# Towards extraction of GPDs: Unbiased determination of DVCS Compton Form Factors

Paweł Sznajder
National Centre for Nuclear Research, Warsaw

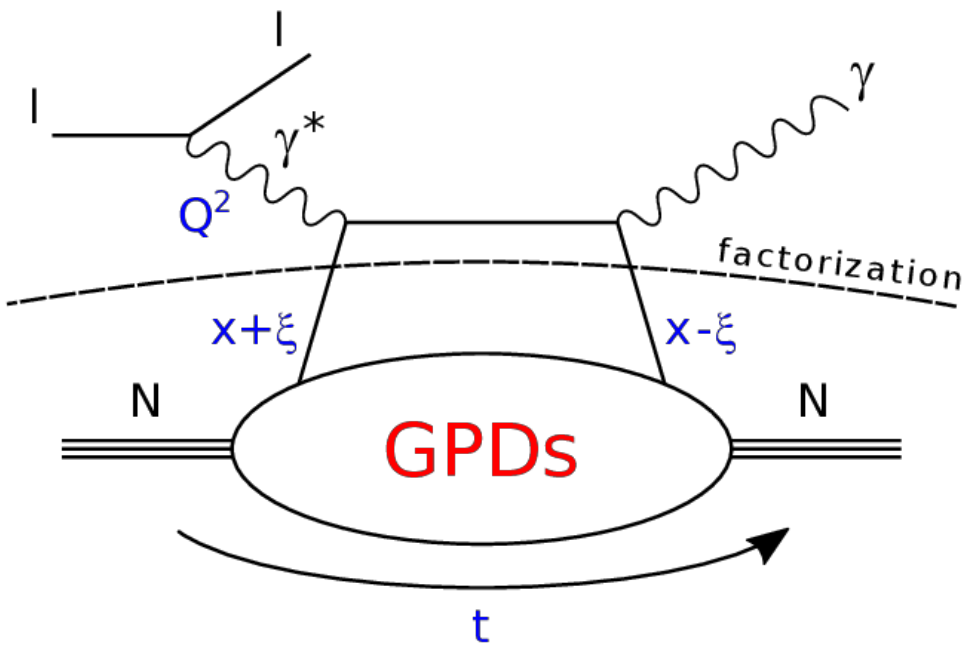
NATIONAL CENTRE FOR NUCLEAR RESEARCH
ŚWIERK

JLab, March 6, 2019

# Outline

- Introduction
- Methodology
- Global analysis of DVCS data - “classic” approach
- Global analysis of DVCS data - ANN approach
- Summary

## Deeply Virtual Compton Scattering (DVCS)

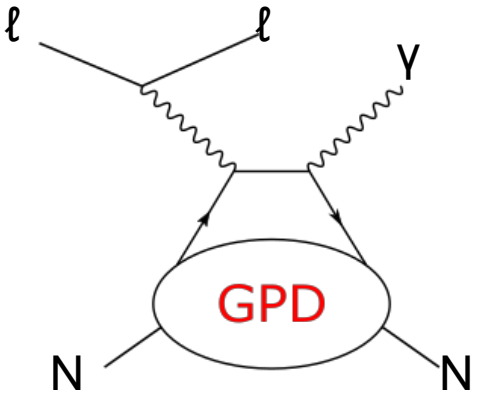



*factorization for $|t|/Q^2 \ll 1$*

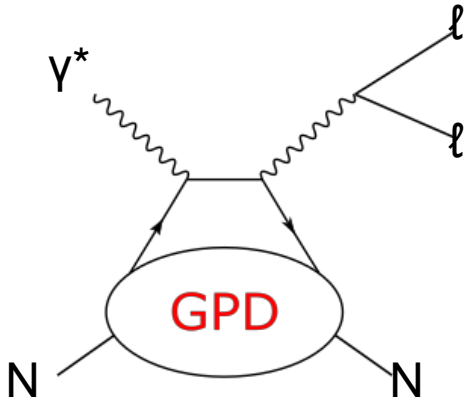
Chiral-even GPDs:
(helicity of parton conserved)

| | | |
|---|---|---|
| $H^{q,g}(x,\xi,t)$ | $E^{q,g}(x,\xi,t)$ | *for sum over parton helicities* |
| $\widetilde{H}^{q,g}(x,\xi,t)$ | $\widetilde{E}^{q,g}(x,\xi,t)$ | *for difference over parton helicities* |
| *nucleon helicity conserved* | *nucleon helicity changed* | |

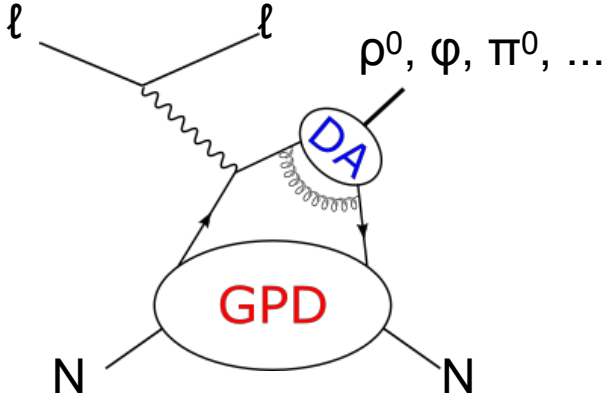
**GPDs accessible in various production channels and observables**
**→ experimental filters**




**DVCS**
*Deeply Virtual Compton Scattering*




**TCS**
*Timelike Compton Scattering*



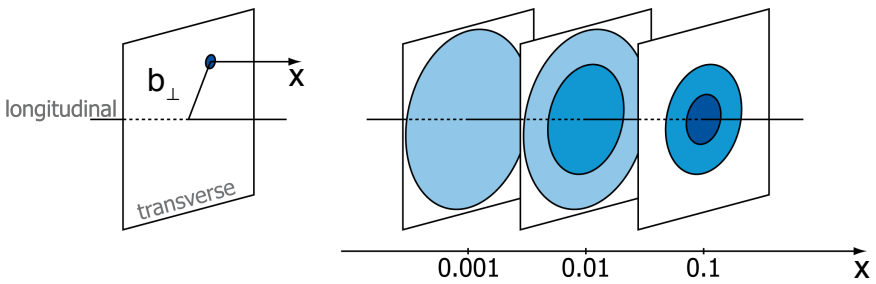

**HEMP**
*Hard Exclusive Meson Production*

*more production channels sensitive to GPDs exist!*

# Introduction

- Nucleon tomography

$$q(x, \mathbf{b}_\perp) = \int \frac{\mathrm{d}^2\mathbf{\Delta}}{4\pi^2} e^{-i\mathbf{b}_\perp \cdot \mathbf{\Delta}} H^q(x, 0, t = -\mathbf{\Delta}^2)$$



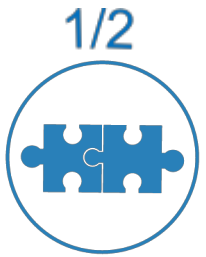

- Study of long. polarization with GPD $\tilde{H}$
- Study of distortion in transv. polarized nucleon with GPD E

Energy momentum tensor in terms of form factors:

$$\langle p', s'|\widehat{T}^{\mu\nu}|p, s\rangle = \bar{u}(p', s')\left[\frac{P^\mu P^\nu}{M}A(t) + \frac{\Delta^\mu\Delta^\nu - \eta^{\mu\nu}\Delta^2}{M}C(t) + M\eta^{\mu\nu}\bar{C}(t) + \frac{P^\mu i\sigma^{\nu\lambda}\Delta_\lambda}{4M}[A(t) + B(t) + D(t)] + \frac{P^\nu i\sigma^{\mu\lambda}\Delta_\lambda}{4M}[A(t) + B(t) - D(t)]\right]u(p, s)$$

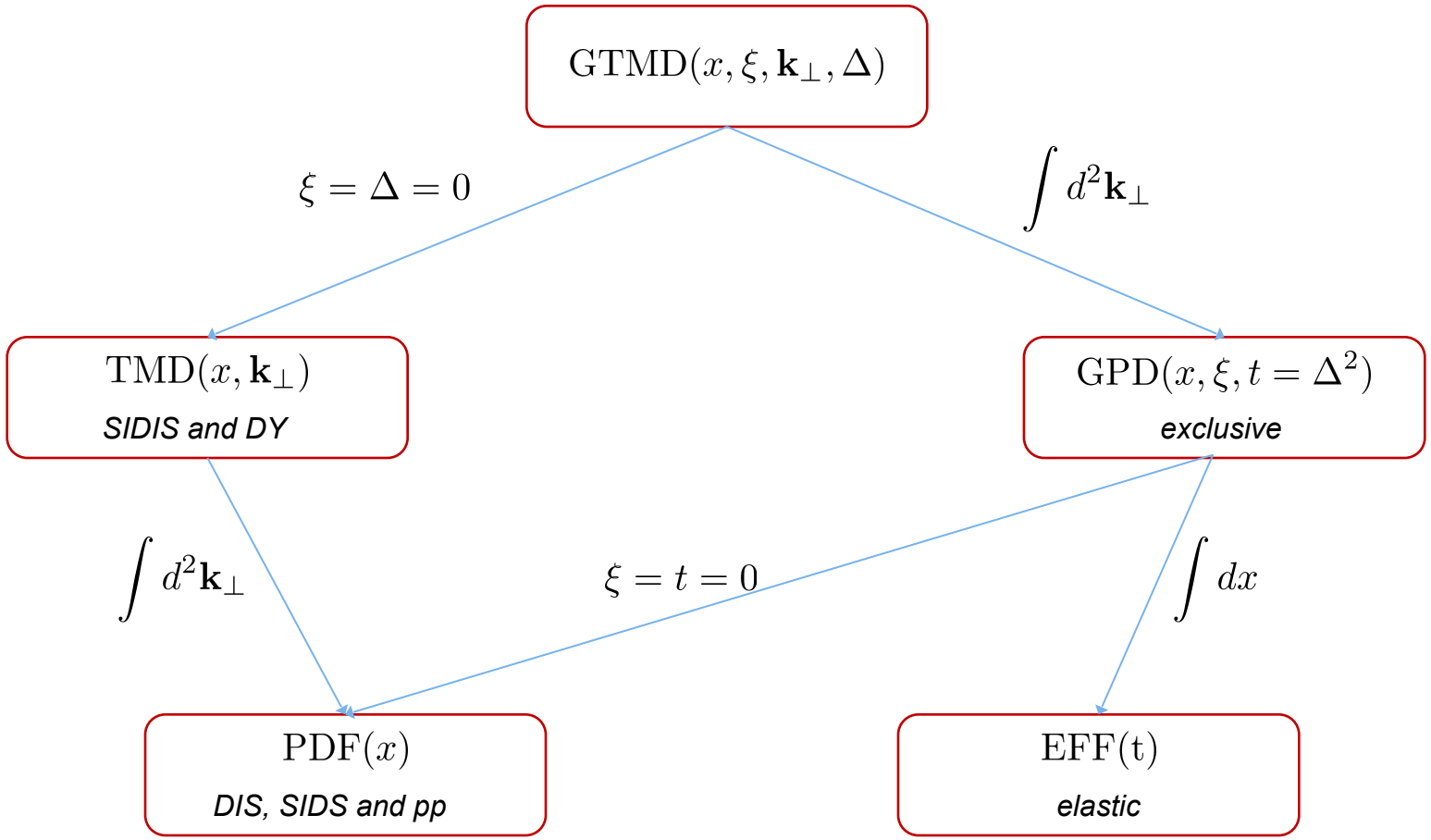
Access to total angular momentum and “mechanical” forces acting on quarks

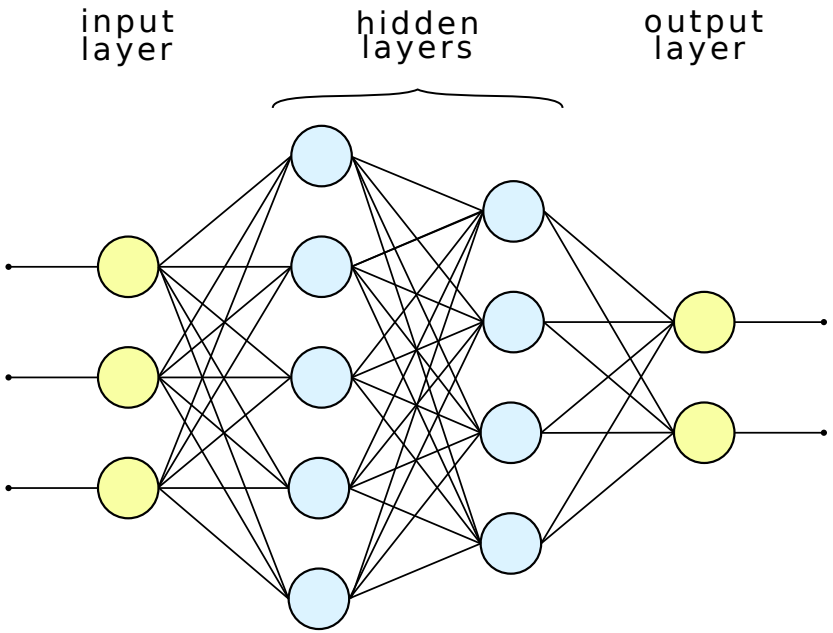
$$A^q(0) + B^q(0) = \int_{-1}^{1} x\,[H^q(x, \xi, 0) + E^q(x, \xi, 0)] = 2J^q$$




*Ji's sum rule*

# Introduction



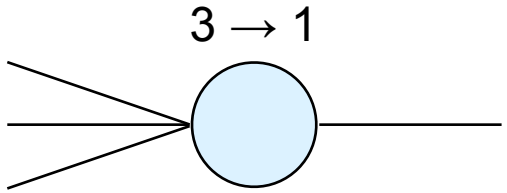


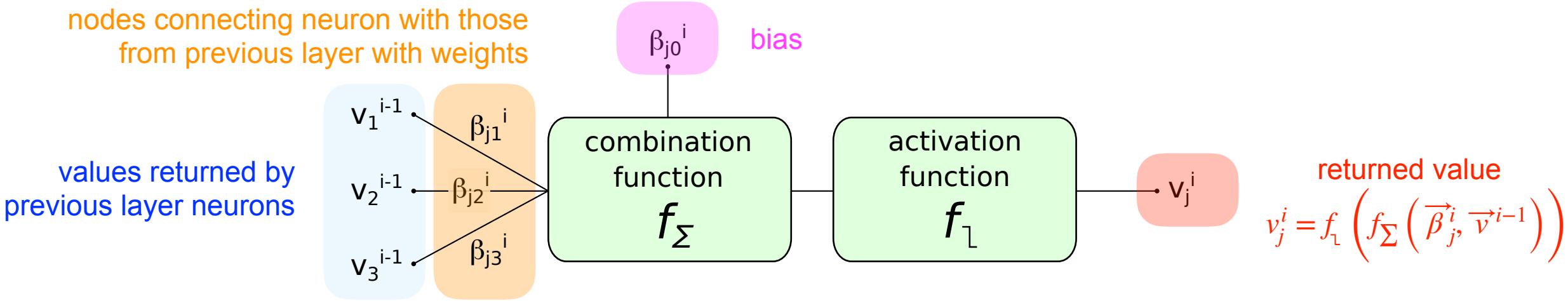

**Artificial neural networks (ANNs)**

- Data processing technique inspired by Nature
- Network made out of simple but highly connected elements → connectionists system
- Many variations, but most popular deep feed-forward ANNs
  - data processed layer by layer
    input layer → hidden layers→ output layer
  - generalisation capability given by consecutive hidden layers
    output of *i-1* layer is more refined than that of *i* layer
- Information containers

# Single neuron (perceptron)

layer *i*, neuron *j*




**Combination function**
*(information aggregation)*

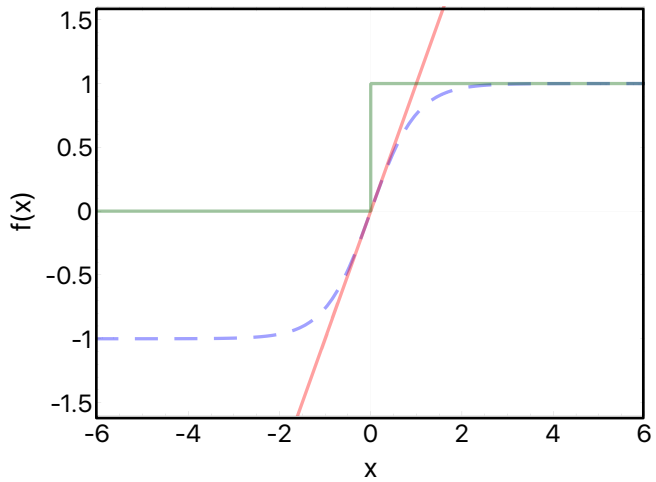
$$f_\Sigma\left(\vec{\beta}_j^i, \vec{v}^{i-1}\right) = \beta_{j0}^i + \sum_k v_k^{i-1}\beta_{jk}^i$$

**Activation function**
*(information processing)*

$f_\lambda(x) = \tanh(x)$

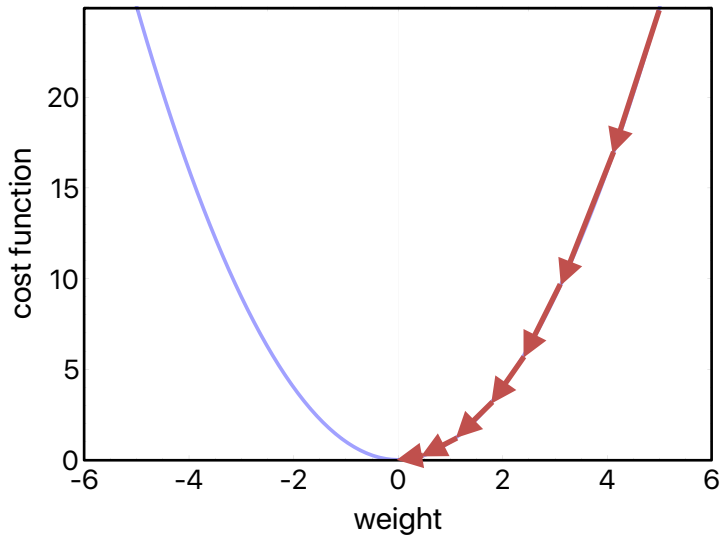
$f_\lambda(x) = x$

$f_\lambda(x) = \Theta(x)$



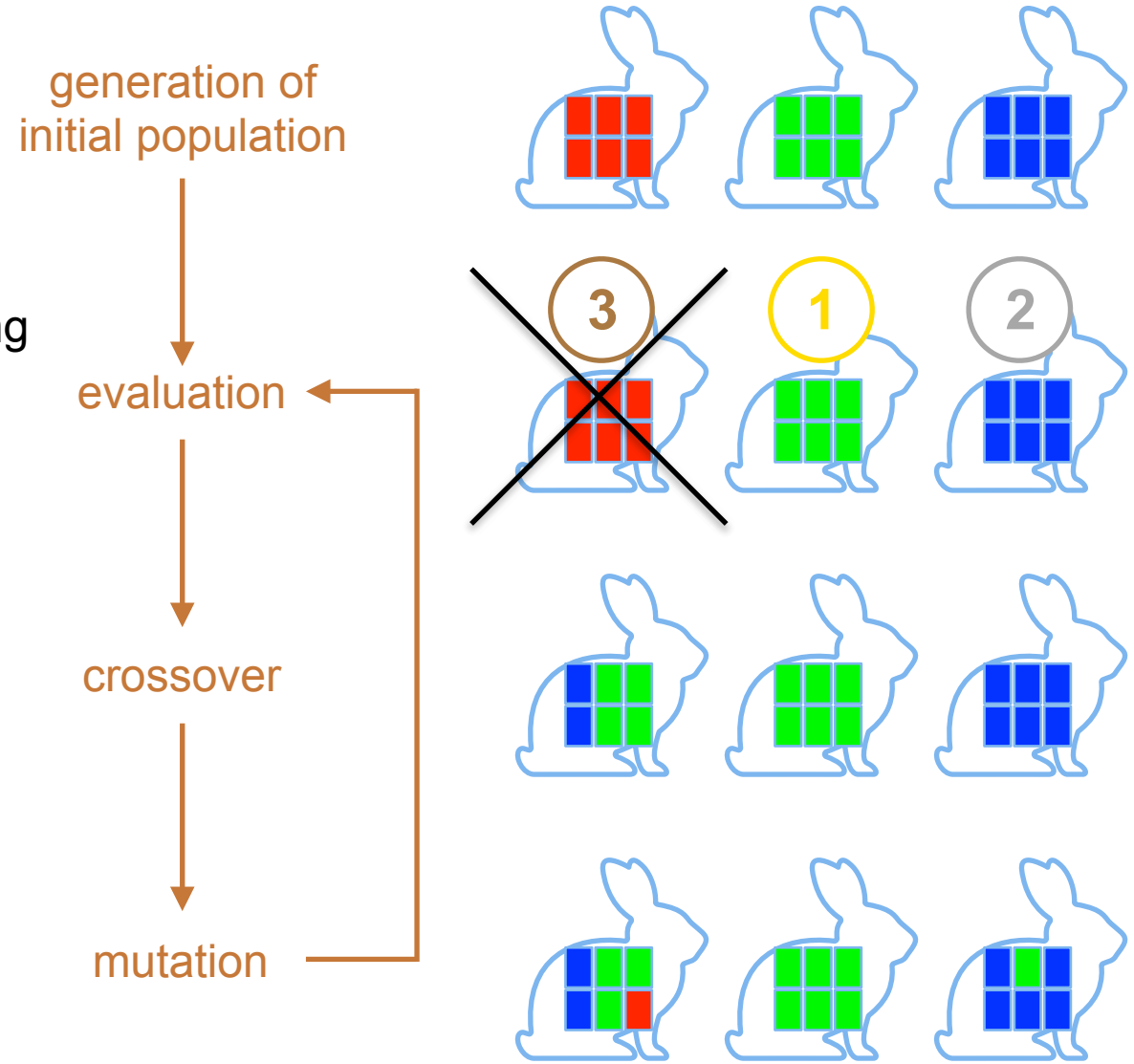
# Training

- Define training sample
- Define cost function, e.g. $\mathrm{RMSE} = \sum_i (v_i - \bar{v}_i)^2$
- Cost function analytically differentiable w.r.t. each parameter of the network
  → straightforward minimisation of cost function
- If not possible, e.g. because ANN not connected directly to training sample
  → compute gradient of cost function numerically
  → use genetic algorithm



# Genetic Algorithm

- Heuristic technique inspired by Nature
- Used in physics, mathematics, engineering, manufacturing
- Ability to
  - minimise w.r.t large number of free parameters
  - solve non-differentiable and non-linear problems
  - avoid local minima
- Iterative algorithm



# Analysis

H. Moutarde, P. S., J. Wagner "*Border and skewness functions from a leading order fit to DVCS data*"
Eur. Phys. J. C78 (2018) 11, 890

Goal: global extraction of Compton Form Factors (CFFs) from DVCS data using LO/LT formalism

Analysis done within PARTONS project

# Compton Form Factors

- imaginary part

$$Im\mathcal{G}(\xi,t) \quad = \pi G^{(+)}(\xi,\xi,t) = \pi \sum_q e_q^2 G^{q(+)}(\xi,\xi,t)$$

$$G^{q(+)}(x,\xi,t) = G^q(x,\xi,t) \mp G^q(-x,\xi,t)$$
$$G^{q(+)}(\xi,\xi,t) = G^{q_{\rm val}}(\xi,\xi,t) + 2G^{q_{\rm sea}}(\xi,\xi,t)$$

"−" for $G \in \{H, E\}$
"+" for $G \in \{\widetilde{H}, \widetilde{E}\}$

- real part

$$Re\mathcal{G}(\xi,t) = \text{P.V.} \int_0^1 G^{(+)}(x,\xi,t) \left( \frac{1}{\xi - x} \mp \frac{1}{\xi + x} \right) \mathrm{d}x$$

$$Re\mathcal{G}(\xi,t) = \text{P.V.} \int_0^1 G^{(+)}(x,x,t) \left( \frac{1}{\xi - x} \mp \frac{1}{\xi + x} \right) \mathrm{d}x + C_G(t)$$

$$C_H(t) = -C_E(t) \qquad C_{\widetilde{H}}(t) = C_{\widetilde{E}}(t) = 0$$

connected to EMT FF

# Subtraction Constant

Subtraction constant as analytic continuation of Mellin moments to *j* = -1

$$C_G^q(t) = C_{G,-1}^q(t) = 2\int_{(0)}^{1}\left(G^{q(+)}(x,x,t) - G^{q(+)}(x,0,t)\right)\frac{1}{x}dx$$

Analytic regularization prescription

$$\int_{(0)}^{1}\frac{f(x)}{x^{a+1}} = \int_{0}^{1}\frac{f(x)-f(0)-xf'(0)-\dots}{x^{a+1}} + f(0)\int_{(0)}^{1}\frac{dx}{x^{a+1}} + f'(0)\int_{(0)}^{1}\frac{dx}{x^{a}} + \dots =$$

$$\int_{0}^{1}\frac{f(x)-f(0)-xf'(0)-\dots}{x^{a+1}} - \frac{f(0)}{a} - \frac{f'(0)}{a-1} + \dots$$

applicable if f(x) analytic and not-vanishing at x = 0

# Ansatz for H and $\tilde{H}$

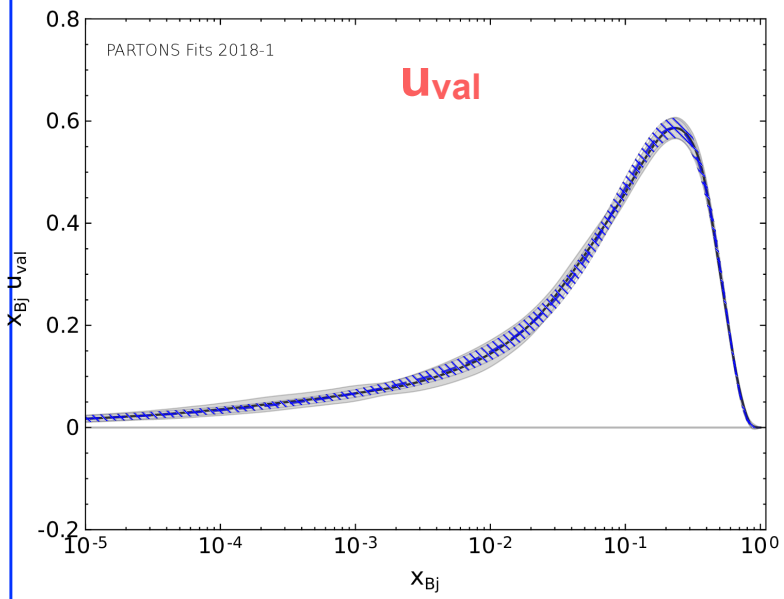
$$G^q(x,0,t) = \mathrm{pdf}_G^q(x)\ \exp(f_G^q(x)t) \qquad f_G^q(x) = A_G^q \log(1/x) + B_G^q(1-x)^2 + C_G^q(1-x)x$$

- reduction to PDFs and correspondence to EFFs
- modify "classical" log(1/x) term by $B_G^q(1\text{-}x)^2$ in low-x and by $C_G^q(1\text{-}x)x$ in high-x regions
- polynomials found in analysis of EFF data → good description of data
- allow to use the analytic regularisation prescription
- finite proton size at x → 1

$$G^q(x,x,t) = G^q(x,0,t)\ g_G^q(x,x,t) \qquad g_G^q(x,x,t) = \frac{a_G^q}{(1-x^2)^2}\,(1 + t(1-x)(b_G^q + c_G^q \log(1+x)))$$

- at x → 0 constant skewness effect
- at x → 1 reproduce power behaviour predicted for GPDs in Phys. Rev. D69, 051501 (2004)
- t-dependence similar to DD-models with (1-x) to avoid any t-dep. at x = 1

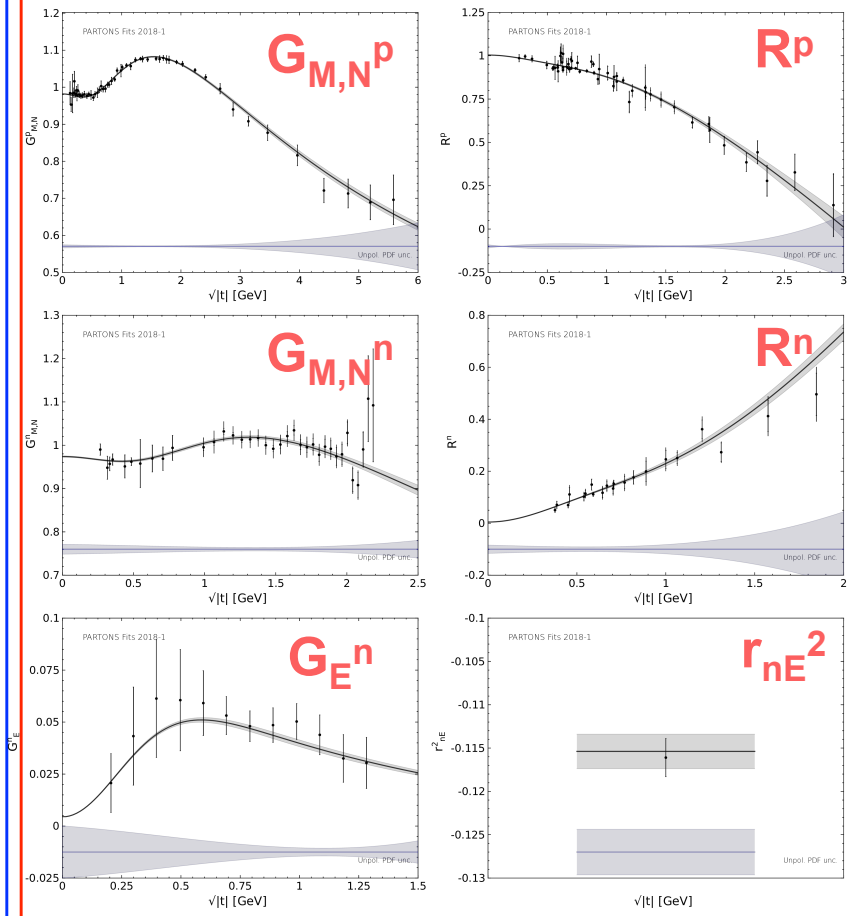
## 1. Analysis of PDF parameterisations




$$\mathrm{pdf}(x, Q^2) = x^{-g(\delta_p, \delta_q, Q^2)}(1-x)^{\alpha} \times \sum_{i=0}^{4} g(p_i, q_i, Q^2) x^i$$

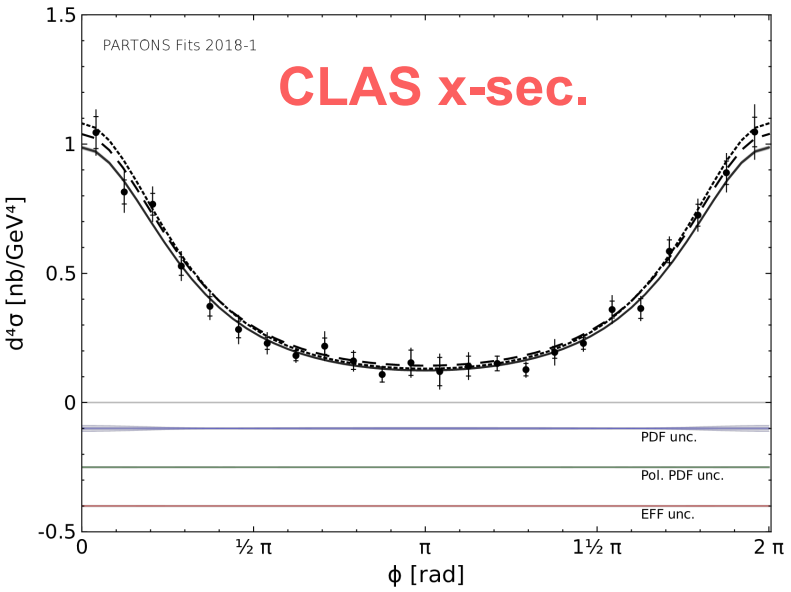
$$g(p, q, Q^2) = p + q \log \frac{Q^2}{Q_0^2}$$

## 2. Analysis of Elastic Form Factor data




$$\chi^2/\mathrm{ndf} = 129.6/(178-9) \approx 0.77$$

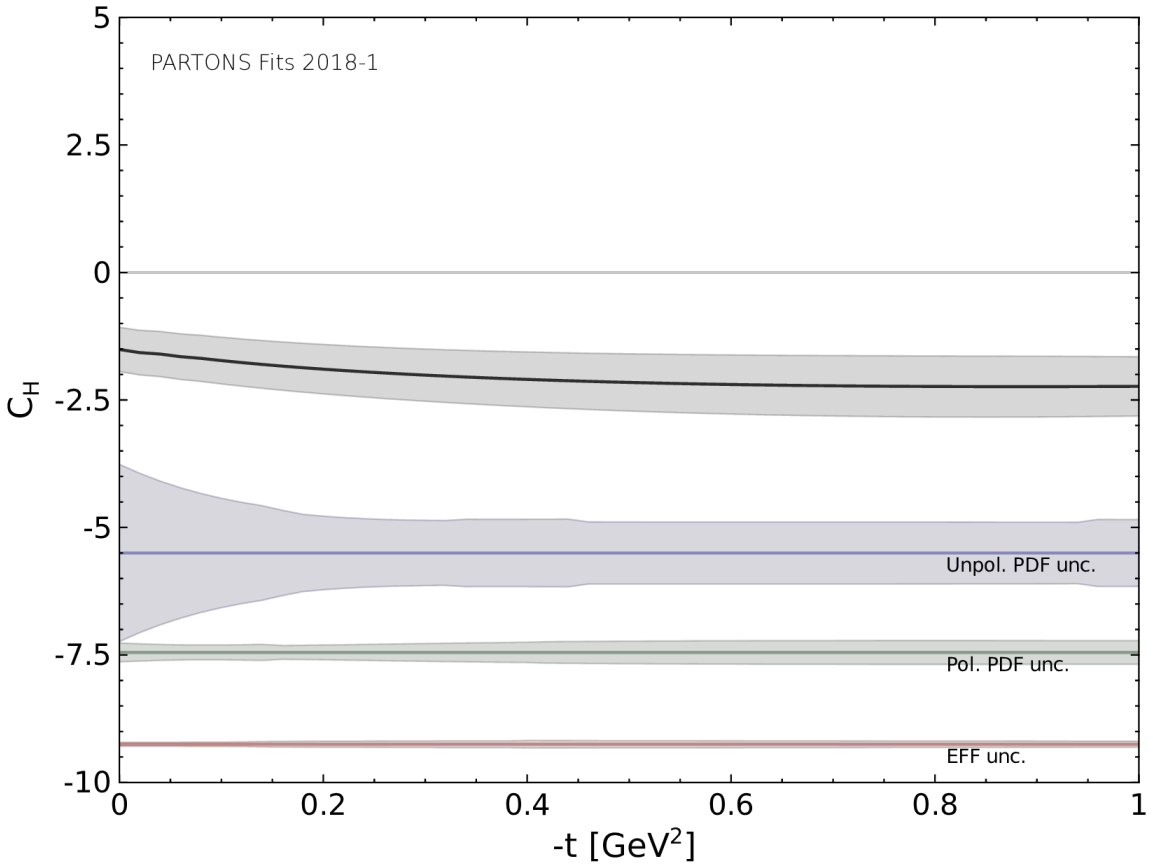
## 3. Analysis of DVCS data



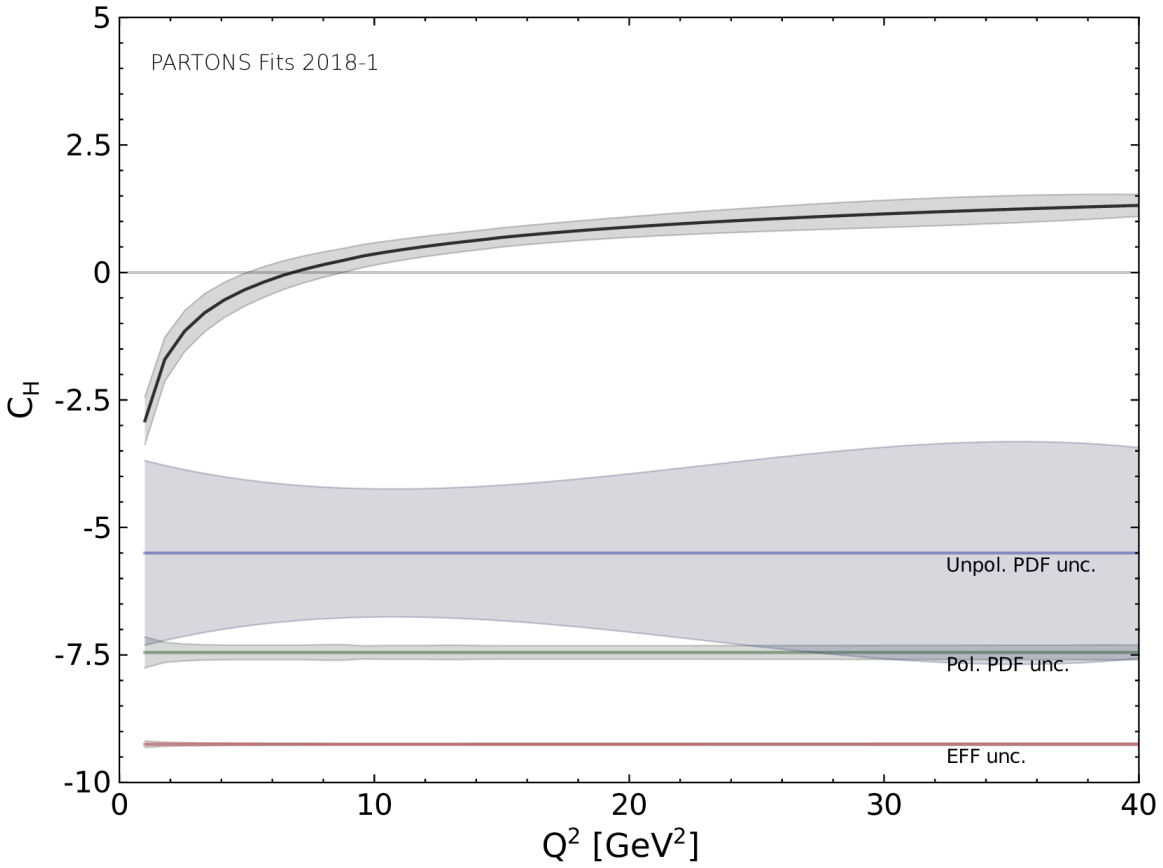

$$\chi^2/\mathrm{ndf} = 2346.3/(2600-13) \approx 0.91$$

# Results

Subtraction constant:



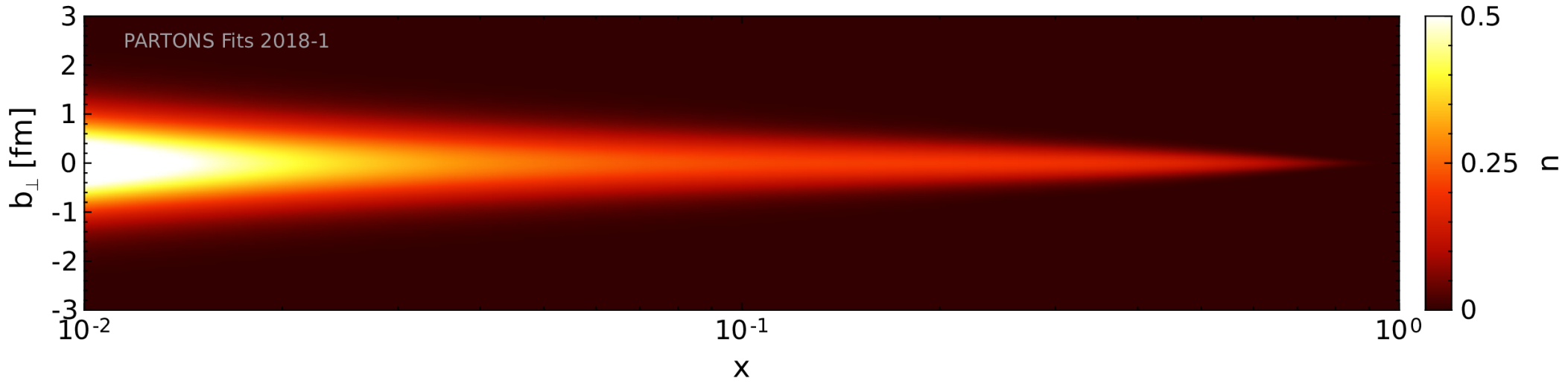

$Q^2 = 2\ \mathrm{GeV}^2$

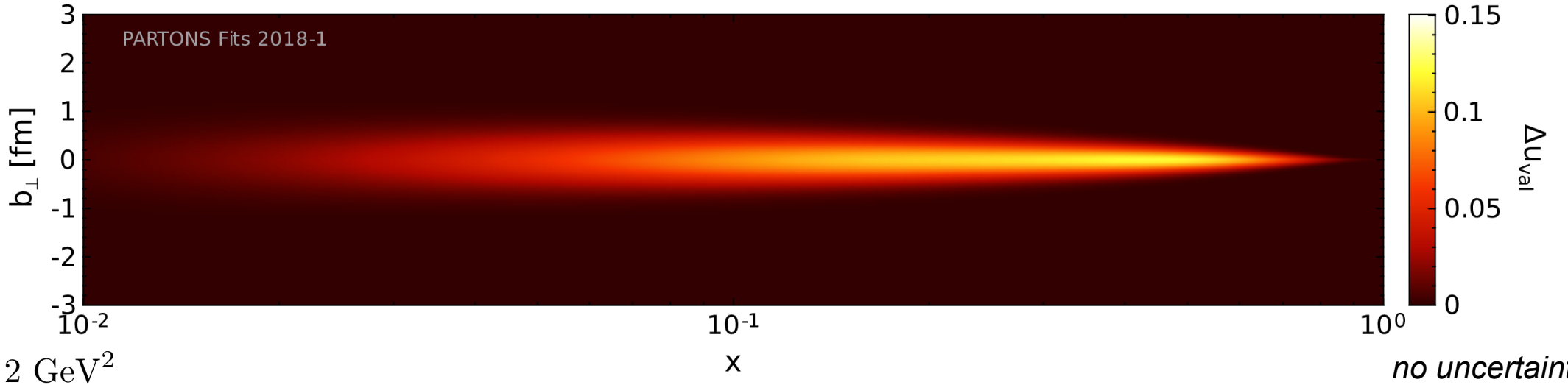



$t = 0$

# Results

Nucleon tomography:



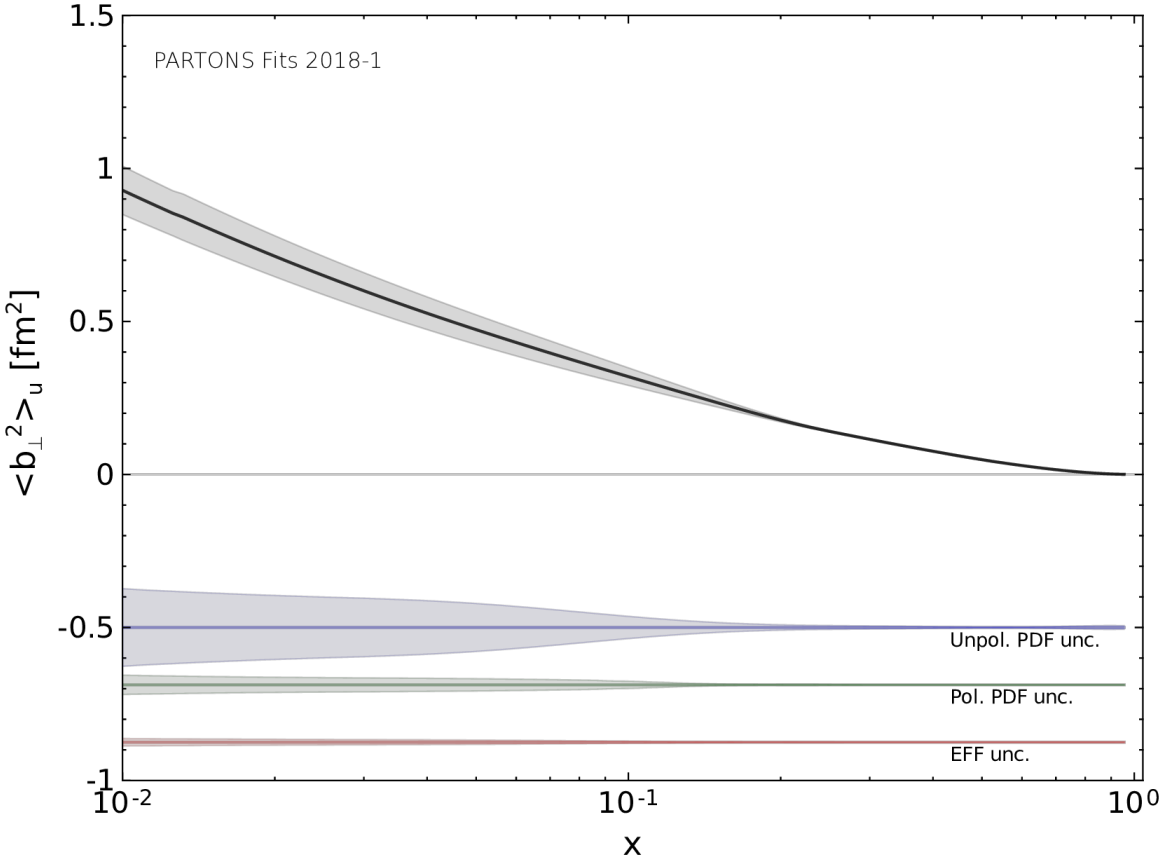

$Q^2 = 2\ \mathrm{GeV}^2$

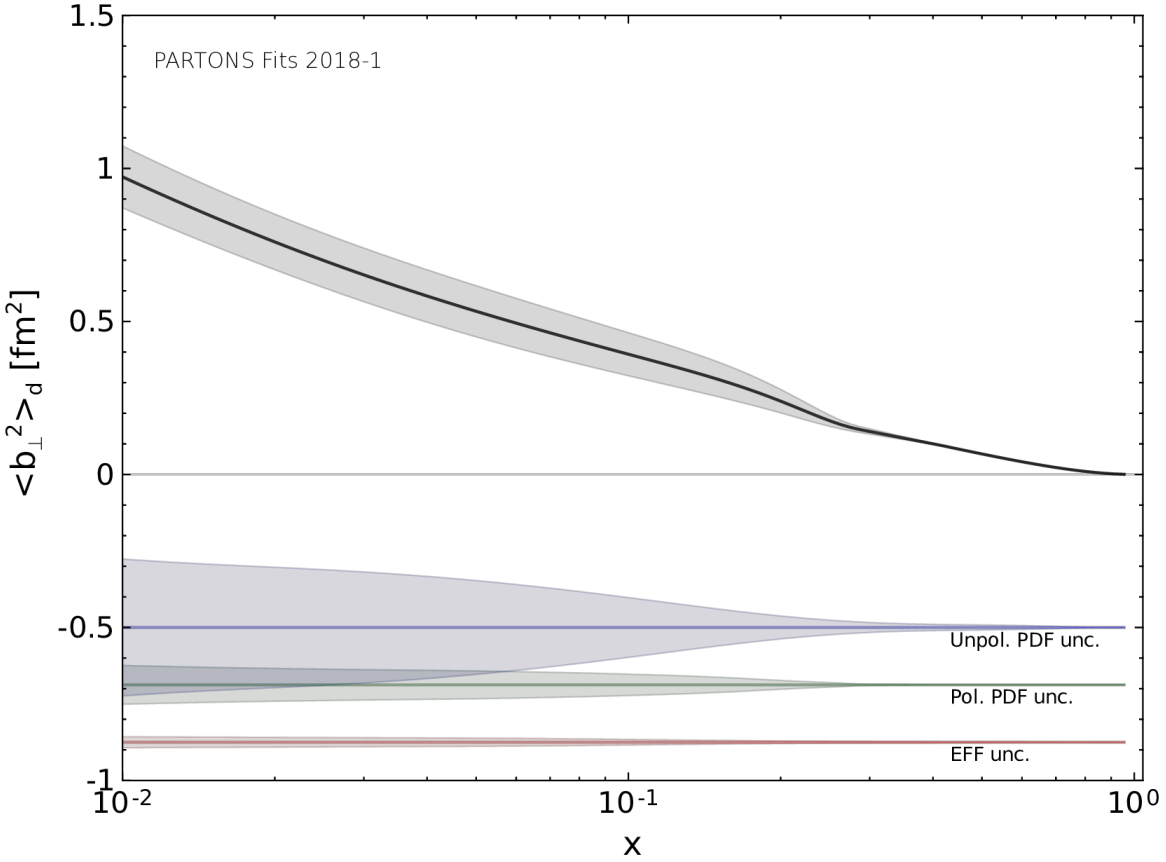
*no uncertainties!*

# Results

Nucleon tomography:

$$\langle b_\perp^2 \rangle_q(x) = \frac{\int d^2\mathbf{b}_\perp \; \mathbf{b}_\perp^2 q(x, \mathbf{b}_\perp)}{\int d^2\mathbf{b}_\perp \; q(x, \mathbf{b}_\perp)}$$




$Q^2 = 2 \text{ GeV}^2$

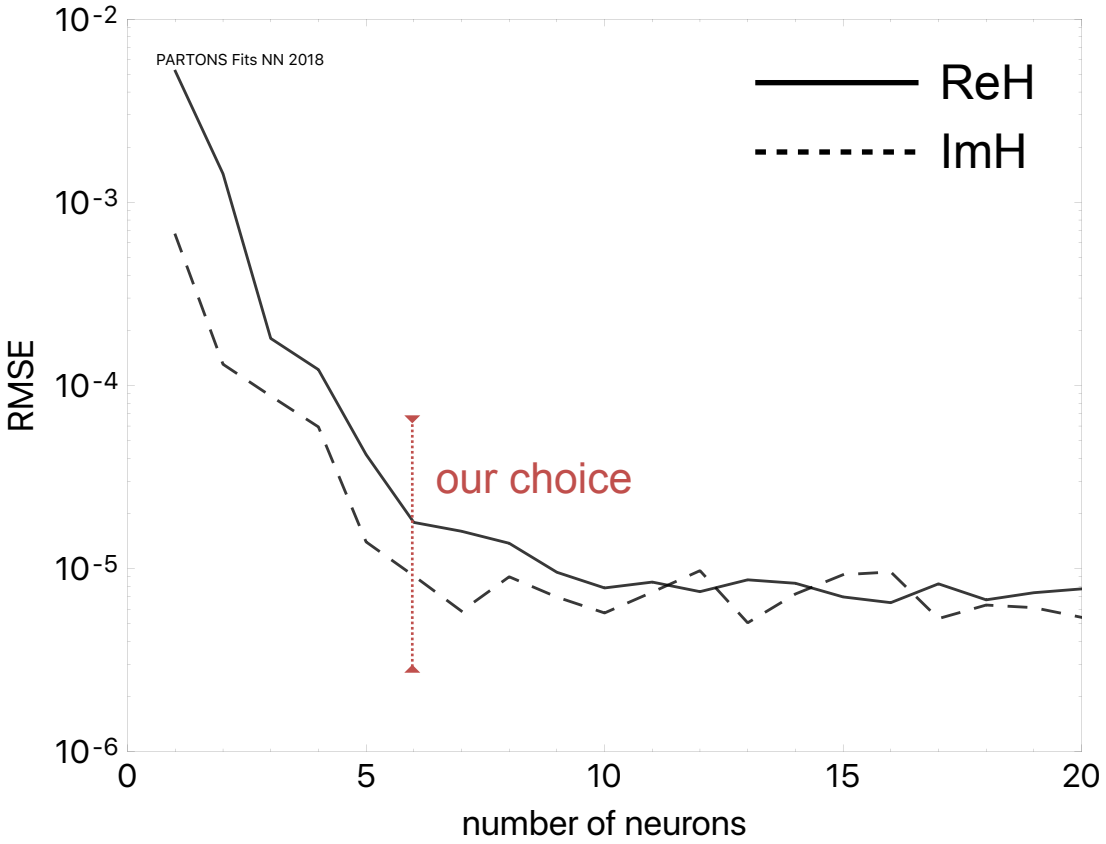
# Analysis

H. Moutarde, P. S., J. Wagner "*Unbiased determination of Compton Form Factors*"
preliminary results

Goal: global extraction of Compton Form Factors (CFFs) from DVCS data using ANN technique

Analysis done within PARTONS project

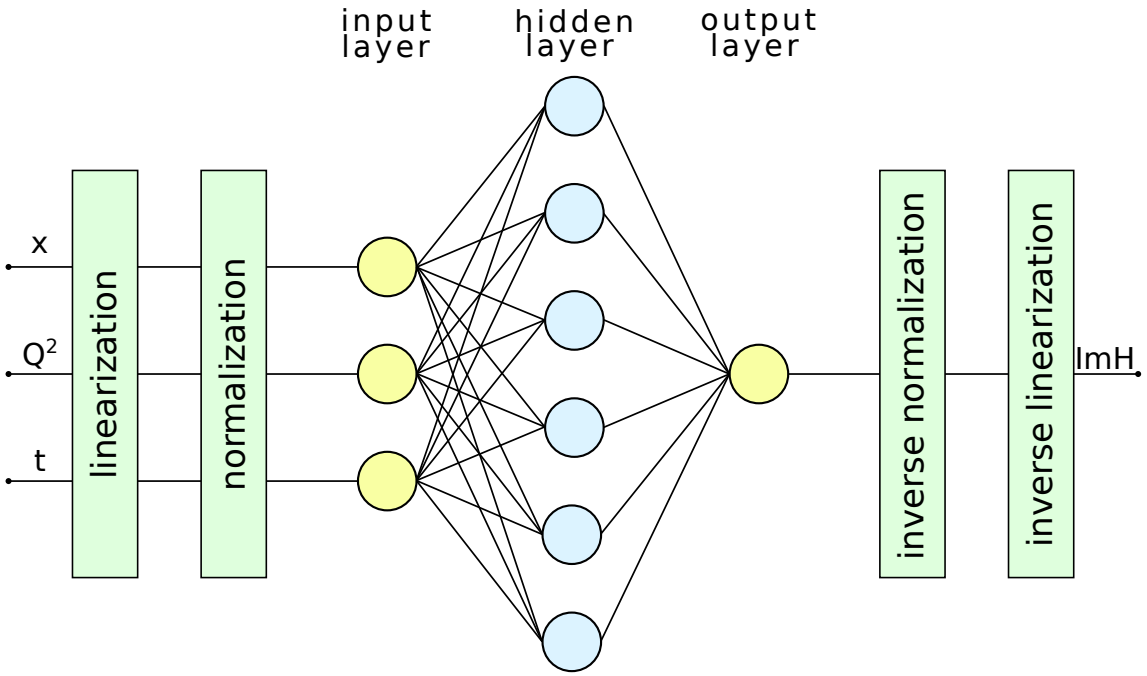
# NN architecture




**How many perceptrons needed to represent CFF sample?**

Exercise done with 1000 randomly generated CFF points using GK GPD model

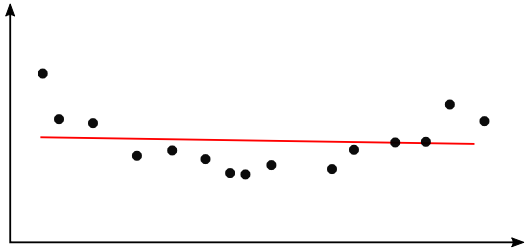
Input and output variables are normalised and linearised



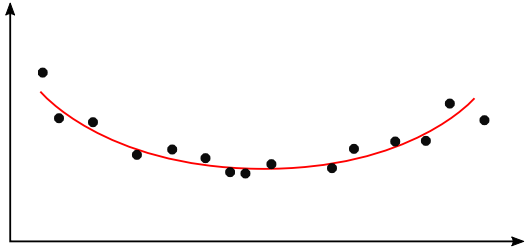
# Regularisation

## Need for regularisation:

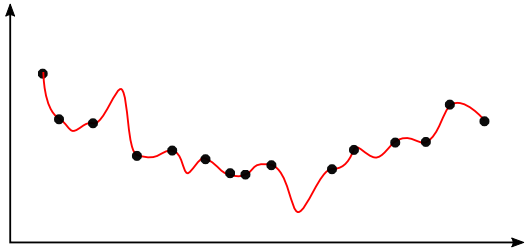
(a) under-fitting
bad description of training sample
no prediction power



(b) proper fitting
good description of training sample
prediction power

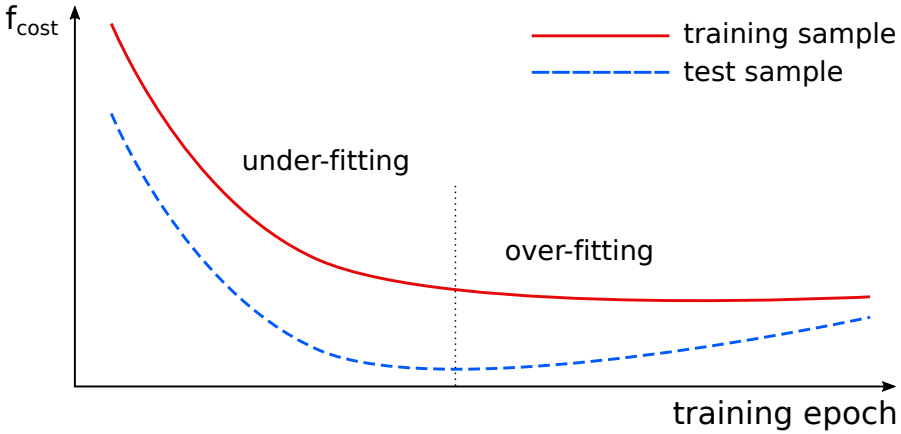


(c) over-fitting
good description of training sample
no prediction power



## Supervised training:

- Divide your sample into training and test subsamples (e.g. 3:1)
  - use training subsample for ANN training
  - used test subsample to detect over-fitting
- Many regularisation techniques available
  - early stopping used in this analysis



# Input data

Performance:

$$\chi^2/\text{nPoints} = 2243.5/2624 \approx 0.85$$

| No. | Collab. | Year | Ref. | $\chi^2$ | $n$ | $\chi^2/n$ |
|---|---|---|---|---|---|---|
| 1 | HERMES | 2001 | [17] | 10.7 | 10 | 1.07 |
| 2 | | 2006 | [18] | 5.5 | 4 | 1.38 |
| 3 | | 2008 | [19] | 18.5 | 18 | 1.03 |
| 4 | | 2009 | [20] | 34.7 | 35 | 0.99 |
| 5 | | 2010 | [21] | 40.7 | 18 | 2.26 |
| 6 | | 2011 | [22] | 16.7 | 24 | 0.70 |
| 7 | | 2012 | [23] | 22.4 | 35 | 0.64 |
| 8 | CLAS | 2001 | [24] | — | 0 | — |
| 9 | | 2006 | [25] | 1.0 | 2 | 0.52 |
| 10 | | 2008 | [26] | 376.4 | 283 | 1.33 |
| 11 | | 2009 | [27] | 28.3 | 22 | 1.29 |
| 12 | | 2015 | [28] | 306.6 | 311 | 0.99 |
| 13 | | 2015 | [29] | 884.7 | 1333 | 0.66 |
| 14 | Hall A | 2015 | [15] | 231.8 | 228 | 1.02 |
| 15 | | 2017 | [16] | 211.4 | 276 | 0.77 |
| 16 | COMPASS | 2018 | [30] | 3.0 | 2 | 1.50 |
| 17 | ZEUS | 2009 | [31] | 5.49 | 4 | 1.38 |
| 18 | H1 | 2005 | [32] | 22.2 | 7 | 3.17 |
| 19 | | 2009 | [33] | 23.4 | 12 | 1.95 |

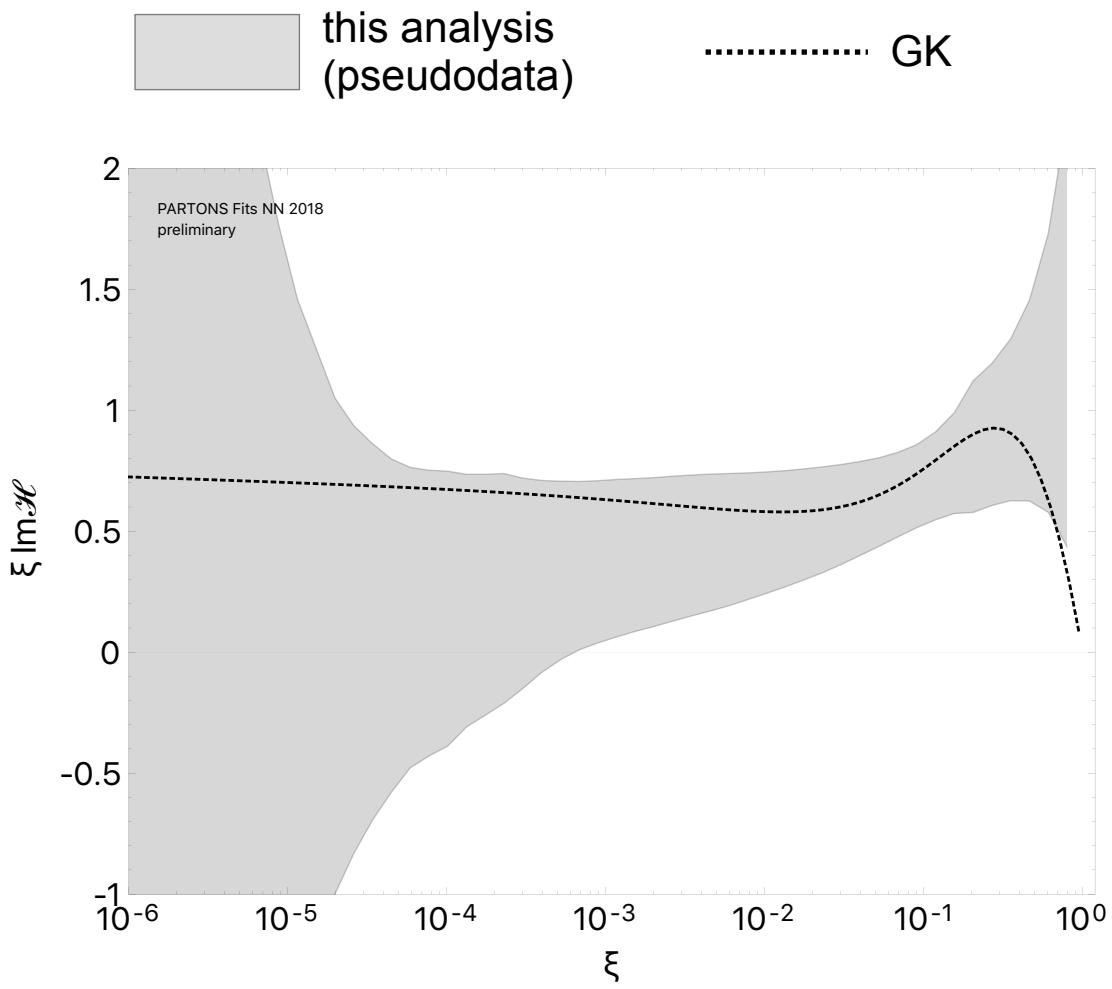
| No. | Collab. | Year | Ref. | Observable | | Kinematic dependence | No. of points used / all |
|---|---|---|---|---|---|---|---|
| 1 | HERMES | 2001 | [17] | $A_{LU}^{+}$ | | $\phi$ | 10 / 10 |
| 2 | | 2006 | [18] | $A_C^{\cos i\phi}$ | $i=1$ | $t$ | 4 / 4 |
| 3 | | 2008 | [19] | $A_C^{\cos i\phi}$ | $i=0,1$ | $x_{\mathrm{Bj}}$ | 18 / 24 |
| | | | | $A_{UT,\mathrm{DVCS}}^{\sin(\phi-\phi_S)\cos i\phi}$ | $i=0$ | | |
| | | | | $A_{UT,\mathrm{I}}^{\sin(\phi-\phi_S)\cos i\phi}$ | $i=0,1$ | | |
| | | | | $A_{UT,\mathrm{I}}^{\cos(\phi-\phi_S)\sin i\phi}$ | $i=1$ | | |
| 4 | | 2009 | [20] | $A_{LU,\mathrm{I}}^{\sin i\phi}$ | $i=1,2$ | $x_{\mathrm{Bj}}$ | 35 / 42 |
| | | | | $A_{LU,\mathrm{DVCS}}^{\sin i\phi}$ | $i=1$ | | |
| | | | | $A_C^{\cos i\phi}$ | $i=0,1,2,3$ | | |
| 5 | | 2010 | [21] | $A_{UL}^{+,\sin i\phi}$ | $i=1,2,3$ | $x_{\mathrm{Bj}}$ | 18 / 24 |
| | | | | $A_{LL}^{+,\cos i\phi}$ | $i=0,1,2$ | | |
| 6 | | 2011 | [22] | $A_{LT,\mathrm{DVCS}}^{\cos(\phi-\phi_S)\cos i\phi}$ | $i=0,1$ | $x_{\mathrm{Bj}}$ | 24 / 32 |
| | | | | $A_{LT,\mathrm{DVCS}}^{\sin(\phi-\phi_S)\sin i\phi}$ | $i=1$ | | |
| | | | | $A_{LT,\mathrm{I}}^{\cos(\phi-\phi_S)\cos i\phi}$ | $i=0,1,2$ | | |
| | | | | $A_{LT,\mathrm{I}}^{\sin(\phi-\phi_S)\sin i\phi}$ | $i=1,2$ | | |
| 7 | | 2012 | [23] | $A_{LU,\mathrm{I}}^{\sin i\phi}$ | $i=1,2$ | $x_{\mathrm{Bj}}$ | 35 / 42 |
| | | | | $A_{LU,\mathrm{DVCS}}^{\sin i\phi}$ | $i=1$ | | |
| | | | | $A_C^{\cos i\phi}$ | $i=0,1,2,3$ | | |
| 8 | CLAS | 2001 | [24] | $A_{LU}^{-,\sin i\phi}$ | $i=1,2$ | — | 0 / 2 |
| 9 | | 2006 | [25] | $A_{UL}^{-,\sin i\phi}$ | $i=1,2$ | — | 2 / 2 |
| 10 | | 2008 | [26] | $A_{LU}^{-}$ | | $\phi$ | 283 / 737 |
| 11 | | 2009 | [27] | $A_{LU}^{-}$ | | $\phi$ | 22 / 33 |
| 12 | | 2015 | [28] | $A_{LU}^{-}$, $A_{UL}^{-}$, $A_{LL}^{-}$ | | $\phi$ | 311 / 497 |
| 13 | | 2015 | [29] | $d^4\sigma_{UU}^{-}$ | | $\phi$ | 1333 / 1933 |
| 14 | Hall A | 2015 | [15] | $\Delta d^4\sigma_{LU}^{-}$ | | $\phi$ | 228 / 228 |
| 15 | | 2017 | [16] | $\Delta d^4\sigma_{LU}^{-}$ | | $\phi$ | 276 / 358 |
| 16 | COMPASS | 2018 | [30] | $d^3\sigma_{UU}^{\pm}$ | | t | 2 / 4 |
| 17 | ZEUS | 2009 | [31] | $d^3\sigma_{UU}^{+}$ | | t | 4 / 4 |
| 18 | H1 | 2005 | [32] | $d^3\sigma_{UU}^{+}$ | | t | 7 / 8 |
| 19 | | 2009 | [33] | $d^3\sigma_{UU}^{\pm}$ | | t | 12 / 12 |
| | | | | | | SUM: | 2624 / 3996 |

Replication of experimental data to propagate corresponding uncertainties:

$$v_i \pm \Delta_i^{\mathrm{tot}} \xrightarrow{\text{replica } j} \left(\mathrm{rnd}_j(v_i, \Delta_i^{\mathrm{tot}}) \pm \Delta_i^{\mathrm{tot}}\right) \times \mathrm{rnd}_j(1, \Delta_i^{\mathrm{norm}}) \qquad \Delta_i^{\mathrm{tot}} = \sqrt{(\Delta_i^{\mathrm{stat}})^2 + (\Delta_i^{\mathrm{sys}})^2}$$

# Feasibility test

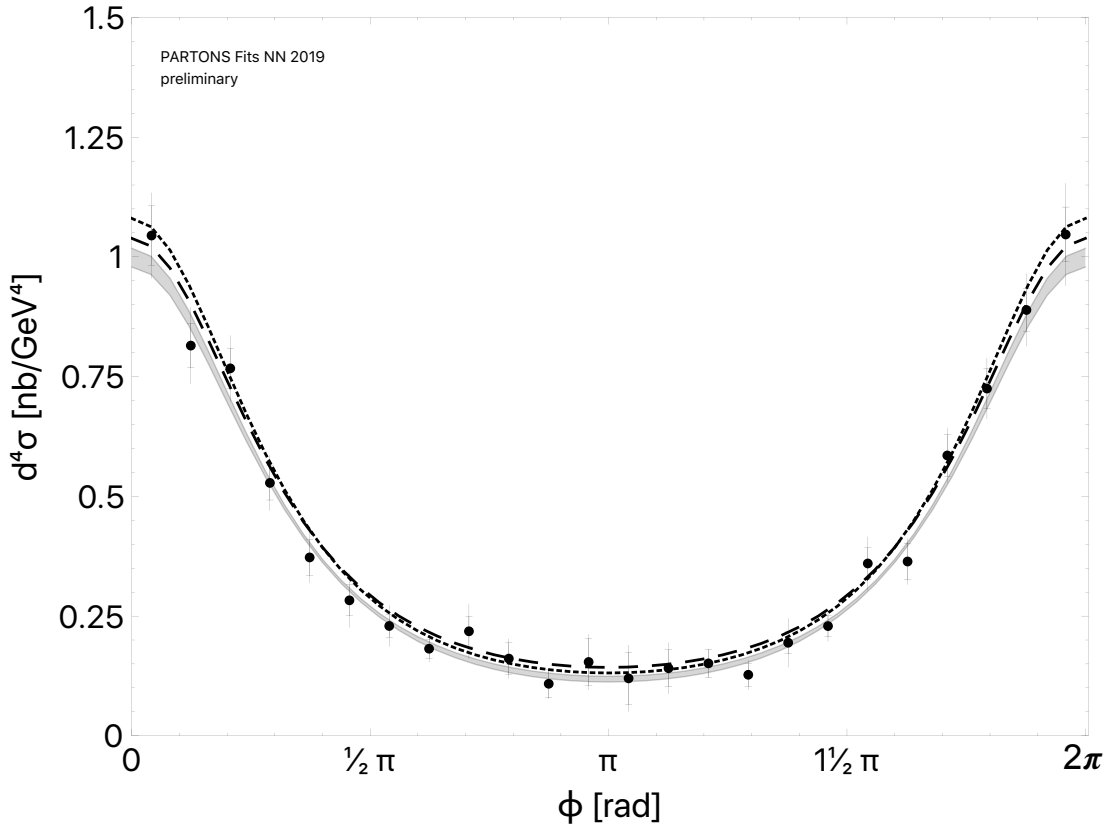
- Pseudo-data generated with GK model
- Kinematics corresponding to real data
- Pseudo-data smeared according to uncertainty of real data
- $\chi^2$ / nPoints:
  all: 2957.2 / 2624 ≈ 1.13
  without low-x data: 2765.1 / 2599 ≈ 1.06

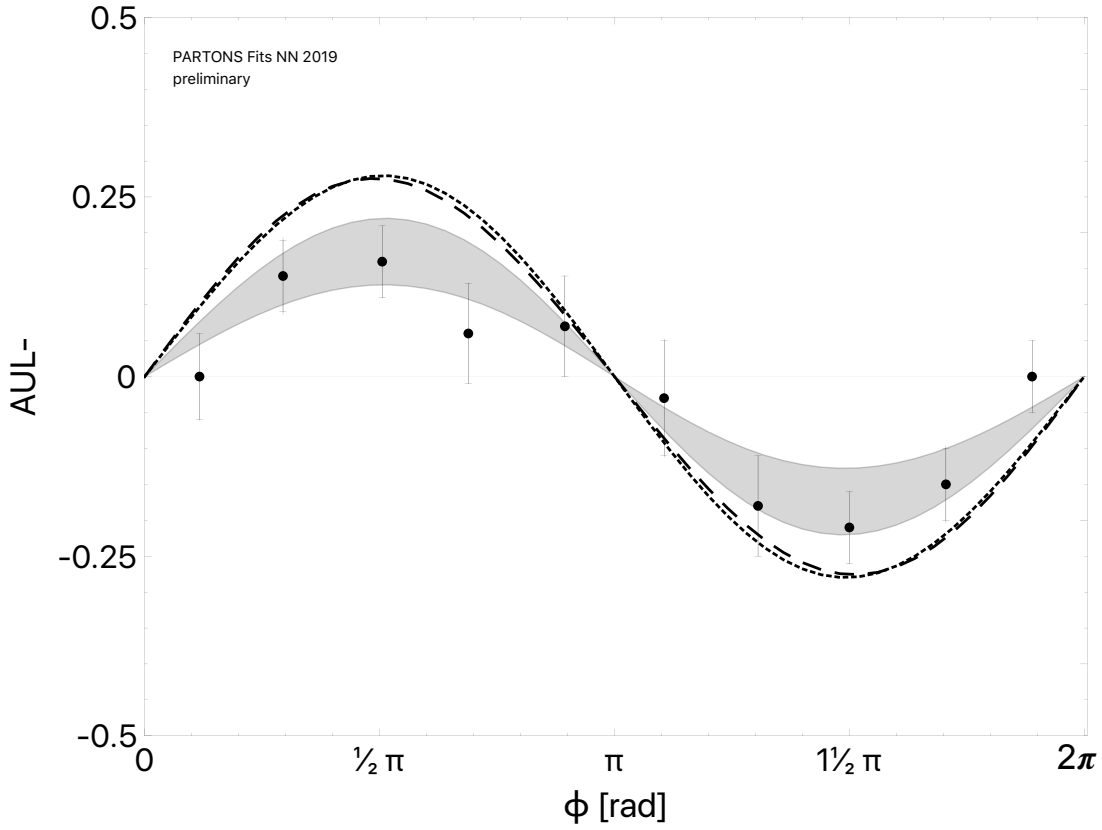



@ t = -0.3 GeV$^2$, Q$^2$ = 2 GeV$^2$

this analysis — — — VGG ····· GK

## CLAS data




Phys. Rev. Lett. 115(21), 212003 (2015)

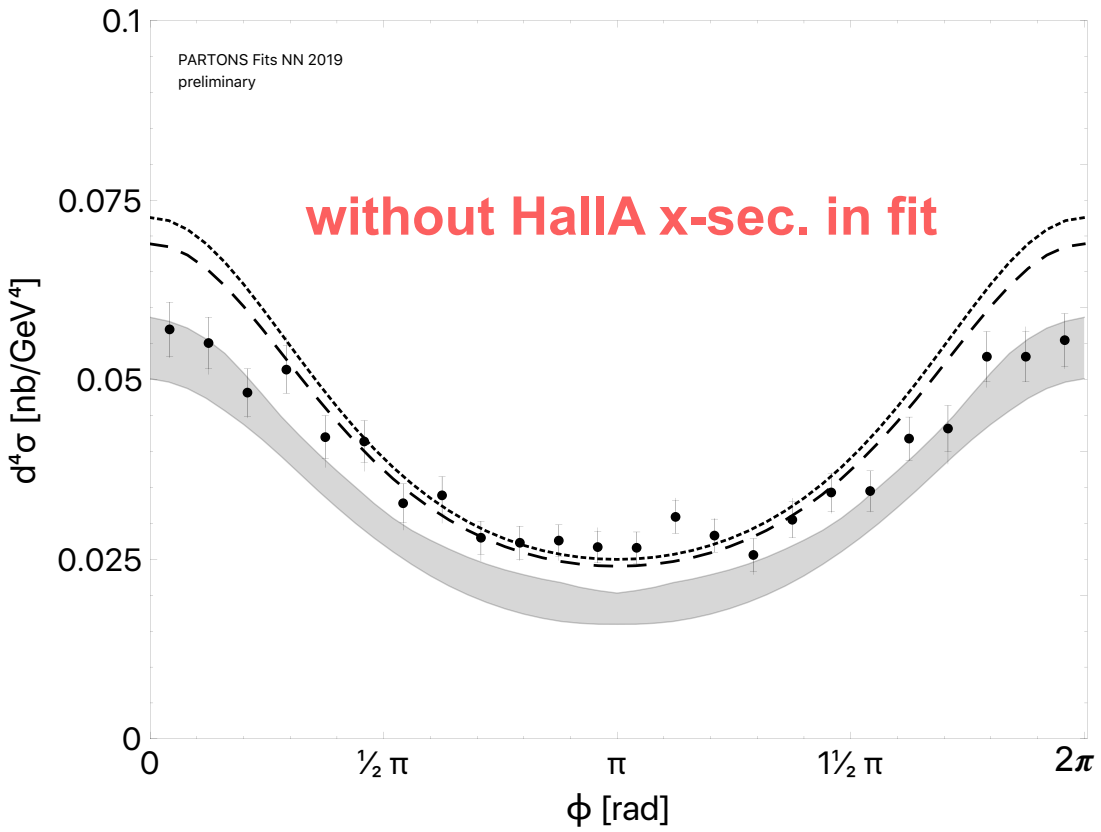
$x_{Bj} = 0.244,\ t = -0.15\ \mathrm{GeV}^2,\ Q^2 = 1.79\ \mathrm{GeV}^2$

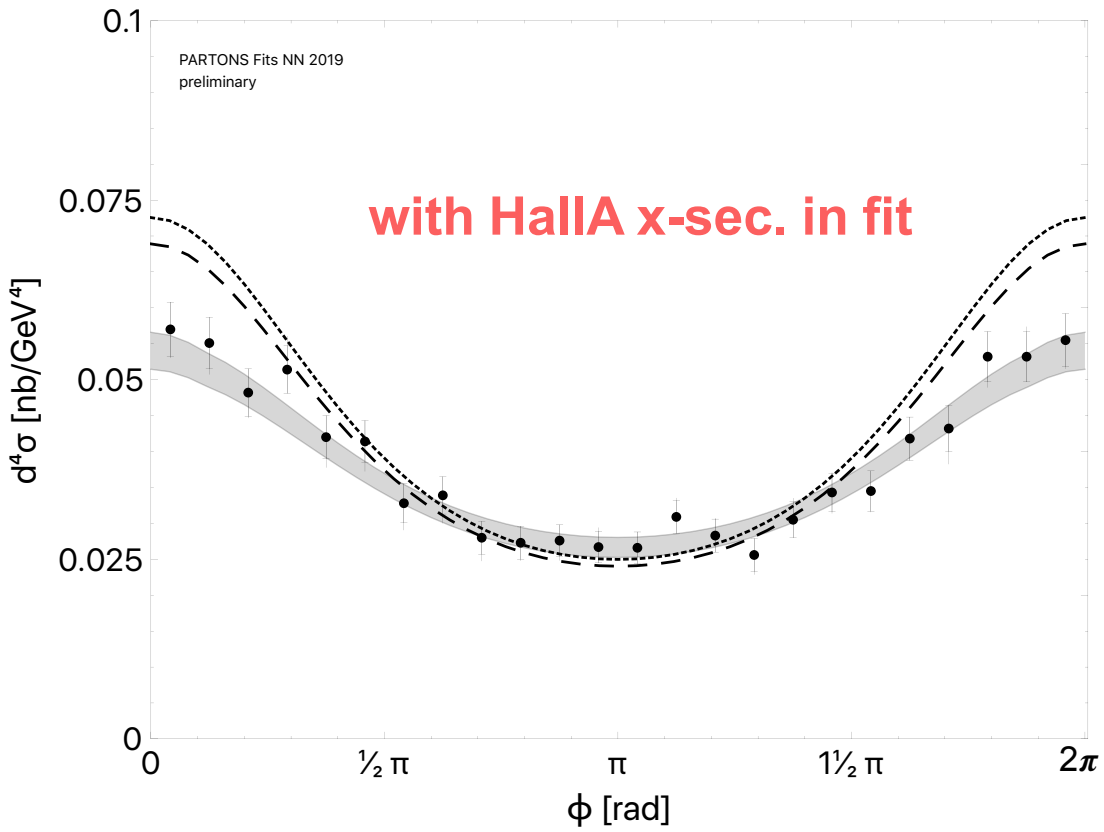
Phys. Rev. D91(5), 052014 (2015)

$x_{Bj} = 0.257,\ t = -0.23\ \mathrm{GeV}^2,\ Q^2 = 2.02\ \mathrm{GeV}^2$

# Hall-A data

this analysis — — — VGG ····· GK



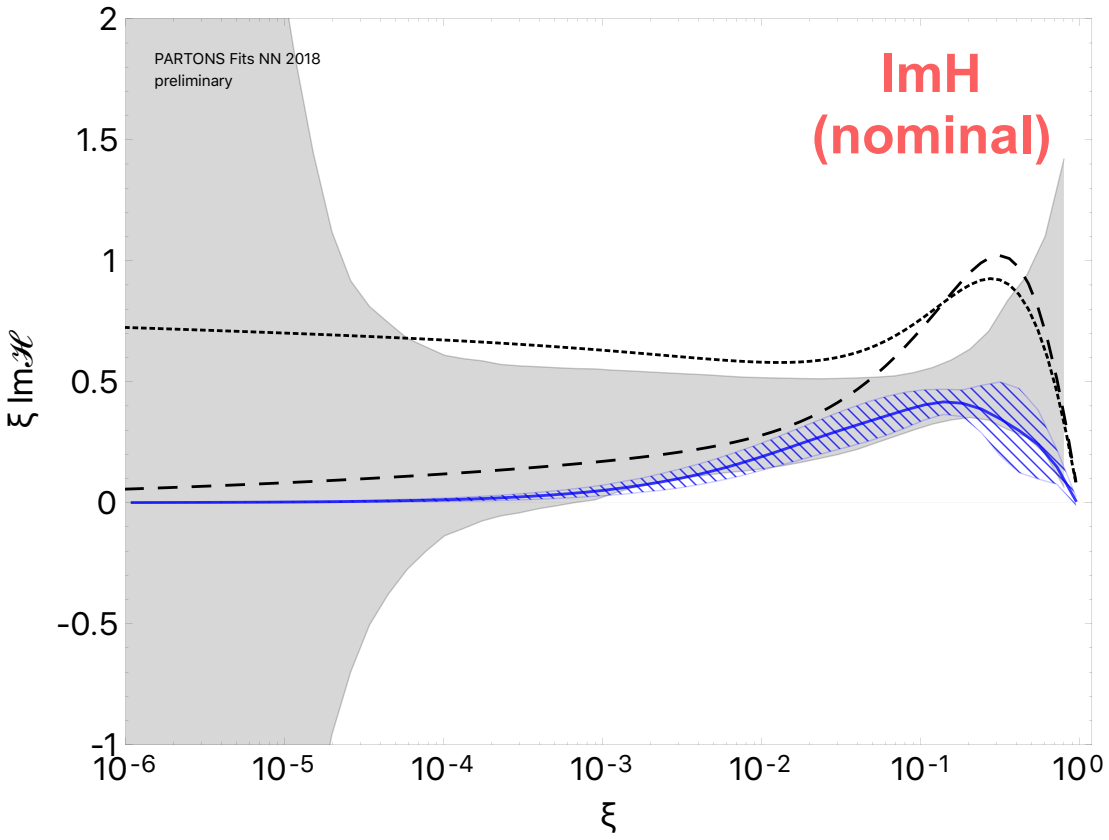

**$\chi^2$/nPoints table →**

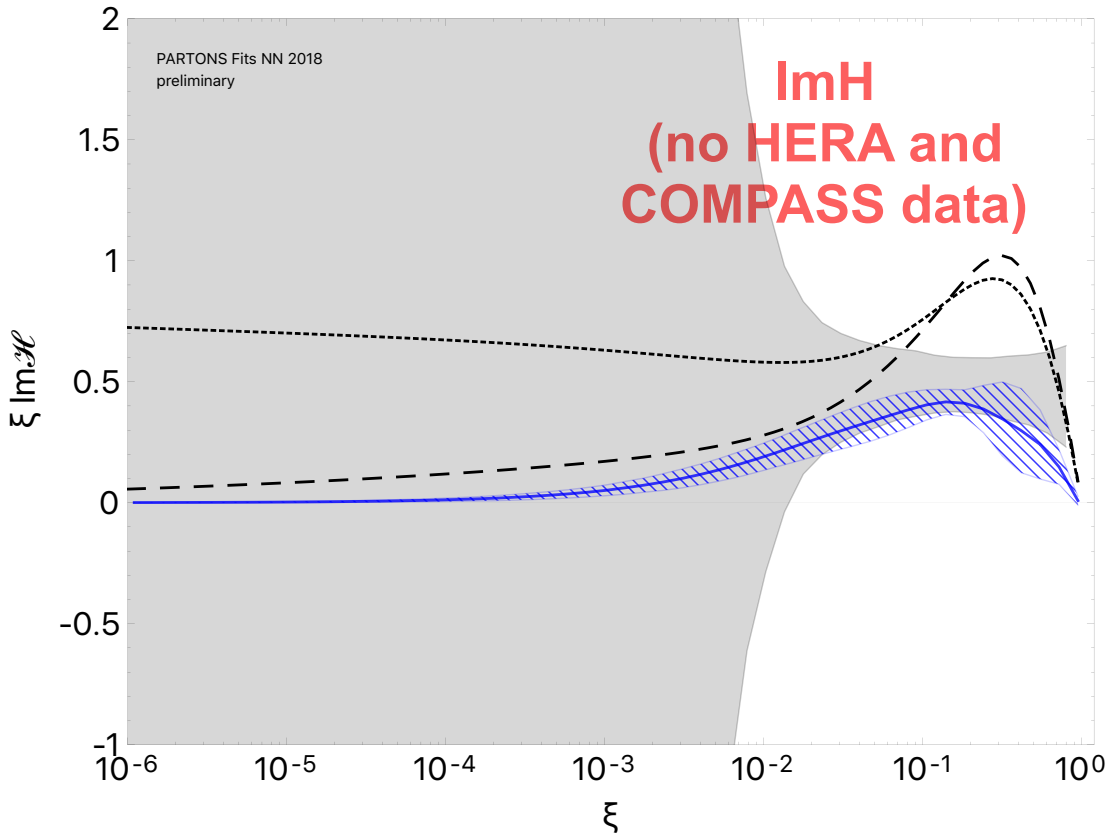
| | witout HallA x-sec. in fit | with HallA x-sec. in fit |
|---|---|---|
| all except low-x data | 2189/2599 = 0.84 | 2481/2599 = 0.95 |
| only CLAS x-sec. | 885/1333 = 0.66 | 1102/1333 = 0.83 |
| only HallA x-sec. | 5917/594 = 9.96 | 664/594 = 1.12 |

Phys. Rev. C92(5), 055202 (2015)

$x_{\mathrm{Bj}} = 0.392,\ t = -0.233\ \mathrm{GeV}^2,\ Q^2 = 2.054\ \mathrm{GeV}^2$

# Inclusion of low-x data



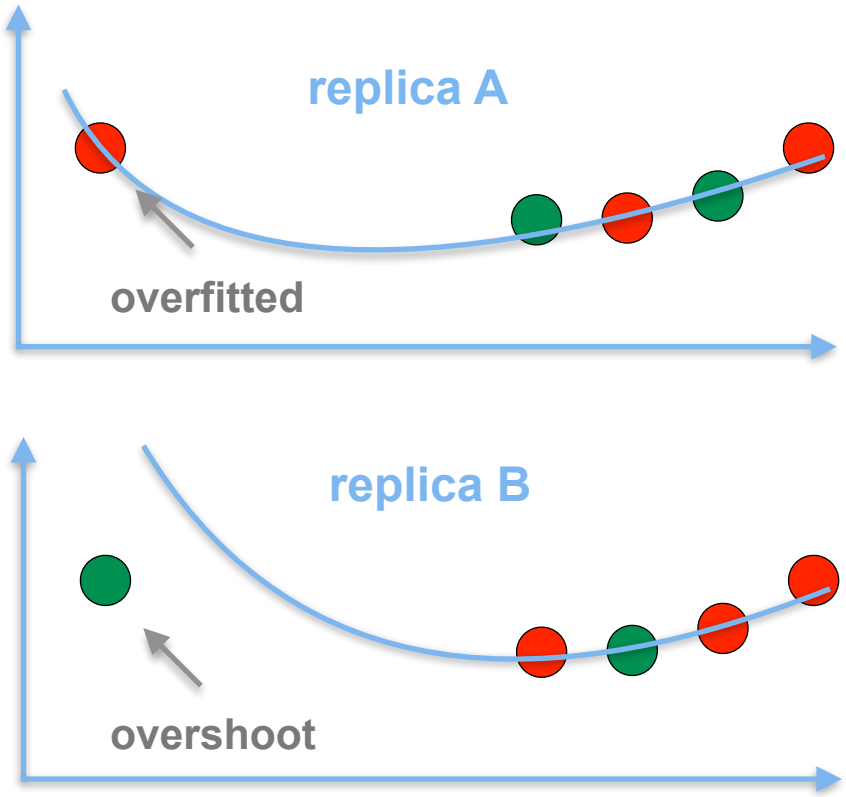




@ t = -0.3 GeV$^2$, Q$^2$ = 2 GeV$^2$

Note: sparsity of low-x data makes supervised training difficult
→ see next slide

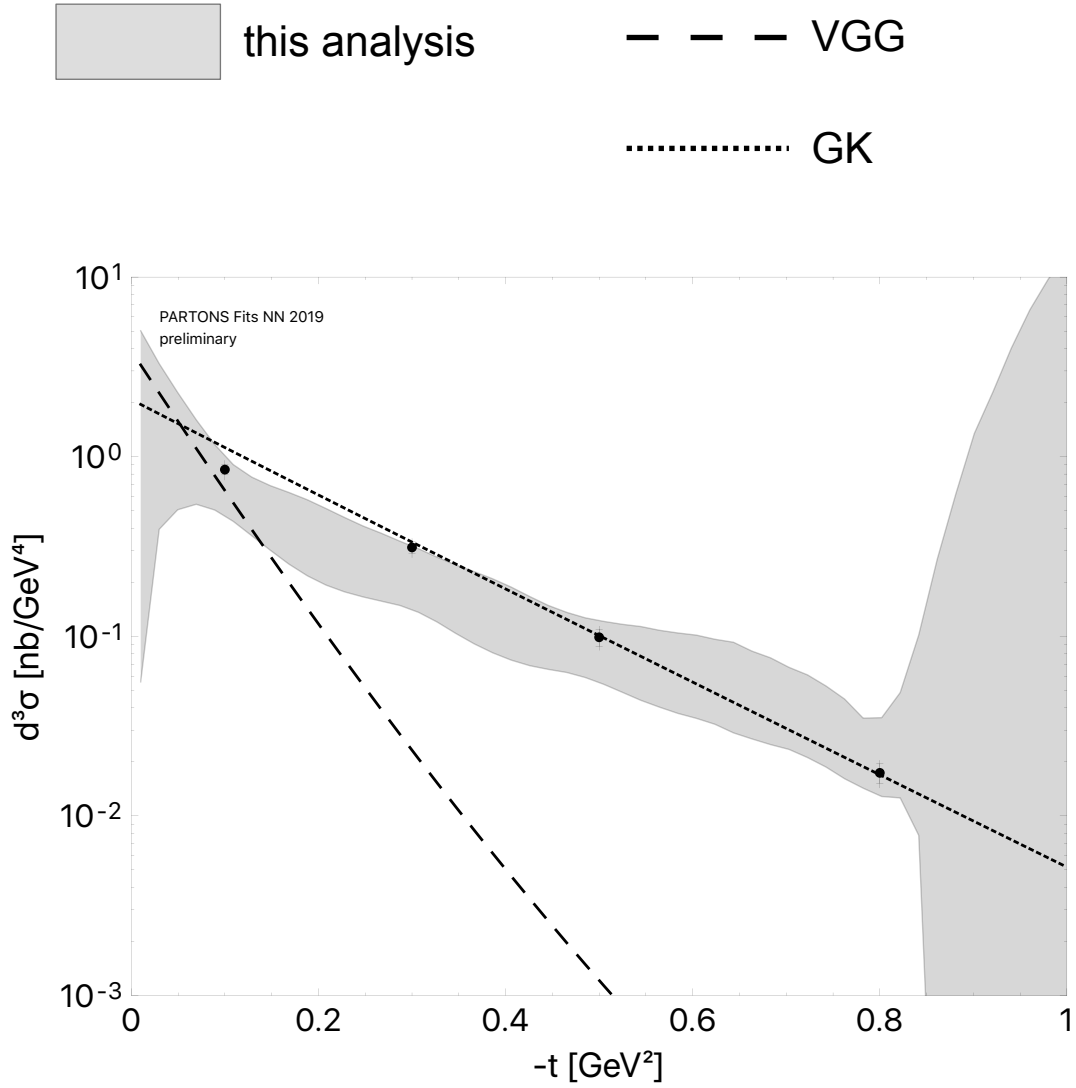
# Inclusion of low-x data

Without decent coverage of phase-space supervised training is challenging




training sample

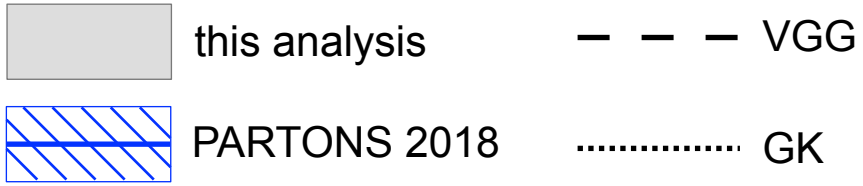
test sample

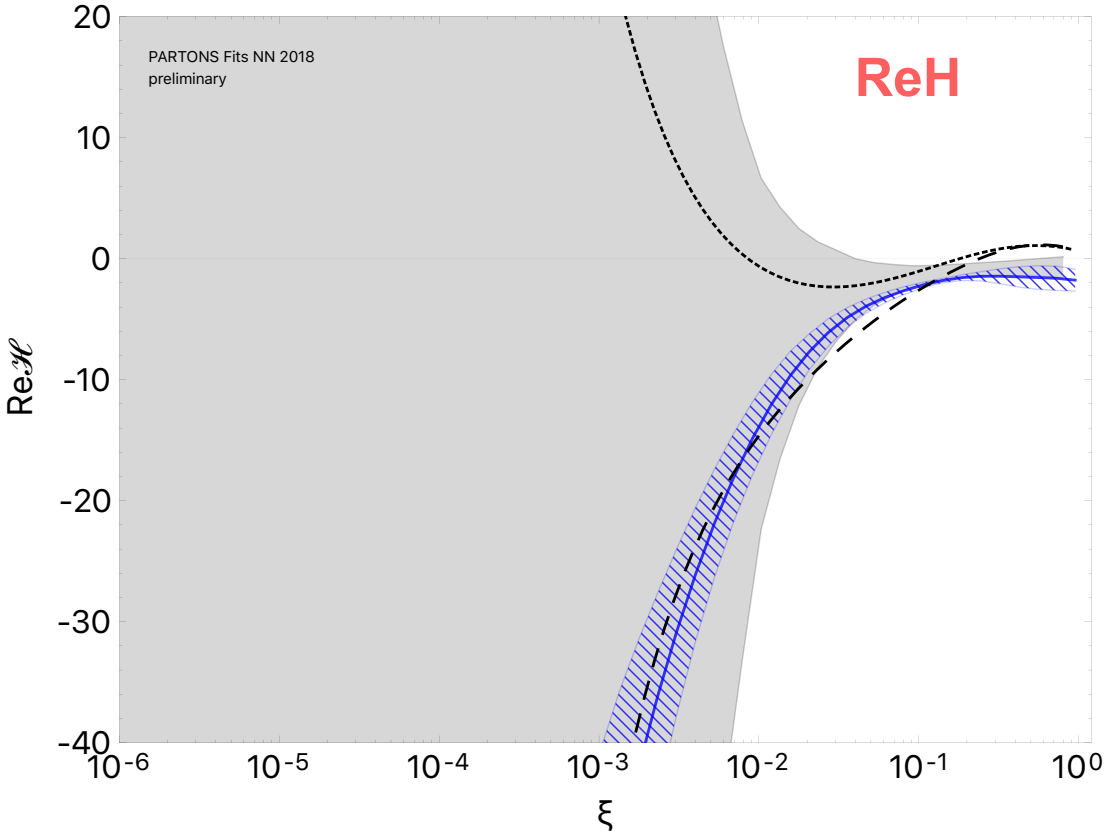


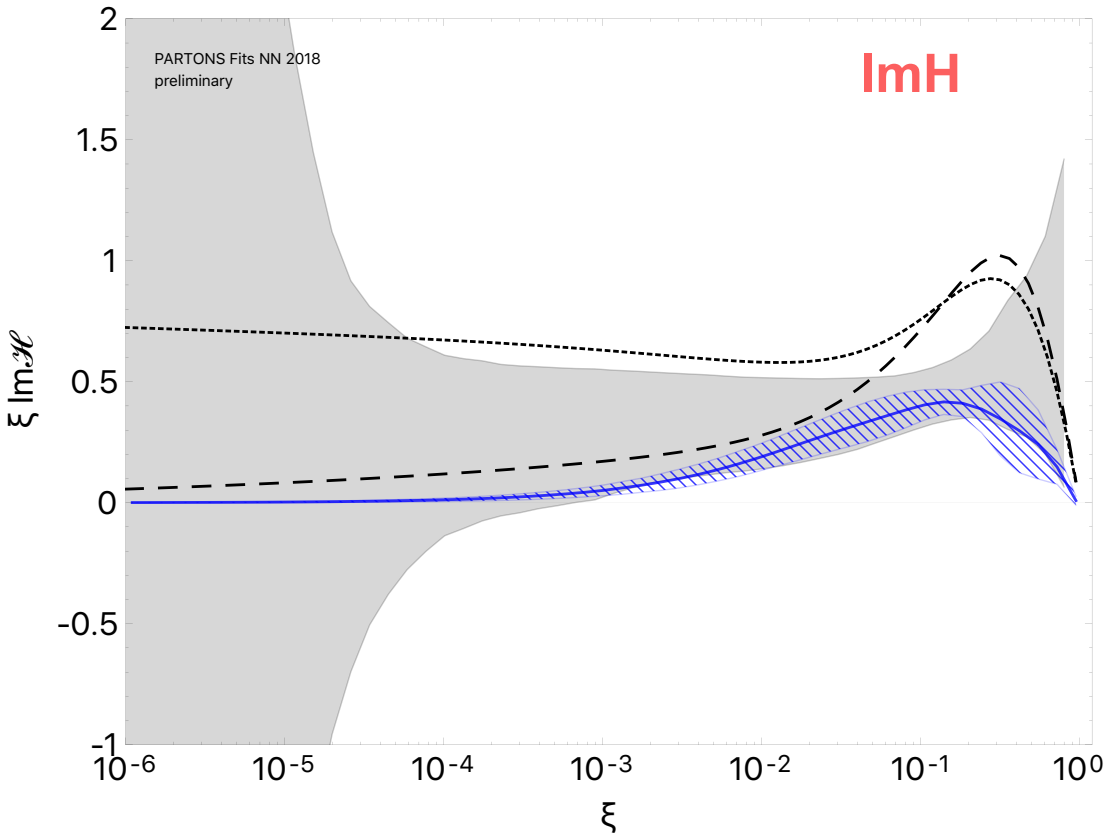

Phys. Lett. B681 391 (2009)

$x_{\mathrm{Bj}} = 0.002,\ Q^2 = 10\ \mathrm{GeV}^2$

# Results

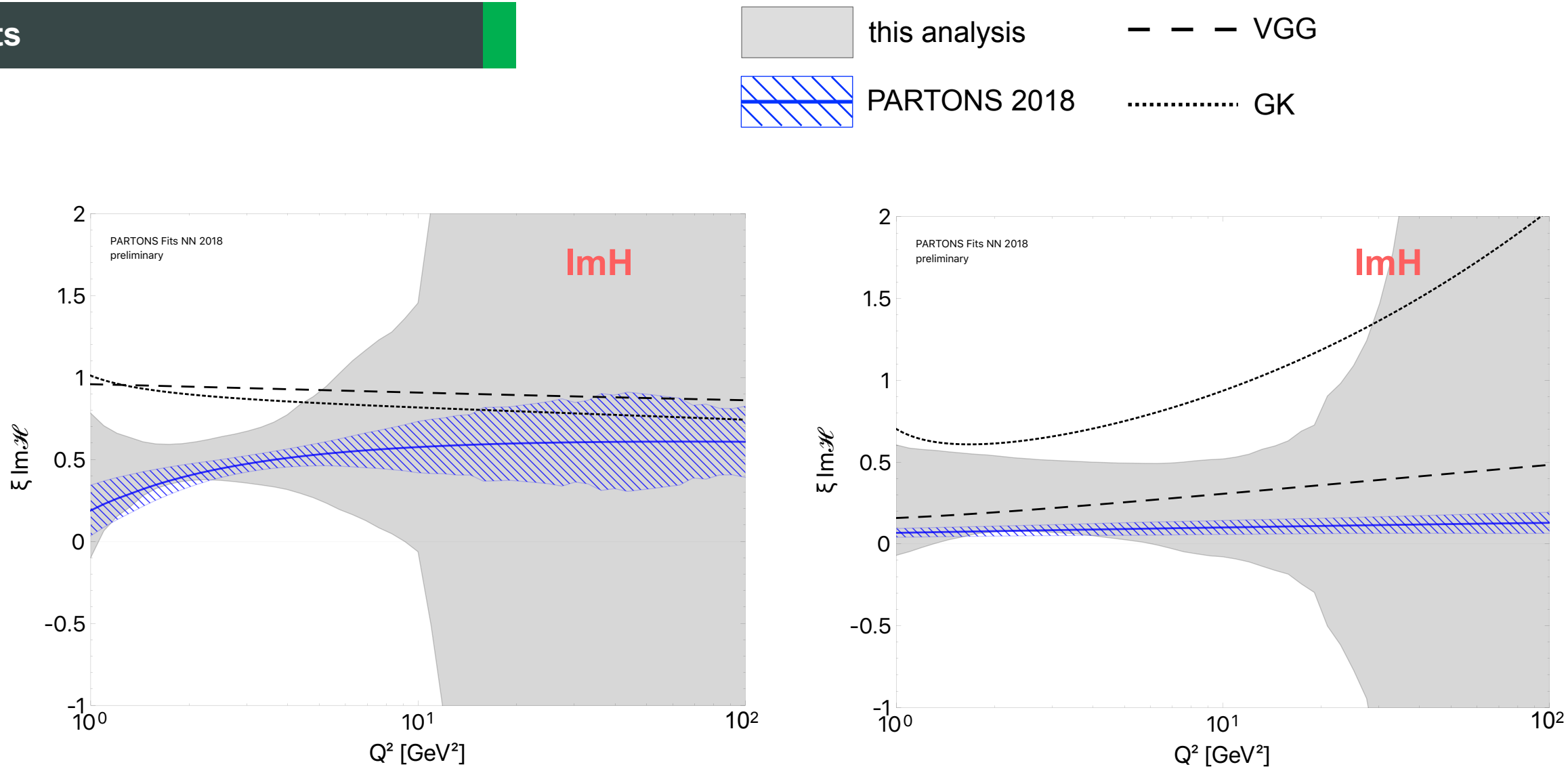



@ t = -0.3 GeV$^2$, Q$^2$ = 2 GeV$^2$

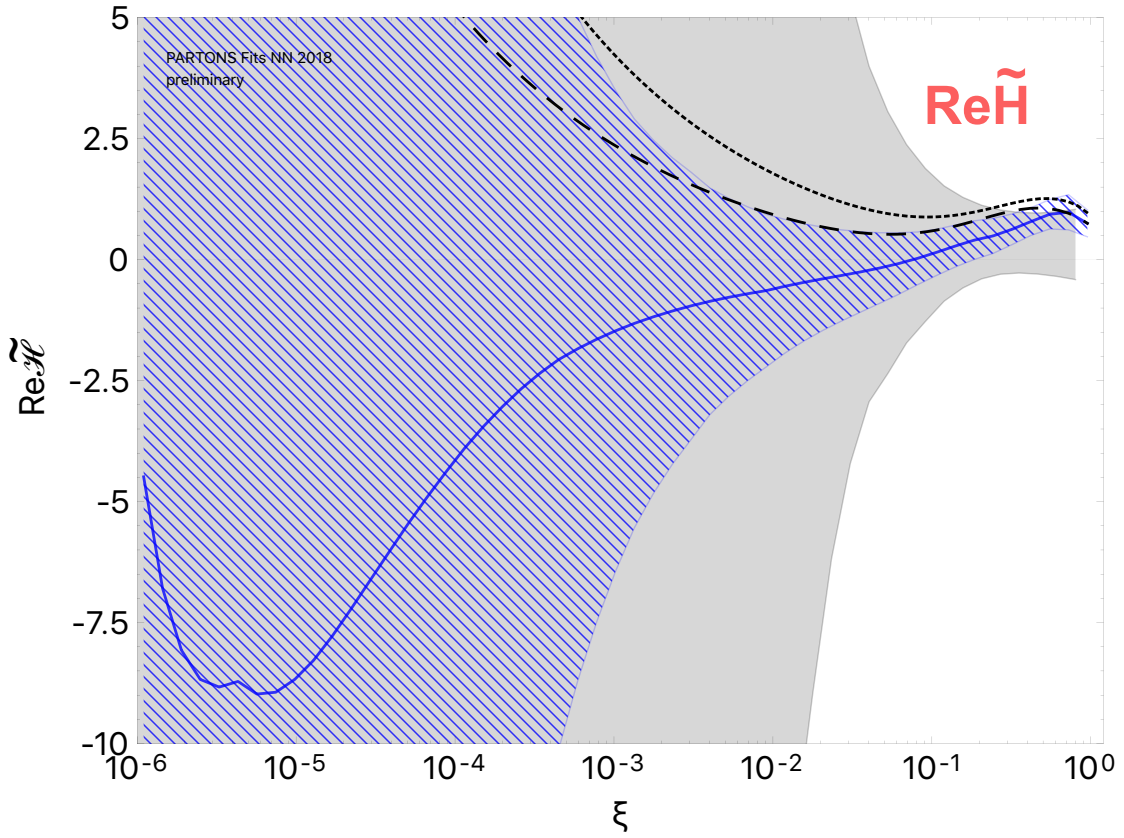
# Results

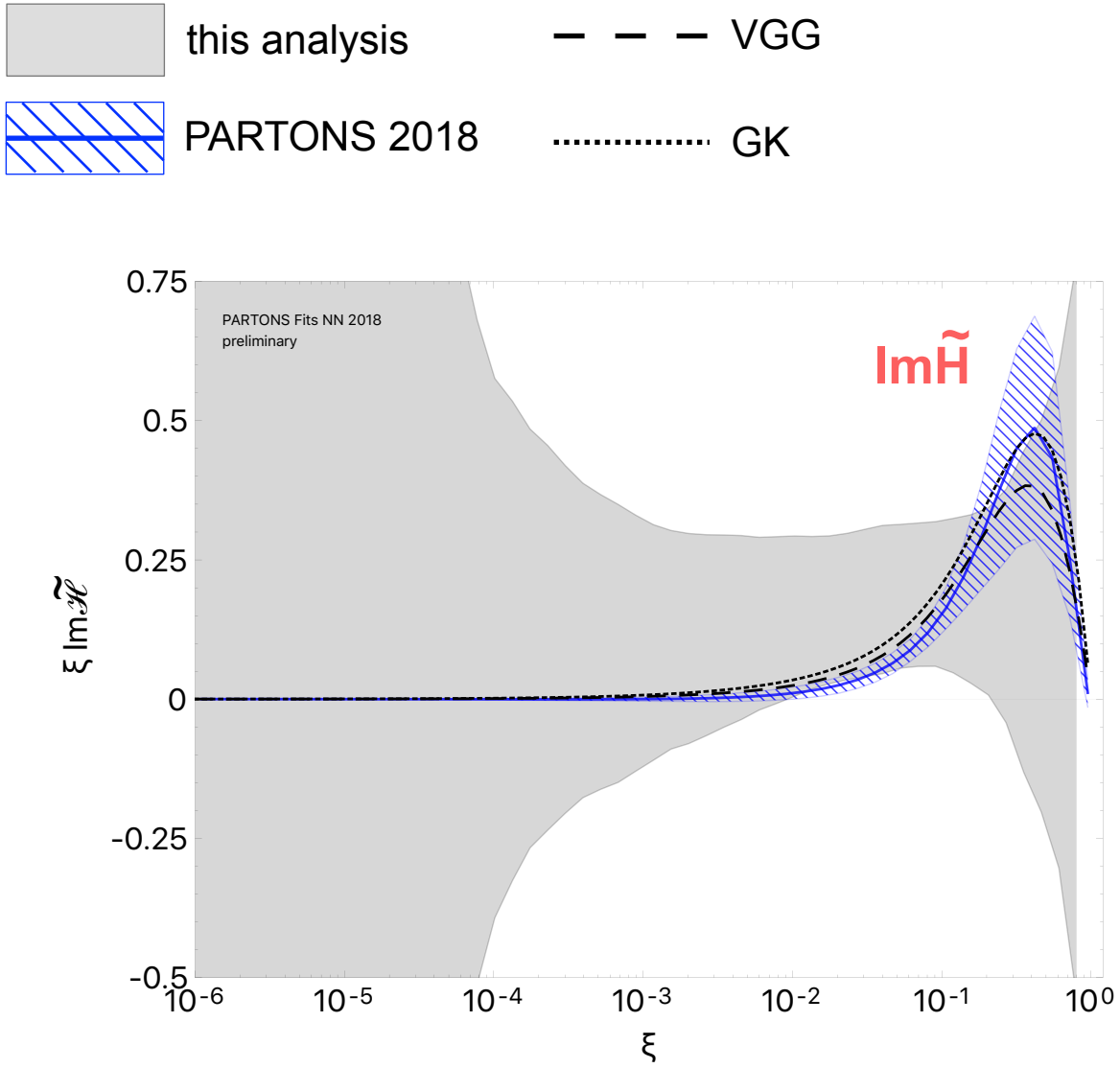



@ ξ = 0.2, t = -0.3 GeV²

@ ξ = 0.002, t = -0.3 GeV²

# Results

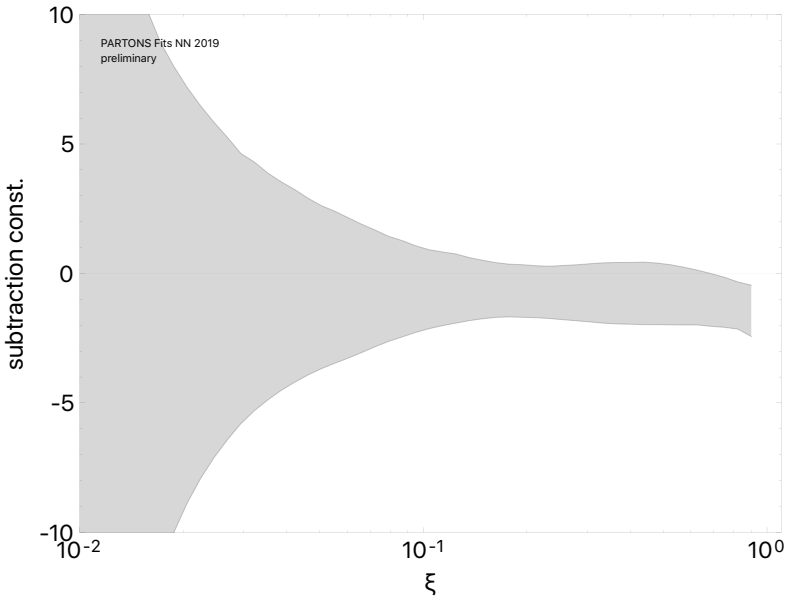


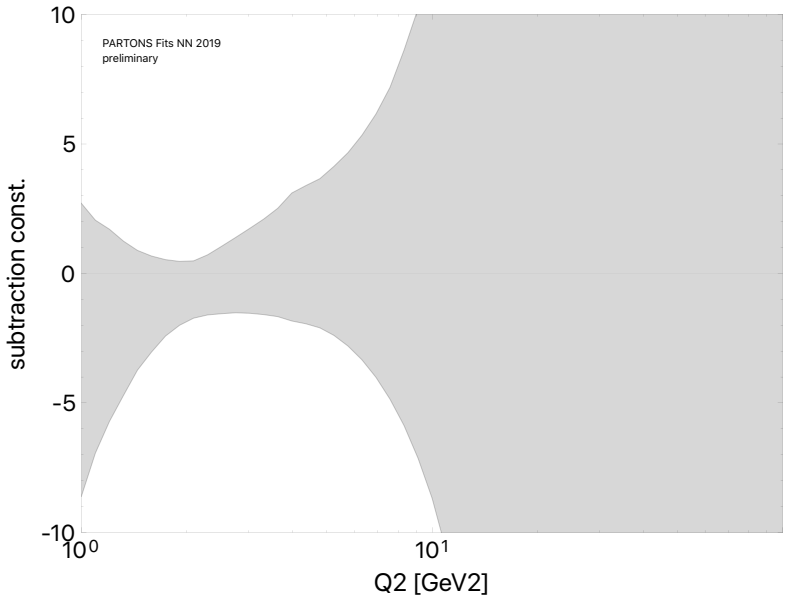

@ t = -0.3 GeV$^2$, Q$^2$ = 2 GeV$^2$

# Subtraction constant

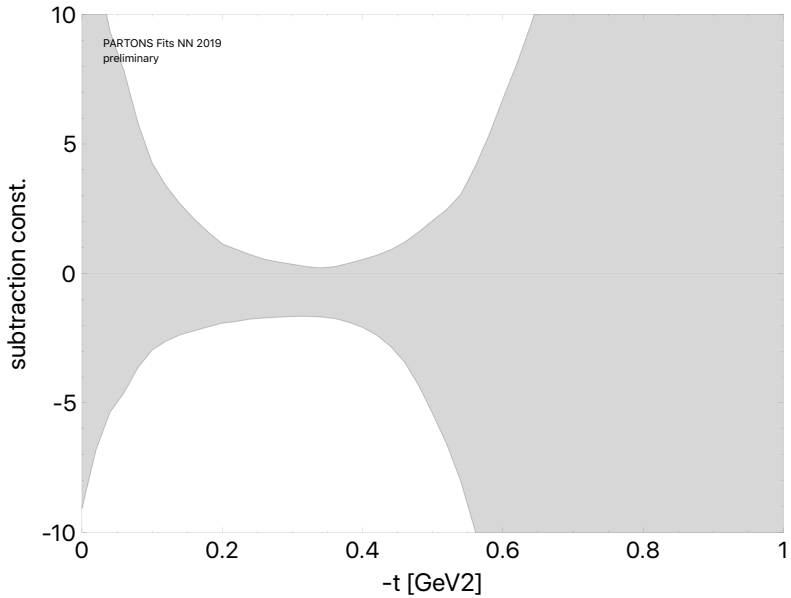
**as function of ξ**
**@ |t| = 0.3 GeV², Q² = 2 GeV²**




**as function of Q²**
**@ ξ = 0.2, |t| = 0.3 GeV²**




**as function of |t|**
**@ ξ = 0.2, Q² = 2 GeV²**




- Direct extraction of subtraction constant → encouraging precision
- As expected, no ξ behaviour observed
- Strong, model independent constraints on modeling of this quantity

# SUMMARY

- Parameterizations of border and skewness function

  → basic properties of GPD as building blocks

  → small number of parameters

  → encoded access to nucleon tomography and subtraction constant

- Neural network parameterization of CFFs

  → powerful tool to study GPDs

  → model independent

  → many applications:

    → extraction of subtraction constant

    → analysis of TCS

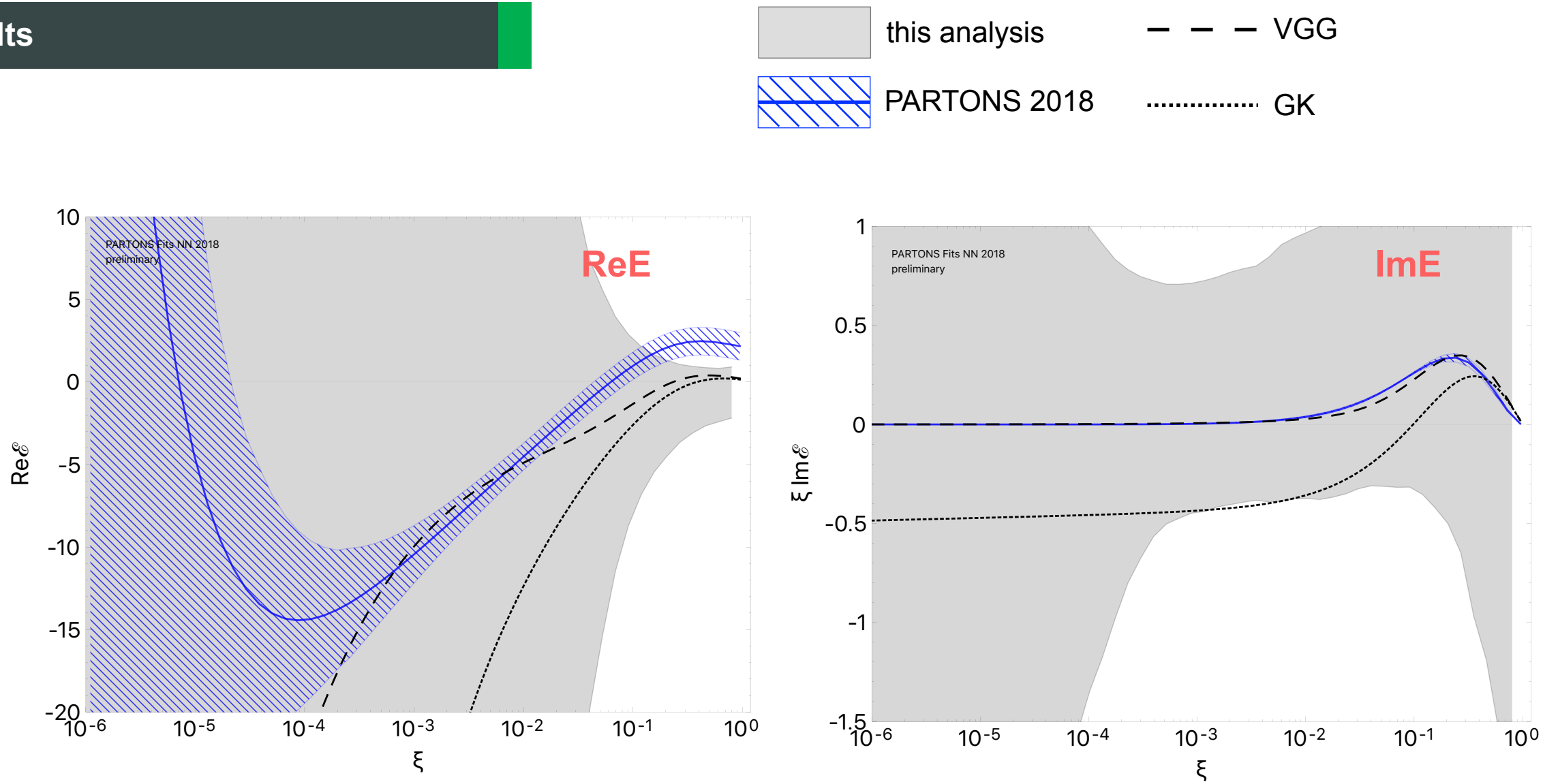
    → impact of future experiments

    → fast generation of DVCS cross sections

    → …

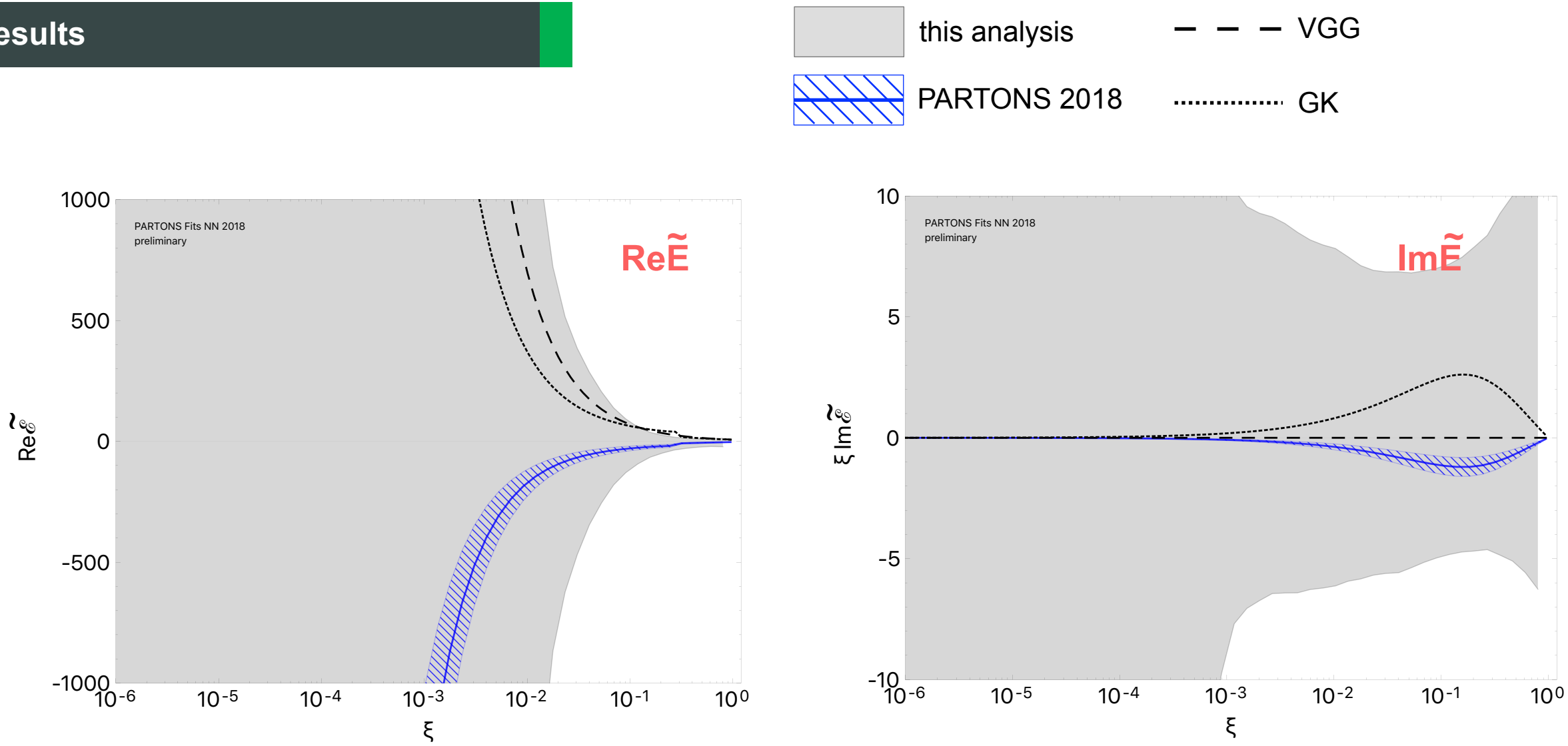
- Need more tools to visualise and compare multi-dim. distributions → Center for Nucl. Femtography

# Results




@ t = -0.3 GeV$^2$, Q$^2$ = 2 GeV$^2$

# Results




@ t = -0.3 GeV$^2$, Q$^2$ = 2 GeV$^2$

# Slope exp(bt)

this analysis

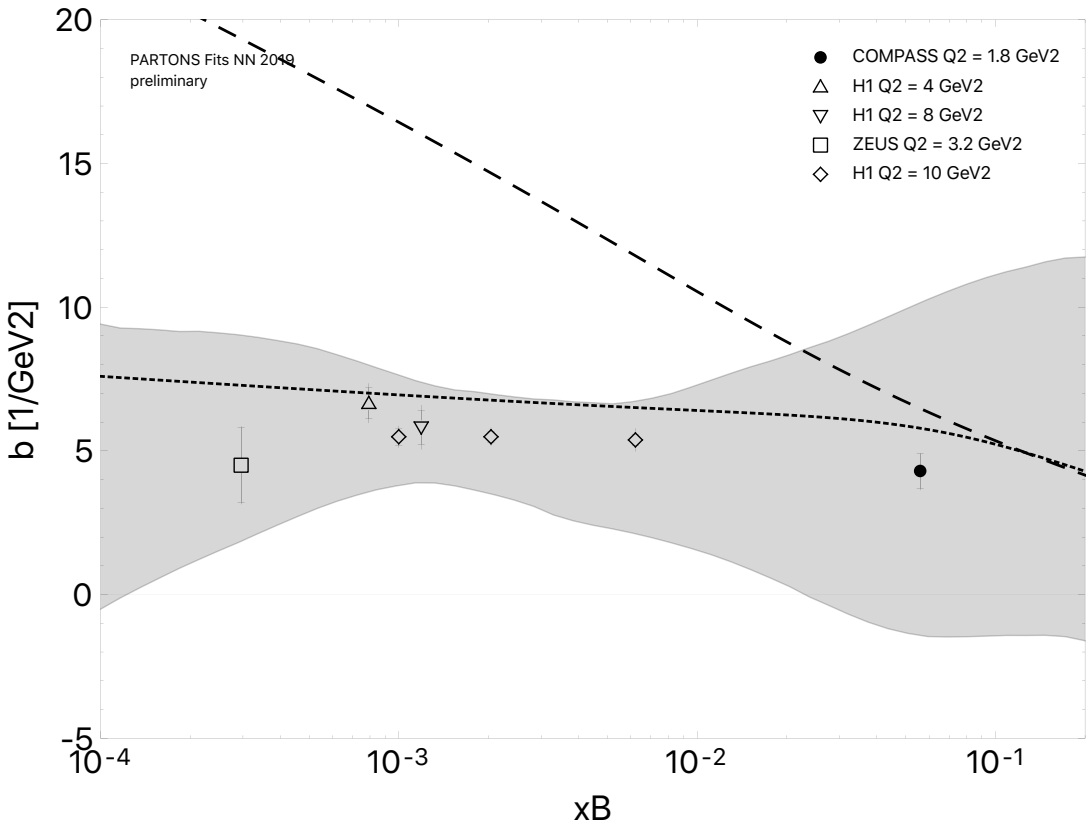
VGG

GK

$$\frac{d\sigma^{\gamma^* p}}{dt} \propto \exp(b \times t)$$

Under assumption of Im$\mathcal{H}$ dominance
observable directly related to nucleon tomography

Estimation of slope from ANN replicas difficult



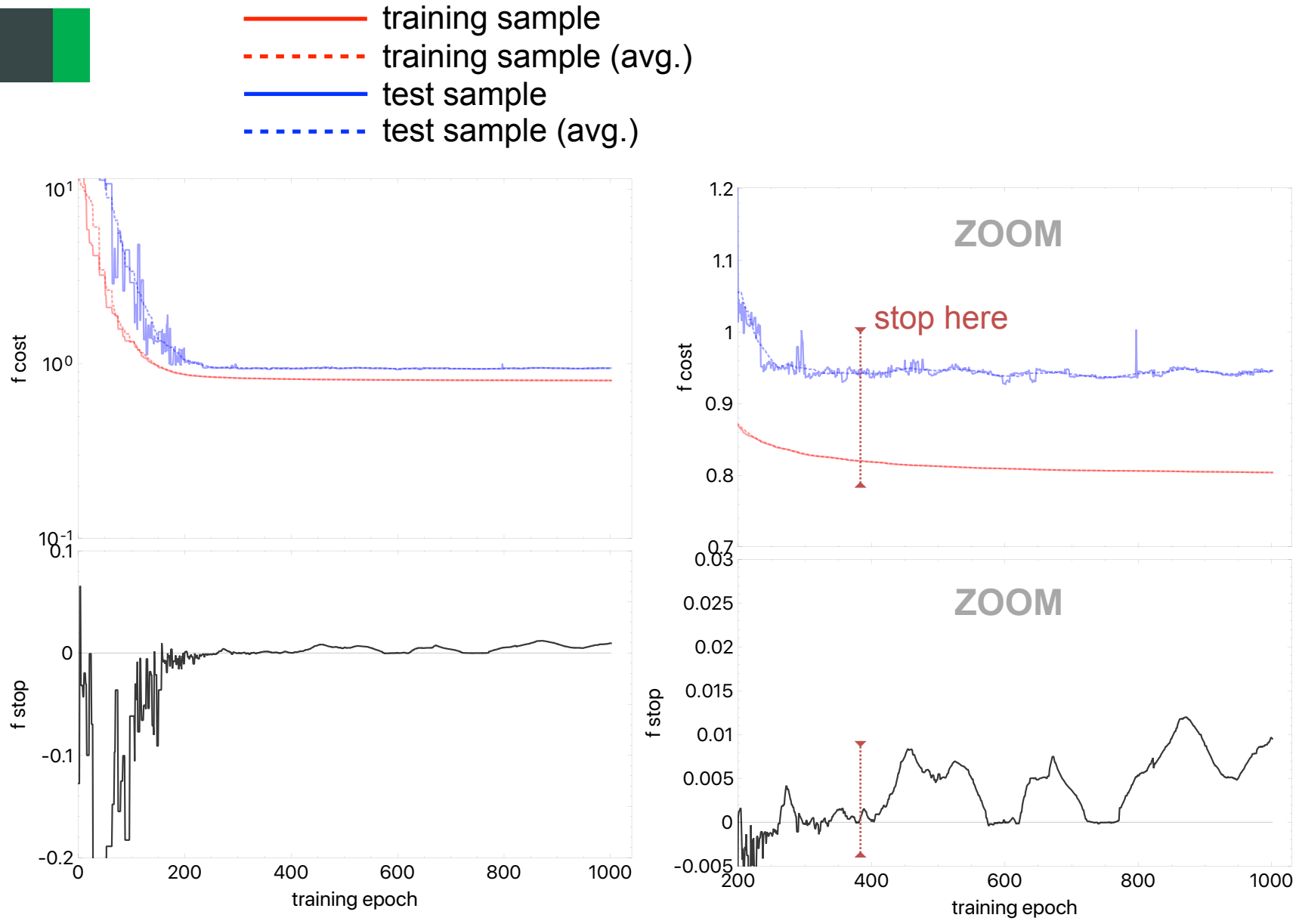

@ $Q^2$ = 10 $GeV^2$

# Early stopping

- Stop training if

$$f_{\text{stop}}(t_i) = \frac{f^{\text{av}}_{\text{cost}}(t_i)}{\min\limits_{j<i} f^{\text{av}}_{\text{cost}}(t_j)} - 1$$

> 0 for more than 100 iterations

- Running average to suppress fluctuations:

$$f^{\text{av}}_{\text{cost}}(t_i) = \frac{f_{\text{cost}}(t_i) + \sum_{j=1}^{n} \left( f_{\text{cost}}(t_{i-j}) + f_{\text{cost}}(t_{i+j}) \right)}{2n+1}$$