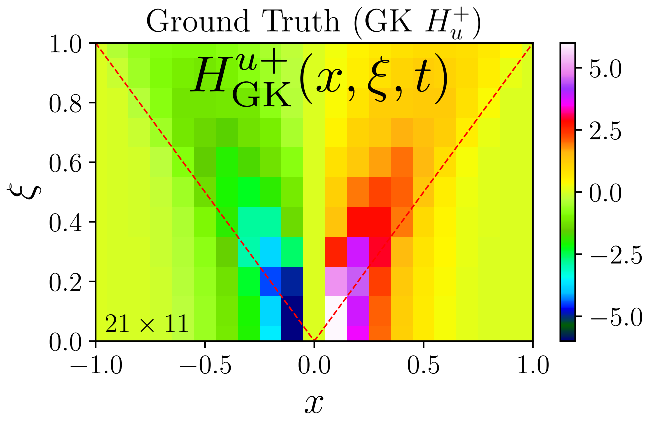


 # independent DD points: $N_{\mathrm{DD}} = N_u^2/2$ $>$ **# independent GPD points:** $N_{\mathrm{GPD}} = 2N_\xi^2$

Consider 3 scenarios: $$\frac{N_{\mathrm{DD}}}{N_{\mathrm{GPD}}} = \left(\frac{N_u}{2N_\xi}\right)^2 = \frac{1}{2},\ 1,\ 2$$

Polynomiality enhances sensitivity

Fix GPD to $(N_x \times N_\xi) = (21 \times 11)$



NN

DVCS

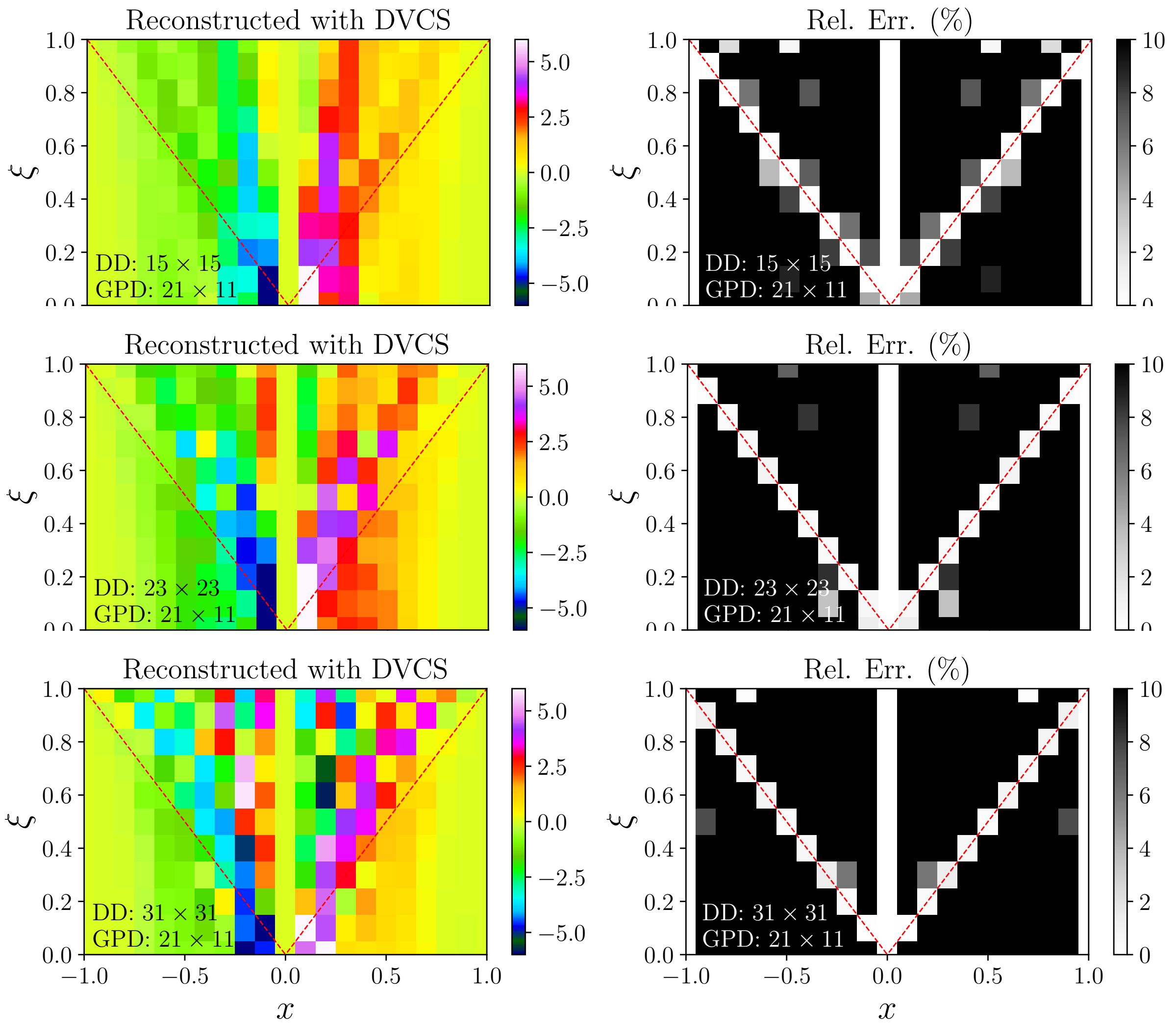
The extra DD layer induces extra inter-pixel correlation

$$\frac{N_{\mathrm{DD}}}{N_{\mathrm{GPD}}} = \frac{1}{2}$$

$$\frac{N_{\mathrm{DD}}}{N_{\mathrm{GPD}}} = 1$$

$$\frac{N_{\mathrm{DD}}}{N_{\mathrm{GPD}}} = 2$$

More faithful



Jefferson Lab

Extract with observables: DVCS + evolution

Evolution induces an array of observables in Q^2

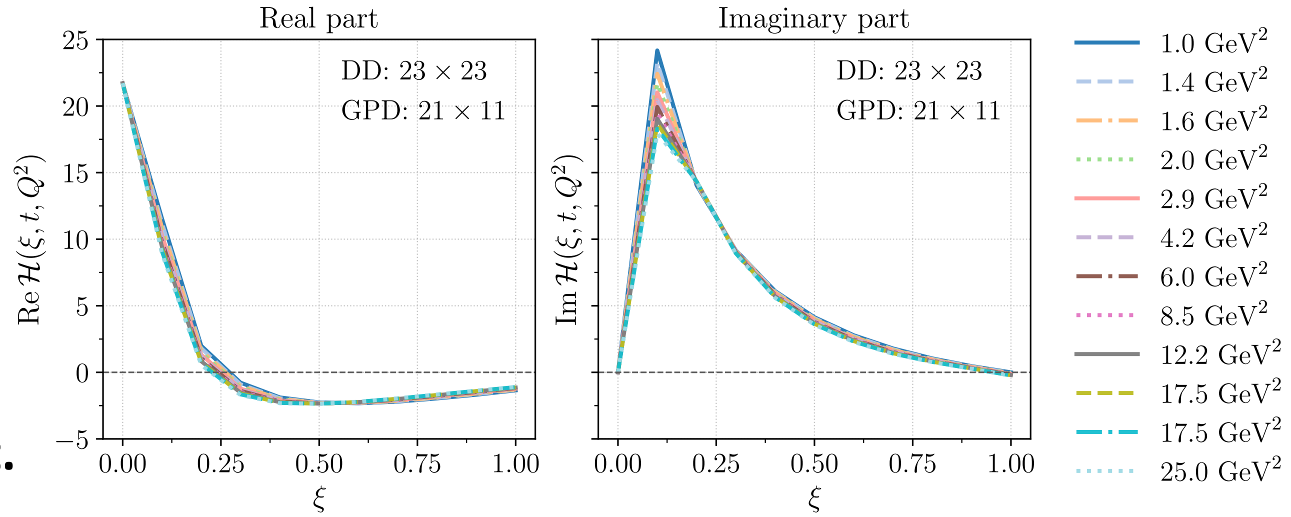
$$\mathcal{M}^{\mathrm{DVCS}'}(\xi, t, Q^2) = \int_{-1}^{1} dx\, \frac{H^{+}(x, \xi, t; \mu^2 = Q^2)}{x - \xi + i\epsilon}$$

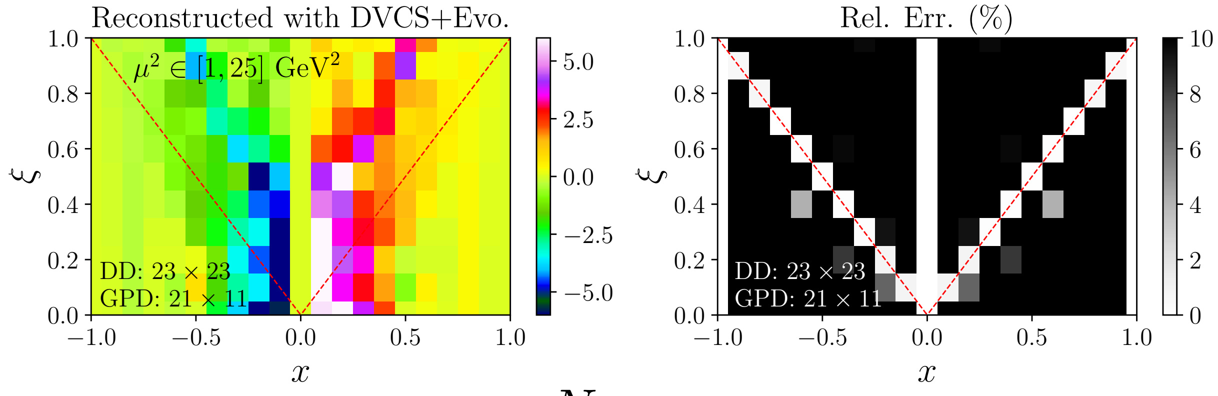
Breaks shadow GPD in principle …

… but quantitatively insignificant.

[Bertone et al. PRD `21]

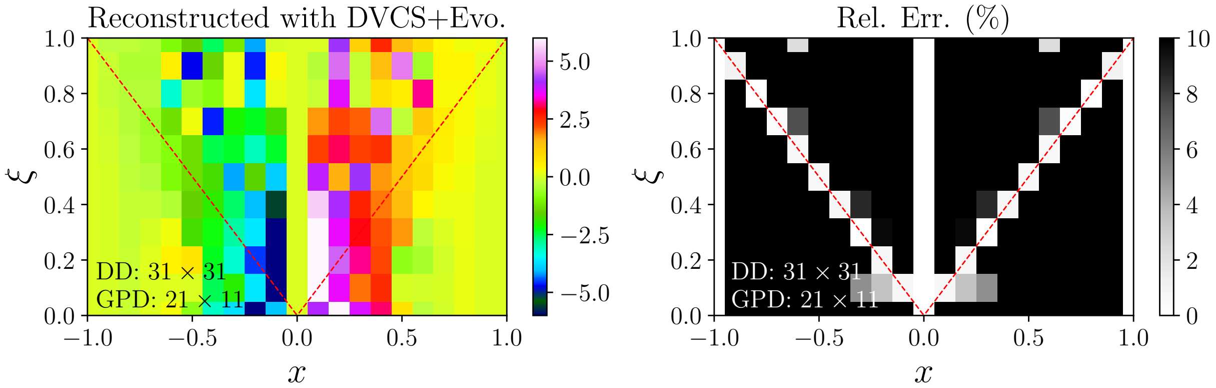
$$Q \in [1, 5]\ \mathrm{GeV}$$





$$\frac{N_{\mathrm{DD}}}{N_{\mathrm{GPD}}} = 1$$

$$\frac{N_{\mathrm{DD}}}{N_{\mathrm{GPD}}} = 2$$



Jefferson Lab

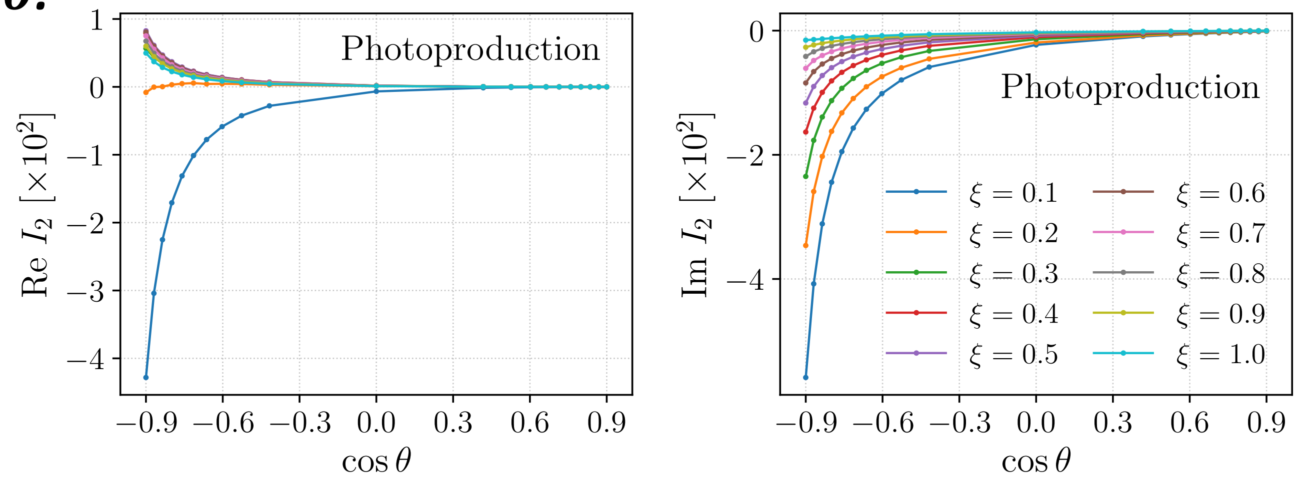
Extract with observables: DVCS + photoproduction integral

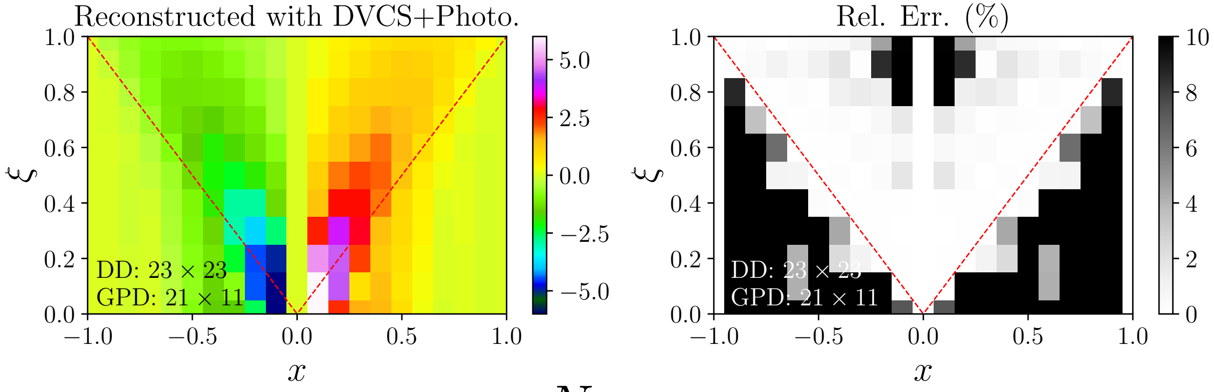
Non-scaling integral entangles x with the observed θ.

$$\mathcal{M}^{\rm NSI}(\xi,t,\theta) = \int_{-1}^{1} dx\, H^{+}(x,\xi,t)\, K(x,\xi,\theta)$$

Breaks shadow GPD more significantly.

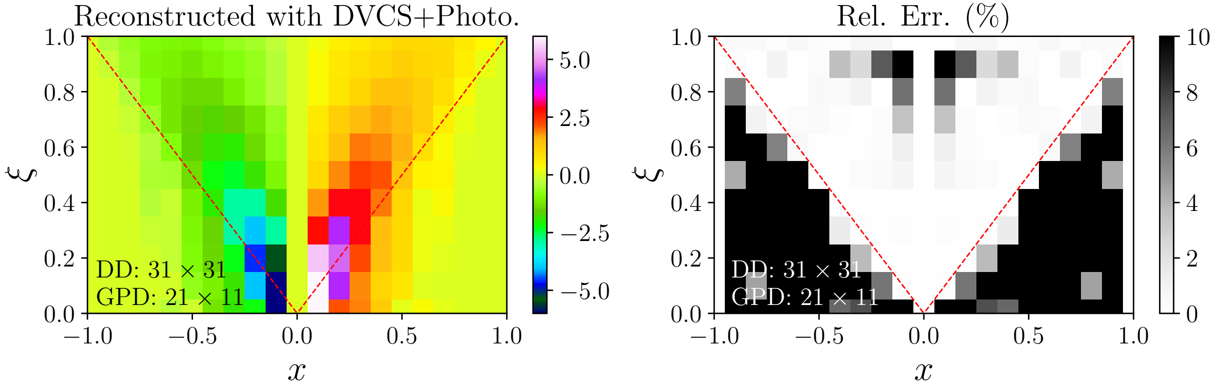
$$\cos\theta \in [-0.9, 0.9]$$





$$\frac{N_{\rm DD}}{N_{\rm GPD}} = 1$$

$$\frac{N_{\rm DD}}{N_{\rm GPD}} = 2$$



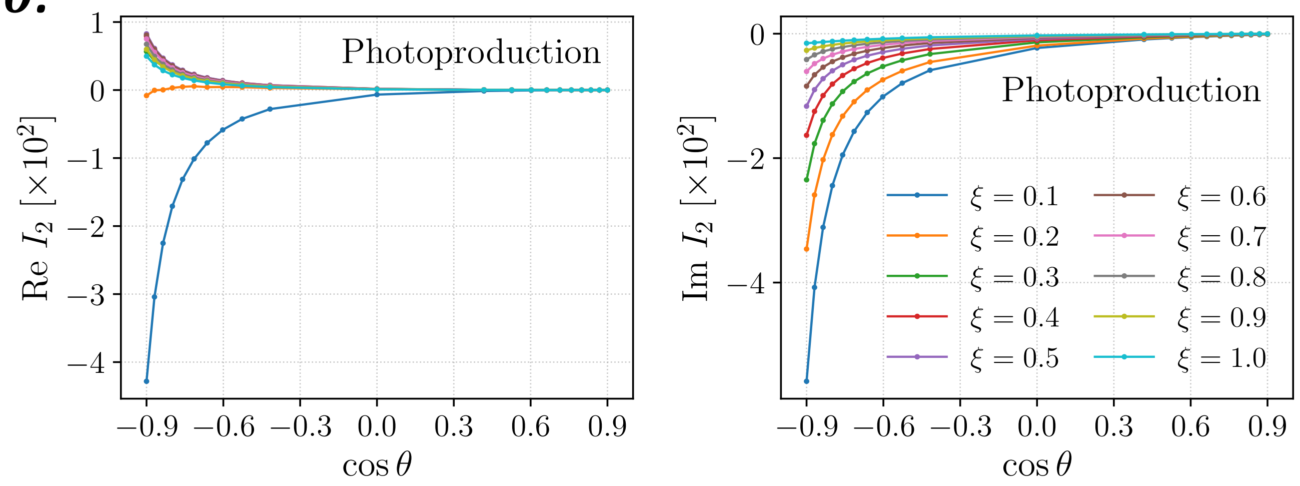
Extract with observables: DVCS + photoproduction integral

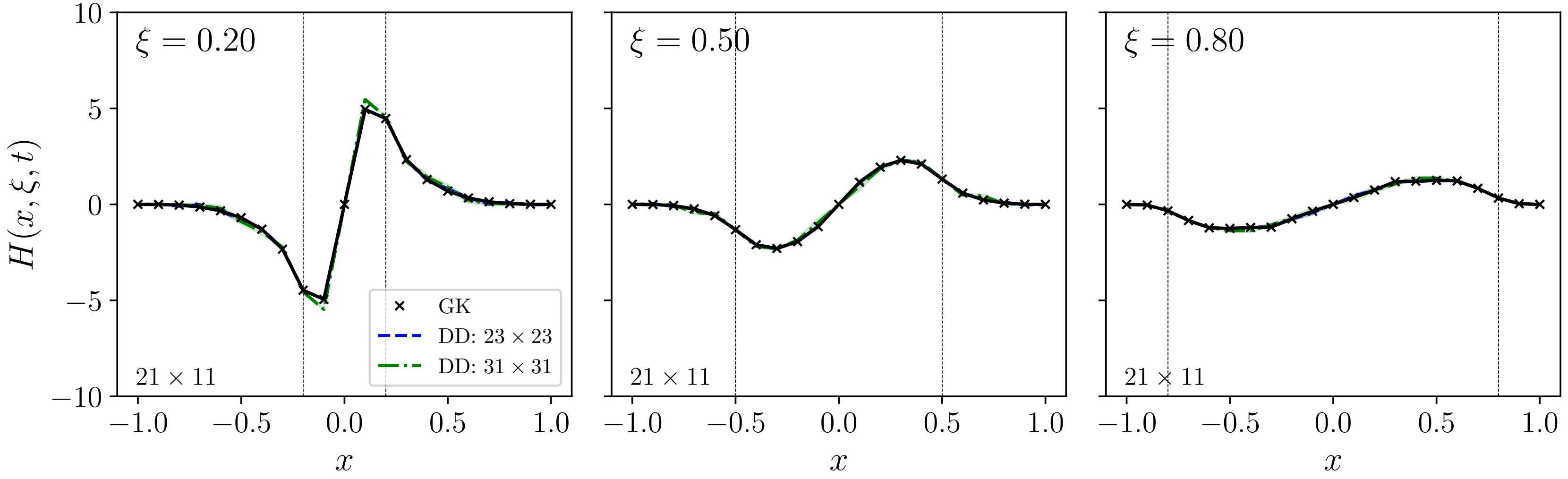
Non-scaling integral entangles x with the observed θ.

$$\mathcal{M}^{\mathrm{NSI}}(\xi,t,\theta)=\int_{-1}^{1}dx\,H^{+}(x,\xi,t)\,K(x,\xi,\theta)$$

Breaks shadow GPD more significantly.

$$\cos\theta\in[-0.9,0.9]$$





Jefferson Lab

Summary

- Extracting x-dependence of GPDs has difficulty from exclusiveness
- Pixelation + Normalizing Flow provides a way to *visualize* the fitting process
- Allows to address the sensitivity region, resolution, and model ambiguity
- *Scaling* integrals such as DVCS mostly constrain the ridge $x = \pm\xi$ (even with evolution)
- *Non-scaling* integrals are needed to constrain other regions
- Future: incorporate correlation observables from Lattice QCD

Thank you!