Techniques of GPD extraction from data

Krešimir Kumerički

University of Zagreb, Croatia

INT-17-3 Workshop: Tomography of Hadrons and Nuclei at Jefferson Lab
August 28 – September 29, 2017
Seattle, WA

Outline

Introduction — DVCS

ϕ vs. harmonics

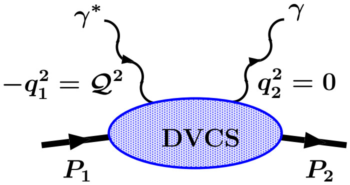
Global fits

Neural nets

Conclusion

Probing the proton with two photons

- Deeply virtual Compton scattering (DVCS) [Müller '92, et al. '94]



$$P = P_1 + P_2 \,, \qquad t = (P_2 - P_1)^2$$

$$q = (q_1 + q_2)/2$$

Generalized Bjorken limit:

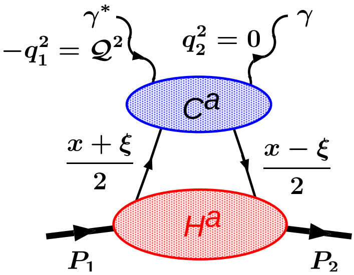
$$-q^2 \simeq \mathcal{Q}^2/2 \to \infty$$

$$\xi = \frac{-q^2}{2P \cdot q} \to \text{const}$$

- To leading twist-two accuracy cross-section can be expressed in terms of Compton form factors (CFFs)

$$\mathcal{H}(\xi, t, \mathcal{Q}^2), \mathcal{E}(\xi, t, \mathcal{Q}^2), \tilde{\mathcal{H}}(\xi, t, \mathcal{Q}^2), \tilde{\mathcal{E}}(\xi, t, \mathcal{Q}^2), \ldots$$

Factorization of DVCS $\longrightarrow$ GPDs



$$P = P_1 + P_2 \,, \qquad t = (P_2 - P_1)^2$$
$$q = (q_1 + q_2)/2$$

$$-q^2 \simeq \mathcal{Q}^2/2 \to \infty$$
$$\xi = \frac{-q^2}{2P \cdot q} \to \text{const}$$

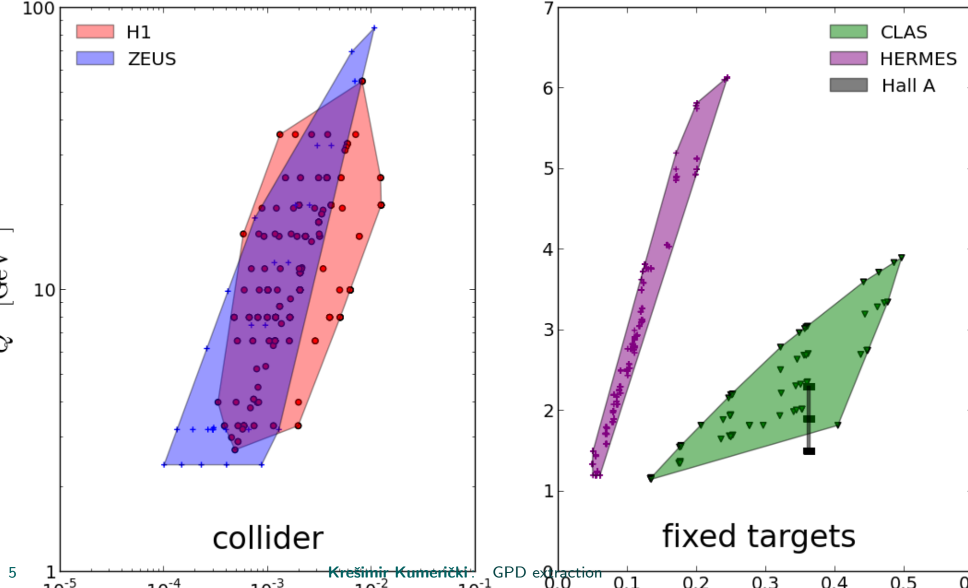
- Compton form factor is a convolution:

$$^a\mathcal{H}(\xi, t, \mathcal{Q}^2) = \int \mathrm{d}x \; C^a(x, \xi, \mathcal{Q}^2/\mathcal{Q}_0^2) \; H^a(x, \eta = \xi, t, \mathcal{Q}_0^2)$$

a=NS,S(Σ,G)

- $H^a(x, \eta, t, \mathcal{Q}_0^2)$ — Generalized parton distribution (GPD)

DVCS experimental coverage



Experimental coverage (fixed target)

Collab.	Year	Observables	Kinematics			No. of points	
			x_B	Q^2 [GeV^2]	$\|t\|$ [GeV^2]	total	indep.
HERMES	2001	$A_{LU}^{\sin\phi}$	0.11	2.6	0.27	1	1
CLAS	2001	$A_{LU}^{\sin\phi}$	0.19	1.25	0.19	1	1
CLAS	2006	$A_{UL}^{\sin\phi}$	0.2–0.4	1.82	0.15–0.44	6	3
HERMES	2006	$A_{C}^{\cos\phi}$	0.08–0.12	2.0–3.7	0.03–0.42	4	4
Hall A	2006	$\sigma(\phi)$, $\Delta\sigma(\phi)$	0.36	1.5–2.3	0.17–0.33	4×24+12×24	4×24+12×24
CLAS	2007	$A_{LU}(\phi)$	0.11–0.58	1.0–4.8	0.09–1.8	62×12	62×12
HERMES	2008	$A_{C}^{\cos(0,1)\phi}$, $A_{UT,DVCS}^{\sin(\phi-\phi_S)}$, $A_{UT,I}^{\sin(\phi-\phi_S)\cos(0,1)\phi}$, $A_{UT,I}^{\cos(\phi-\phi_S)\sin\phi}$	0.03–0.35	1–10	<0.7	*12*+12+12 12+12 *12*	*4*+4+4 4+4 *4*
CLAS	2008	$A_{LU}(\phi)$	0.12–0.48	1.0–2.8	0.1–0.8	66	33
HERMES	2009	$A_{LU,I}^{\sin(1,2)\phi}$, $A_{LU,DVCS}^{\sin\phi}$, $A_{C}^{\cos(0,1,2,3)\phi}$	0.05–0.24	1.2–5.75	<0.7	18+18+18 18+18+*18*+*18*	6+6+6 6+6+*6*+*6*
HERMES	2010	$A_{UL}^{\sin(1,2,3)\phi}$, $A_{LL}^{\cos(0,1,2)\phi}$	0.03–0.35	1–10	<0.7	12+12+*12* 12+*12*+12	4+4+*4* 4+*4*+4
HERMES	2011	$A_{LT,I}^{\cos(\phi-\phi_S)\cos(0,1,2)\phi}$, $A_{LT,I}^{\sin(\phi-\phi_S)\sin(1,2)\phi}$, $A_{LT,BH+DVCS}^{\cos(\phi-\phi_S)\cos(0,1)\phi}$, $A_{LT,BH+DVCS}^{\sin(\phi-\phi_S)\sin\phi}$	0.03–0.35	1–10	<0.7	*12+12+12* *12+12* *12*+12 12	*4+4+4* *4+4* *4*+4 4
HERMES	2012	$A_{LU,I}^{\sin(1,2)\phi}$, $A_{LU,DVCS}^{\sin\phi}$, $A_{C}^{\cos(0,1,2,3)\phi}$	0.03–0.35	1–10	<0.7	18+*18*+*18* 18+18+*18*+*18*	6+*6*+*6* 6+6+*6*+*6*
CLAS	2015	$A_{LU}(\phi)$, $A_{UL}(\phi)$, $A_{LL}(\phi)$	0.17–0.47	1.3–3.5	0.1–1.4	166+166+166	166+166+166
CLAS	2015	$\sigma(\phi)$, $\Delta\sigma(\phi)$	0.1–0.58	1–4.6	0.09–0.52	2640+2640	2640+2640
Hall A	2015	$\sigma(\phi)$, $\Delta\sigma(\phi)$	0.33–0.40	1.5–2.6	0.17–0.37	480+600	240+360
Hall A	2017	$\sigma(\phi)$, $\Delta\sigma(\phi)$	[0.36]	[1.5–2.0]	0.17–0.37	450+360	450+360

Electroproduction of photon

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

$$|\mathcal{T}_{\rm BH}|^2 = K^{\rm BH} \frac{1}{\mathcal{P}_1(\phi)\mathcal{P}_2(\phi)} \left\{ c_0^{\rm BH} + \sum_{n=1}^{2} c_n^{\rm BH} \cos(n\phi) + s_1^{\rm BH} \sin\phi \right\},$$

$$\mathcal{I} = K^{\mathcal{I}} \frac{-e_\ell}{\mathcal{P}_1(\phi)\mathcal{P}_2(\phi)} \left\{ c_0^{\mathcal{I}} + \sum_{n=1}^{3} \left[c_n^{\mathcal{I}} \cos(n\phi) + s_n^{\mathcal{I}} \sin(n\phi) \right] \right\},$$

$$|\mathcal{T}_{\rm DVCS}|^2 = K^{\rm DVCS} \left\{ c_0^{\rm DVCS} + \sum_{n=1}^{2} \left[c_n^{\rm DVCS} \cos(n\phi) + s_n^{\rm DVCS} \sin(n\phi) \right] \right\},$$

$$c^{\mathcal{I}}_{1,\rm unpol.} \propto \left[F_1 \,\mathfrak{Re}\,\mathcal{H} - \frac{t}{4M_p^2} F_2 \,\mathfrak{Re}\,\mathcal{E} + \frac{x_{\rm B}}{2 - x_{\rm B}} (F_1 + F_2) \,\mathfrak{Re}\,\widetilde{\mathcal{H}} \right]$$

$$c_n^{\rm DVCS} \propto \quad \text{quadratic in} \;\; \mathcal{H}, \mathcal{E}, \tilde{\mathcal{H}}, \cdots$$

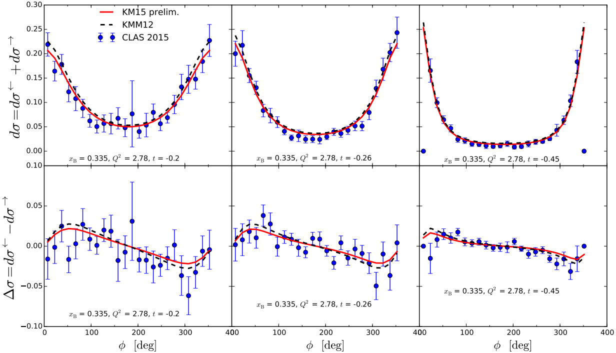
[Belitsky, Müller et. al '01–'14]

Separation of form factors

- One would like to obtain separation between:
  1. different terms in cross section: $|\mathcal{T}_{\rm BH}|^2$, $|\mathcal{T}_{\rm DVCS}|^2$, $\mathcal{I}$
  2. then different harmonic coefficients: $c_0^{\mathcal{I}}$, $s_1^{\mathcal{I}}$, $c_0^{\rm DVCS}$, ...
  3. then different CFFs: $\mathcal{H}$, $\mathcal{E}$, $\tilde{\mathcal{H}}$, ...
- ...and then one could try to extract GPDs
- To do that in experiments one chooses
  1. different beam charges
  2. different beam and target polarizations
  3. different beam energies, utilizing dependence of harmonics c_n and kinematical factors K on y ("Rosenbluth separation"):

$$K = K(x_{\rm B}, Q^2, t|y), \qquad y = \frac{Q^2}{x_{\rm B}(s - M_p^2)}$$

Case #1: CLAS cross-section (ϕ-space)



- KM15 model (global refit including this data):
  $\chi^2/\mathrm{npts} = 1032.0/1014$ for $d\sigma$ and $936.1/1012$ for $\Delta\sigma$
- looks excellent

Going from ϕ- to n-space

- Because of simple dependence of cross-section σ on ϕ, harmonics (Fourier transforms) of σ provide very direct access to coefficients c_0, c_1, ...
- It is often convenient to work with weighted harmonics

$$\sigma^{\sin n\phi, w} \equiv \frac{1}{\pi} \int_{-\pi}^{\pi} dw \, \sin(n\phi) \, \sigma(\phi) \, ,$$

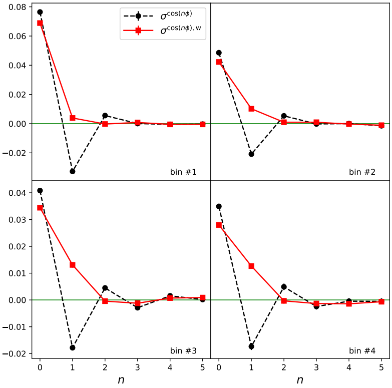
with specialy weighted Fourier integral measure

$$dw \equiv \frac{2\pi \mathcal{P}_1(\phi)\mathcal{P}_2(\phi)}{\int_{-\pi}^{\pi} d\phi \, \mathcal{P}_1(\phi)\mathcal{P}_2(\phi)} d\phi \, ,$$

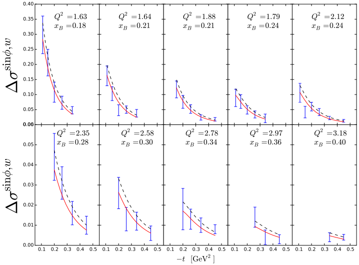
thus cancelling strongly oscillating factors $1/(\mathcal{P}_1(\phi)\mathcal{P}_2(\phi))$ in Bethe-Heitler and interference terms in $d\sigma$.
- (Since $\mathcal{P}_{1,2}(\phi)$ depend also on y, weighting may be undesirable if you are after Rosenbluth separation, see [Achenauer, Fazio, K.K., Müller '13] for EIC case.)

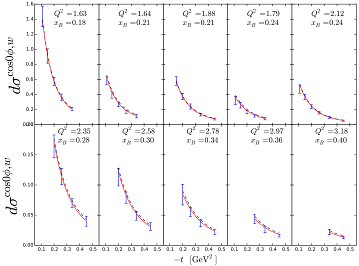
Weighted harmonics

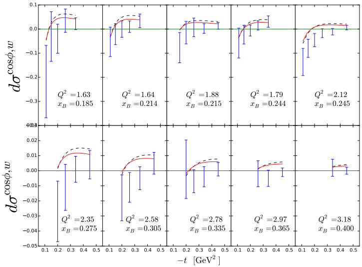
- Series of such weighted harmonic terms converges then faster with increasing n than normal Fourier series.
- Example of first four bins of Hall A (2015) cross-section measurement:



Case #1: CLAS cross-section (n-space)



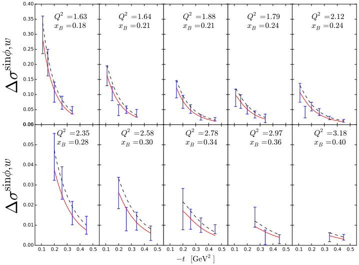


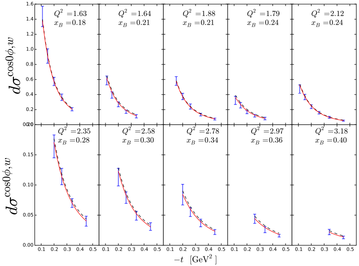


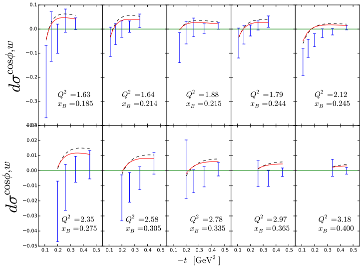
- $\chi^2/\mathrm{npts} = 62.2/48$ for $d\sigma^{\cos\phi,w}$

(O.K. but not so perfect as in ϕ-space)

Case #1: CLAS cross-section (n-space)







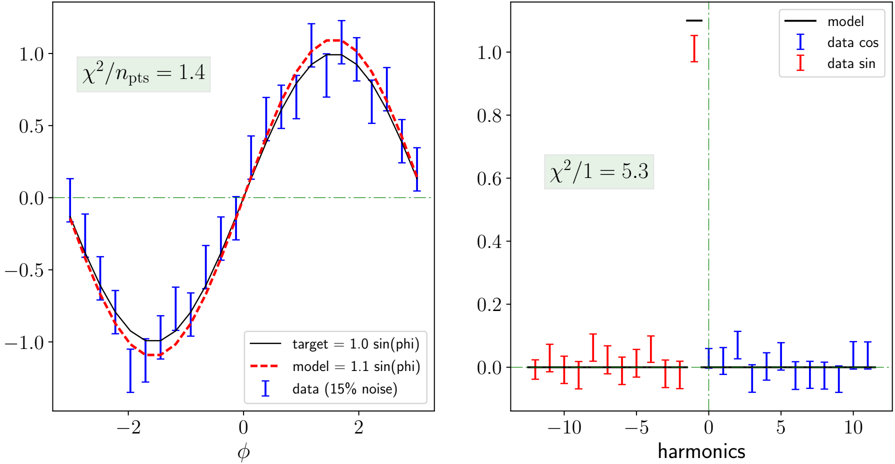
- $\chi^2/\mathrm{npts} = 62.2/48$ for $d\sigma^{\cos\phi,w}$

(O.K. but not so perfect as in ϕ-space)

In addition to $\chi^2/\mathrm{d.o.f.}$ use:

$$\text{pull} \equiv \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \frac{f(x_i) - y_i}{\Delta y_i}$$

Case #2: Sine toy model



- ϕ-space view can be misleading

Extraction of harmonics

- General model for cross section or any other observable

$$f(\phi) = \sum_{k=0}^{c_{\max}} c_k \cos(k\phi) + \sum_{k=1}^{s_{\max}} s_k \sin(k\phi)$$

- $c_{\max} + s_{\max} + 1$ real parameters to be determined from N equidistantly measured values $f(\phi_{i=1,\ldots,N})$.
- least-squares optimization gives standard Fourier formulas, e.g., for $k \geq 1$,

$$c_k = \frac{2}{N} \sum_{i=1}^{N} f(\phi_i) cos(k\phi_i)$$

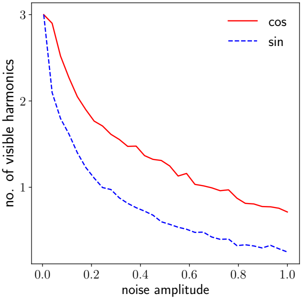
- c_k and their uncertainties are independent of $c_{\max}, s_{\max}$, thanks to orthogonality of trig functions (but only if you have measured full set of ϕ bins!)

How many harmonics are there?

- Case #3: Target function is

$$f(\phi) = 0.42 - 0.28\cos(\phi) + 0.08\cos(2\phi) + 0.02\cos(3\phi) - 0.13\sin(\phi) - 0.03\sin(2\phi) + 0.006\sin(3\phi)$$

  and is used to generate data with variable noise
- Fitting Fourier series with increasing number of harmonics until $\chi^2/\mathrm{d.o.f.}$ w.r.t. target function starts to deteriorate

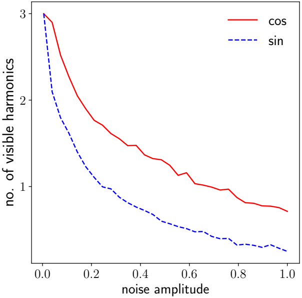


How many harmonics are there?

- Case #3: Target function is

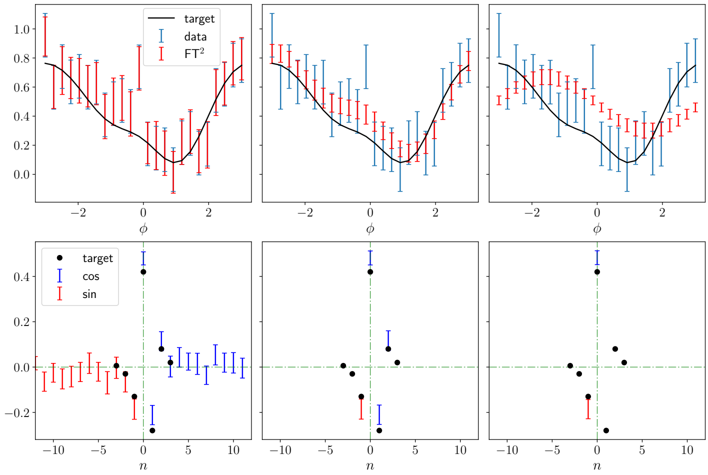
$$f(\phi) = 0.42 - 0.28\cos(\phi) + 0.08\cos(2\phi) + 0.02\cos(3\phi) - 0.13\sin(\phi) - 0.03\sin(2\phi) + 0.006\sin(3\phi)$$

  and is used to generate data with variable noise
- Fitting Fourier series with increasing number of harmonics until $\chi^2/\mathrm{d.o.f.}$ w.r.t. target function starts to deteriorate



But in real life we don't know target function! How to determine number of harmonics using only data?

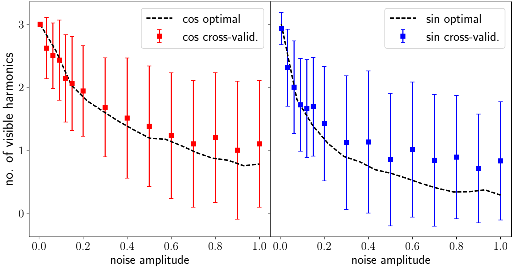
Bias-variance trade-off



Cross-validation procedure

To determine the optimal number of harmonics using only data:

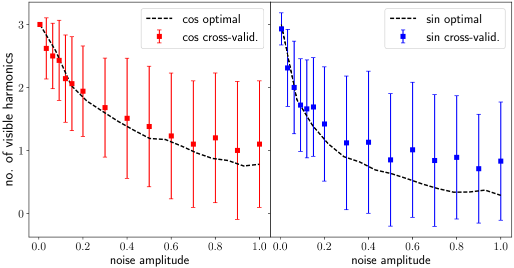
- Divide data randomly in n subsets
- Leave one out and fit the model to remaining $n-1$ subsets
- Evaluate model (calculating χ^2) on the left-out set
- Repeat with leaving out other $n-1$ sets and average the error
- Repeat with increasing number of harmonics until description of left-out sets starts to deteriorate



Cross-validation procedure

To determine the optimal number of harmonics using only data:

- Divide data randomly in n subsets
- Leave one out and fit the model to remaining $n-1$ subsets
- Evaluate model (calculating χ^2) on the left-out set
- Repeat with leaving out other $n-1$ sets and average the error
- Repeat with increasing number of harmonics until description of left-out sets starts to deteriorate



Finaly, determine just optimal harmonics using all data.

Propagation of uncertainties to harmonics

- Consider three types of uncertainty:
  1. uncorrelated point-to-point uncertainty (absolute size ϵ)
  2. correlated normalization uncertainty (relative size ϵ)
  3. correlated modulated (ϕ-dependent) uncertainty (e.g., relative size $\epsilon \cos(\phi)$)
- Uncorrelated uncertainty: $\Delta c_k = \sqrt{2/N}\,\epsilon$
- Normalization uncertainty: $\Delta c_k / c_k = \epsilon$
- Correlated modulated uncertainty: more complicated, but for hierarchical case $c_0 \gg c_1 \gg \cdots$ one obtains

$$\frac{\Delta c_0}{c_0} = \frac{c_1}{2c_0}\epsilon, \qquad \frac{\Delta c_1}{c_1} = \frac{c_0}{c_1}\epsilon$$

i.e. we have enhancement of uncertainty for subleading harmonics!

$$(c_0 + c_1 \cos\phi + \cdots) \times (1 + \epsilon \cos\phi) = c_0 \left(1 + \frac{c_1}{2c_0}\epsilon\right) + c_1 \left(1 + \frac{c_0}{c_1}\epsilon\right) \cos\phi$$

Modulated correlated error in the wild

Hall A [M. Defurne et. al 2015] discussing systematic uncertainties:

tematic error from the parameter choice to be 1%.

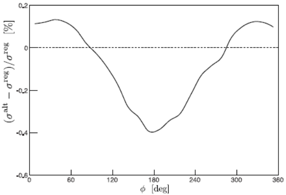
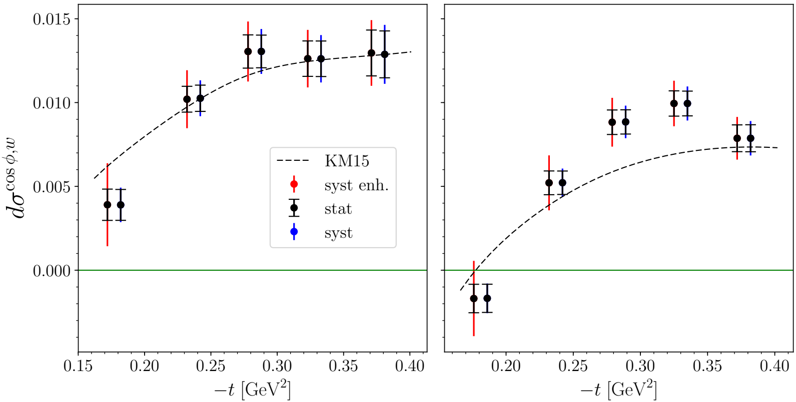


FIG. 20. Difference in % between the cross section extracted with the squared DVCS amplitude term and with the $\Re e\, \mathcal{C}^{\mathcal{I},V}$ term for $x_B = 0.37$, $Q^2 = 2.36$ GeV^2 and $-t = 0.33$ GeV^2. The ϕ-profile of the difference is a consequence of the small $\cos\phi$ and $\cos 2\phi$ dependences of the $\Re e\, \mathcal{C}^{\mathcal{I},V}$ kinematic coefficient. Both extractions give almost the same reduced χ^2/dof=0.94 (nominal) and 0.93 (alternate) for the entire Kin2 setting.

Systematic uncertainty	Value	Section
HRS acceptance cut	1%	IV A
Electron ID	0.5%	IV D
HRS multitrack	0.5%	IV D
Multi-cluster	0.4%	IV D
Corrected luminosity	1%	IV D
Fit parameters	1%	VI B
Radiative corrections	2%	V
Beam polarization	2%	III A 3
Total (helicity-independent)	2.8%	
Total (helicity-dependent)	3.4%	

TABLE V. Normalization systematic uncertainties in the extracted photon electroproduction cross sections. The systematic error coming from the fit parameter choice is not a normalization error per se, but we consider that 1% is an upper limit for this error on all kinematic bins. The helicity-dependent cross sections have an extra uncertainty stemming from the beam polarization measurement. The last column gives the section in which each systematic effect is discussed.

Hall A cos(ϕ) harmonics



(Syst added *linearly* on top of stat.)

Hybrid GPD models for global fits

- Sea quarks and gluons modelled using SO(3) partial wave expansion in conformal GPD moment space + $\mathcal{Q}^2$ evolution.
- Valence quarks — model CFFs directly (ignoring $\mathcal{Q}^2$ evolution):

$$\mathfrak{Im}\,\mathcal{H}(\xi,t) = \pi\left[\frac{4}{9}H^{u_{\rm val}}(\xi,\xi,t) + \frac{1}{9}H^{d_{\rm val}}(\xi,\xi,t) + \frac{2}{9}H^{\rm sea}(\xi,\xi,t)\right]$$

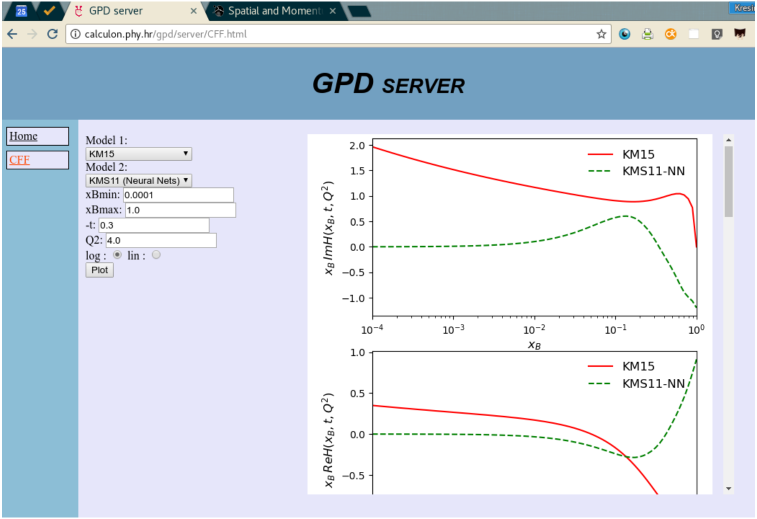
$$H(x,x,t) = n\,r\,2^{\alpha}\left(\frac{2x}{1+x}\right)^{-\alpha(t)}\left(\frac{1-x}{1+x}\right)^{b}\frac{1}{\left(1-\frac{1-x}{1+x}\frac{t}{M^2}\right)^{p}}\,.$$

- $\mathfrak{Re}\,\mathcal{H}$ determined by dispersion relations
- 15 free parameters in total for H, $\tilde{H}$, E, $\tilde{E}$.

Model	KM09a	KM09b	KM10	KM10a	KM10b	KMS11	KMM12	KM15
free params.	{3}+(3)+5	{3}+(3)+6	{3}+15	{3}+10	{3}+15	NNet	{3}+15	{3}+15
χ^2/d.o.f.	32.0/31	33.4/34	135.7/160	129.2/149	115.5/126	13.8/36	123.5/80	240./275
F_2	{85}	{85}	{85}	{85}	{85}		{85}	{85}
$\sigma_{\rm DVCS}$	(45)	(45)	51	51	45		11	11
$d\sigma_{\rm DVCS}/dt$	(56)	(56)	56	56	56		24	24
$A_{LU}^{\sin\phi}$	12+12	12+12	12	16	12+12		4	13
$A_{LU,I}^{\sin\phi}$			18	18		18	6	6
$A_{C}^{\cos 0\phi}$							6	6
$A_{C}^{\cos\phi}$	12	12	18	18	12	18	6	6
$\Delta\sigma^{\sin\phi,w}$			12				12	63
$\sigma^{\cos 0\phi,w}$			4				4	58
$\sigma^{\cos\phi,w}$			4				4	58
$\sigma^{\cos\phi,w}/\sigma^{\cos 0\phi,w}$		4			4			
$A_{UL}^{\sin\phi}$							10	17
$A_{LL}^{\cos 0\phi}$							4	14
$A_{LL}^{\cos\phi}$								10
$A_{UT,I}^{\sin(\phi-\phi_S)\cos\phi}$							4	4

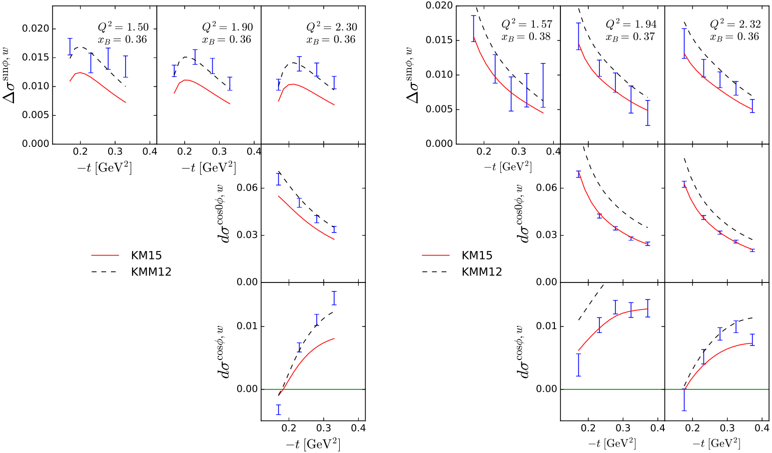
- [K.K., Müller, et al. '09–'15]
- These models are publicly available (google for "gpd page")

KM GPD server



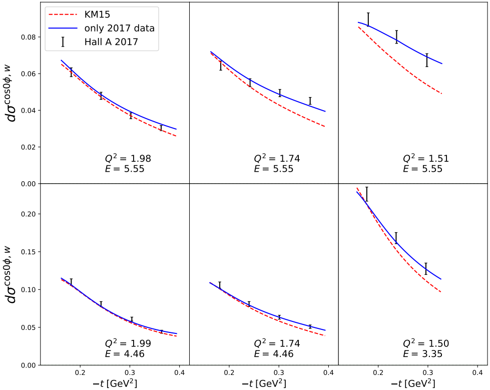
- Plots of all CFFs available; numerical values soon to come . . .

2006 vs 2015 Hall A cross-sections



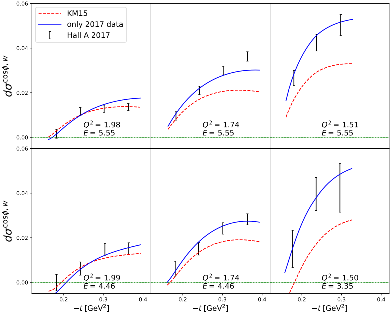
- Improvement of global $\chi^2/\mathrm{d.o.f.}$ 123.5/80 → 240./275

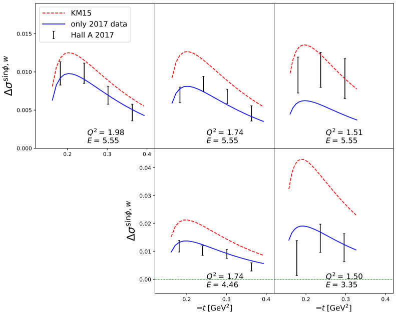
2015 vs 2017 Hall A cross-sections (1/2)



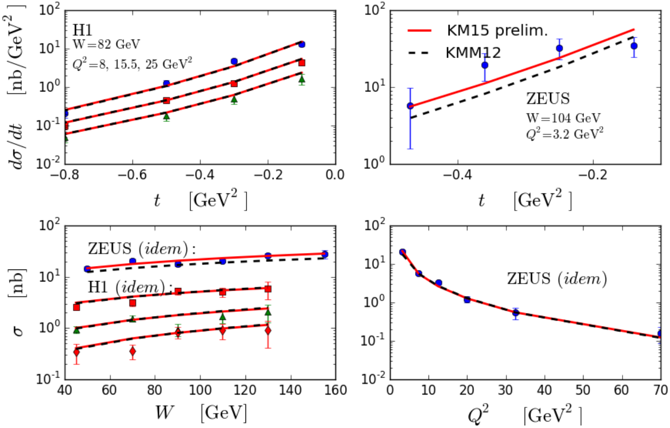
- Predictions of KM15 model fail for low-Q^2 2017 Hall A measurements!

2015 vs 2017 Hall A cross-sections (2/2)

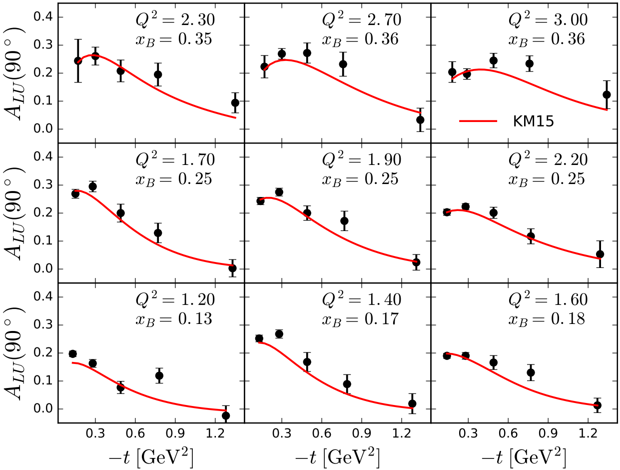




H1 (2007), ZEUS (2008)

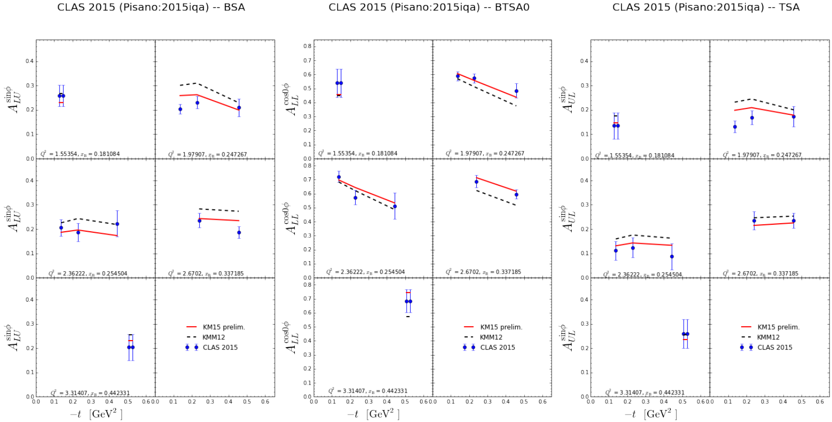


2007 CLAS beam spin asymmetry



- Only data with $|t| \leq 0.3\,\mathrm{GeV}^2$ used for fits.

2015 CLAS asymmetries

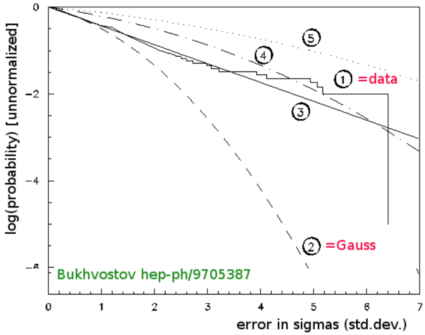


$\chi^2/npts$ and pull values

Collaboration	Observable	$n_{\rm pts}$	KMM12		KM15	
			$\chi^2/n_{\rm pts}$	pull	$\chi^2/n_{\rm pts}$	pull
ZEUS	$\sigma_{\rm DVCS}$	11	0.49	-1.76	0.51	-1.74
ZEUS,H1	$d\sigma_{\rm DVCS}/dt$	24	0.97	0.85	1.04	1.37
HERMES	$A_{\rm C}^{\cos 0\phi}$	6	1.31	0.49	1.24	0.29
HERMES	$A_{\rm C}^{\cos\phi}$	6	0.24	-0.56	0.07	-0.20
HERMES	$A_{\rm LU,I}^{\sin\phi}$	6	2.08	-2.52	1.34	-1.28
CLAS	$A_{\rm LU}^{\sin\phi}$	4	1.28	2.09		
CLAS	$A_{\rm LU}^{\sin\phi}$	13			1.24	0.63
CLAS	$\Delta\sigma^{\sin\phi,w}$	48			0.41	-1.66
CLAS	$d\sigma^{\cos 0\phi,w}$	48			0.16	-0.21
CLAS	$d\sigma^{\cos\phi,w}$	48			1.16	6.36
Hall A	$\Delta\sigma^{\sin\phi,w}$	12	1.06	-2.55		
Hall A	$d\sigma^{\cos 0\phi,w}$	4	1.21	2.14		
Hall A	$d\sigma^{\cos\phi,w}$	4	3.49	-0.26		
Hall A	$\Delta\sigma^{\sin\phi,w}$	15			0.81	-2.84
Hall A	$d\sigma^{\cos 0\phi,w}$	10			0.40	0.92
Hall A	$d\sigma^{\cos\phi,w}$	10			2.52	-2.42
HERMES,CLAS	$A_{\rm UL}^{\sin\phi}$	10	1.90	-1.89	1.10	-1.94
HERMES	$A_{\rm LL}^{\cos 0\phi}$	4	3.44	2.17	3.19	1.99
HERMES	$A_{\rm UT,I}^{\sin(\phi-\phi_S)\cos\phi}$	4	0.90	0.61	0.90	0.71
CLAS	$A_{\rm UL}^{\sin\phi}$	10			0.76	0.38
CLAS	$A_{\rm LL}^{\cos 0\phi}$	10			0.50	-0.22
CLAS	$A_{\rm LL}^{\cos\phi}$	10			1.54	2.40

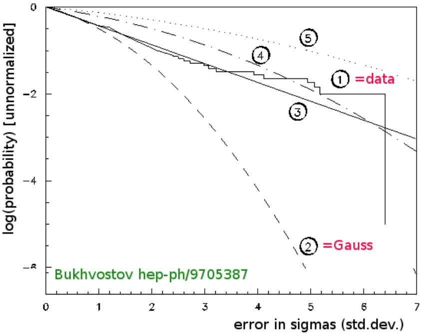
Problems with standard fitting approaches

1. Choice of fitting function introduces theoretical bias leading to systematic error which cannot be estimated (and is likely much larger for GPDs(x, η, t) than for PDFs(x).
2. Propagation of uncertainties from experiment to fitted function is difficult. Errors in actual experiments are not always Gaussian.

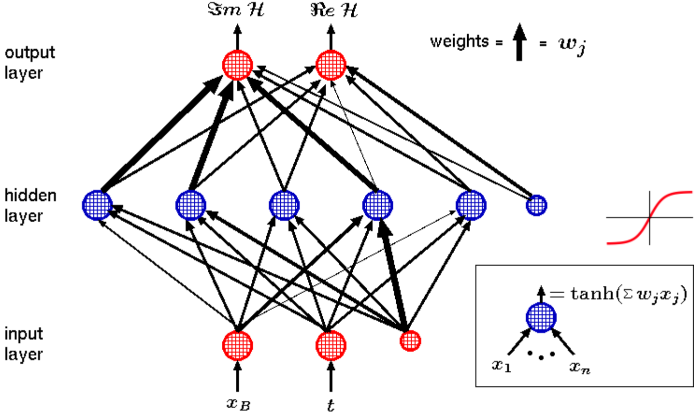


Problems with standard fitting approaches

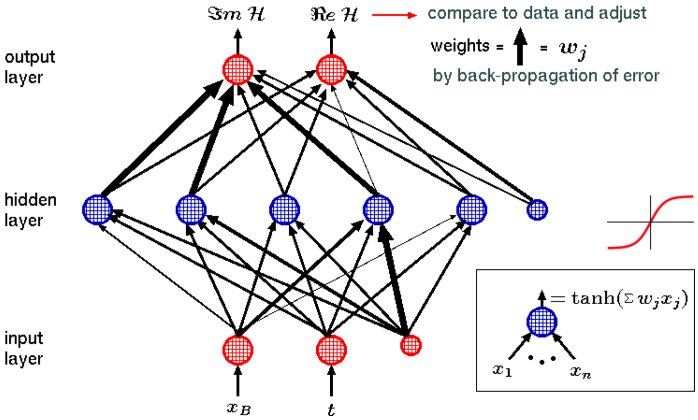
1. Choice of fitting function introduces theoretical bias leading to systematic error which cannot be estimated (and is likely much larger for GPDs(x, η, t) than for PDFs(x).$\rightarrow$ NNets
2. Propagation of uncertainties from experiment to fitted function is difficult. Errors in actual experiments are not always Gaussian.$\rightarrow$ Monte Carlo error propagation



Multilayer perceptron



Multilayer perceptron



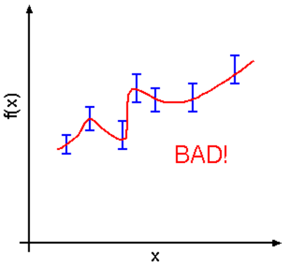
- Essentially a least-square fit of a complicated many-parameter function. $f(x) = \tanh(\sum w_i \tanh(\sum w_j \cdots))$ $\Rightarrow$ no theory bias

Function fitting by a neural net

- Theorem: Given enough neurons, any smooth function $f(x_1, x_2, \cdots)$ can be approximated to any desired accuracy. Single hidden layer is sufficient (but not always most efficient).

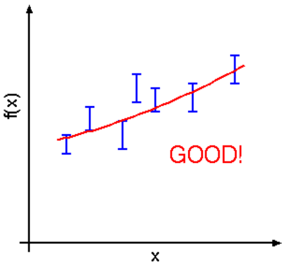
Function fitting by a neural net

- Theorem: Given enough neurons, any smooth function $f(x_1, x_2, \cdots)$ can be approximated to any desired accuracy. Single hidden layer is sufficient (but not always most efficient).
- With simple training of neural nets to data there is a danger of overfitting (a.k.a. overtraining)



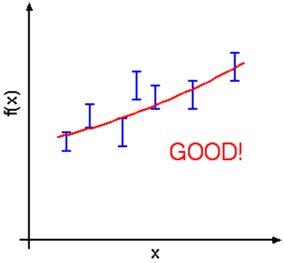
Function fitting by a neural net

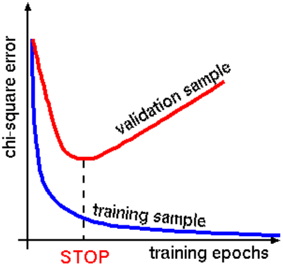
- Theorem: Given enough neurons, any smooth function $f(x_1, x_2, \cdots)$ can be approximated to any desired accuracy. Single hidden layer is sufficient (but not always most efficient).
- With simple training of neural nets to data there is a danger of overfitting (a.k.a. overtraining)



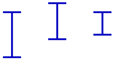
Function fitting by a neural net

- Theorem: Given enough neurons, any smooth function $f(x_1, x_2, \cdots)$ can be approximated to any desired accuracy. Single hidden layer is sufficient (but not always most efficient).
- With simple training of neural nets to data there is a danger of overfitting (a.k.a. overtraining)
- Solution: Divide data (randomly) into two sets: *training* sample and *validation sample*. Stop training when error of validation sample starts increasing.

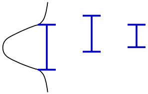




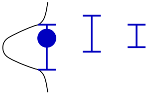
Monte Carlo propagation of errors



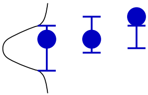
Monte Carlo propagation of errors



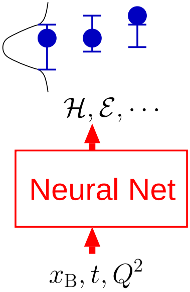
Monte Carlo propagation of errors



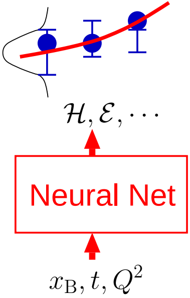
Monte Carlo propagation of errors



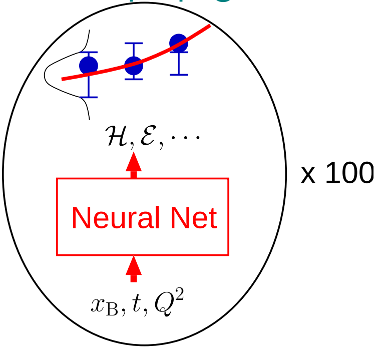
Monte Carlo propagation of errors



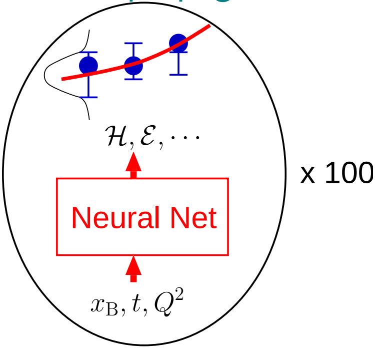
Monte Carlo propagation of errors



Monte Carlo propagation of errors



Monte Carlo propagation of errors



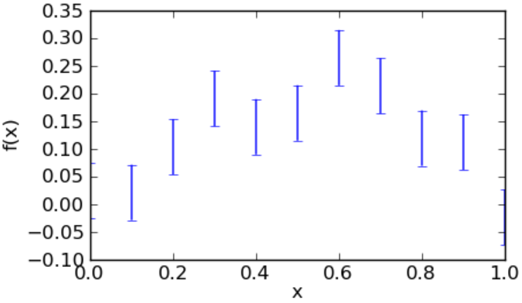
- Set of N_{rep} NNs defines a probability distribution in space of possible CFF functions:

$$\left\langle \mathcal{F}[\mathcal{H}] \right\rangle = \int \mathcal{D}\mathcal{H}\, \mathcal{P}[\mathcal{H}]\, \mathcal{F}[\mathcal{H}] = \frac{1}{N_{rep}} \sum_{k=1}^{N_{rep}} \mathcal{F}[\mathcal{H}^{(k)}]\,, \qquad (1)$$

- Experimental uncertainties and their correlations are preserved [Giele et al. '01]

Fitting example

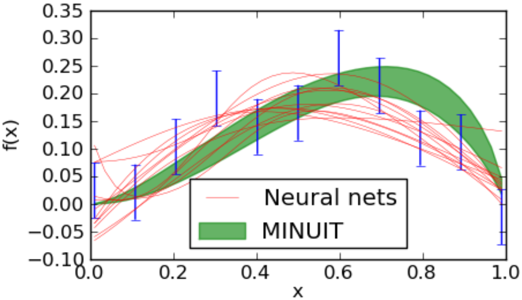
- Take some (fake) data



- Fit with
  1. Standard Minuit fit with ansatz $f(x) = x^a(1-x)^b$
  2. Neural network fit

Fitting example

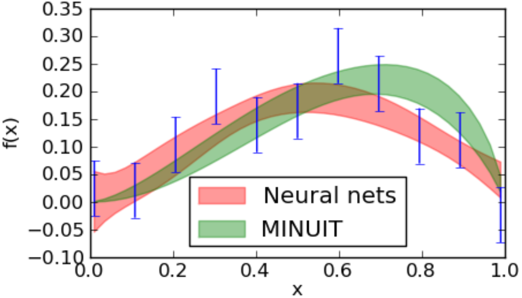
- Take some (fake) data



- Fit with
  1. Standard Minuit fit with ansatz $f(x) = x^a(1-x)^b$
  2. Neural network fit

Fitting example

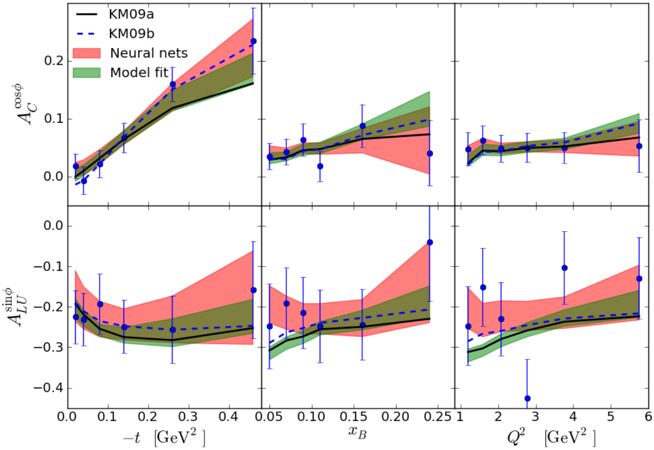
- Take some (fake) data



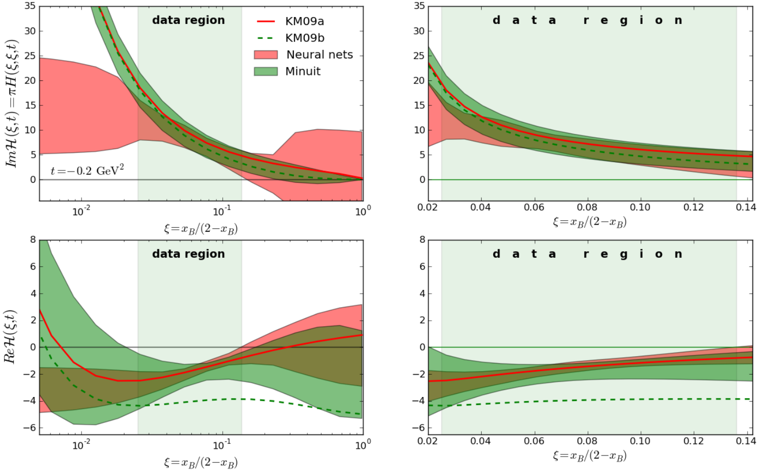
- Fit with
  1. Standard Minuit fit with ansatz $f(x) = x^a(1-x)^b$
  2. Neural network fit

Fit to HERMES BSA+BCA data

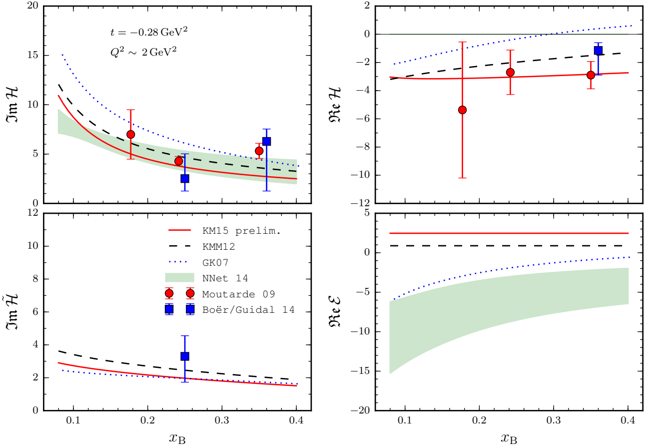
- 50 neural nets with 13 neurons in a single hidden layer



Resulting neural network CFFs



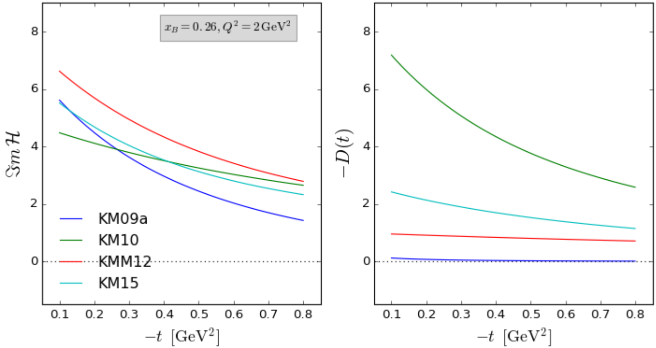
CFFs from various fits



(from [K.K., Moutarde and Liuti '16])

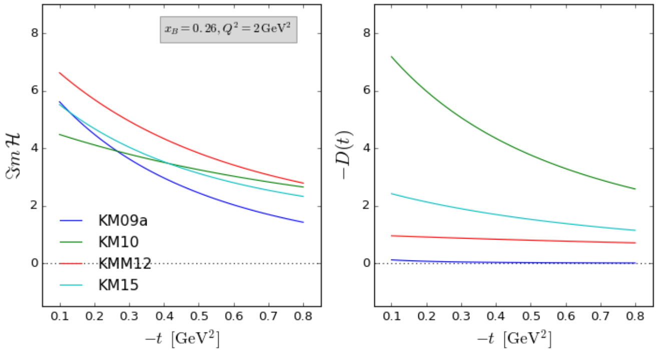
CFFs from various fits (II)

Comparing $\Im\mathfrak{m}\,\mathcal{H}$ and D-term for a typical JLab kinematics:



CFFs from various fits (II)

Comparing $\mathfrak{Im}\,\mathcal{H}$ and D-term for a typical JLab kinematics:



- We need access to ERBL region.

Going beyond first approximations

Published DVCS data (apart from new 2017 Hall A) is well described within leading order, leading twist and other approximations. Some available corrections:

- NLO evolution and NLO coefficient functions
- finite-t and target mass corrections [Braun et al. '14]
- twist-3 GPDs
- massive charm [Noritzsch '03]
- small-x resummation

can maybe be absorbed in redefinition of GPDs (think "DVCS scheme" factorization), but have to be taken into account when working with more processes (striving towards universal GPDs) and with more precise future data.

Summary

- Global fits of all proton DVCS data using flexible hybrid models were in healthy shape until 2017 (some tension for first $cos(\phi)$ harmonic of JLab cross sections).
- New Hall A 2017 data present a challenge
- Standard global model fitting and neural networks approach are complementing each other

Summary

- Global fits of all proton DVCS data using flexible hybrid models were in healthy shape until 2017 (some tension for first $cos(\phi)$ harmonic of JLab cross sections).
- New Hall A 2017 data present a challenge
- Standard global model fitting and neural networks approach are complementing each other

The End