Likelihood Analysis of DVCS Compton Form Factors

Talk by Douglas Adams

within EXCLAIM collaboration, postdoc at UVA,
soon publishing paper with co-authors:

Douglas Q. Adams,[1,*] Joshua Bautista,[1,†] Marija Čuić,[1,‡] Adil Khawaja,[1,§] Saraswati Pandey,[1,¶] Zaki Panjsheeri,[1,**] Gia-Wei Chern,[1,††] Simonetta Liuti,[1,‡‡] Yaohang Li,[2,§§] Marie Boër,[3,§§] Michael Engelhardt,[4,§§] Gary R. Goldstein,[5,§§] Huey-Wen Lin,[6,§§] and Mattew D. Sievert[4,¶¶]

Outline in Story Form

GOAL: Use DVCS data and comparing to cross section model to find CFFs

- We find a CFF result using VAIM: Got some valid CFFs
- Curve fit: A really bad result: Encounter a problem 1!
- Definition of the likelihood: Try to fix the problem
- Canonical Likelihood: Reproduces the problem in explainable way
- Canonical Likelihood Modified: Fix the problem in 2 ways
  - Difference method Likelihood
  - Canonical Likelihood
- Encounter a problem 2!: Poll the audience
- Some results: Table of CFFs and errors

DVCS Kinematic Parameters:

- $s = (k + p)^2$ is the electron-proton center of mass energy squared,
- $Q^2 = -(k - k')^2 = -q^2$ is the four-momentum transfer squared between the incoming and outgoing electrons,
- $x_{Bj} = Q^2/2(pq)$ is Bjorken-x. In the asymptotic limit, disregarding t/Q^2 ans M^2/Q^2 corrections, x_{Bj} is written in terms the skewness parameter, $\xi = -(\Delta q)/[(pq) + (p'q)] = x_{Bj}/(2 - x_{Bj})$
- $t = (p - p')^2 = (q' - q)^2$ is the four-momentum transfer squared between the initial and final protons (q' is the final photon four-momentum),
- ϕ is the azimuthal angle between the planes defined by the electron momenta, by the final proton, p', and photon, q'.

https://github.com/kkumer/gepard/blob/master/src/gepard/datasets/ep2epgamma/BSS-HALLA-18.dat

doi:10.1038/s41567-019-0774-3

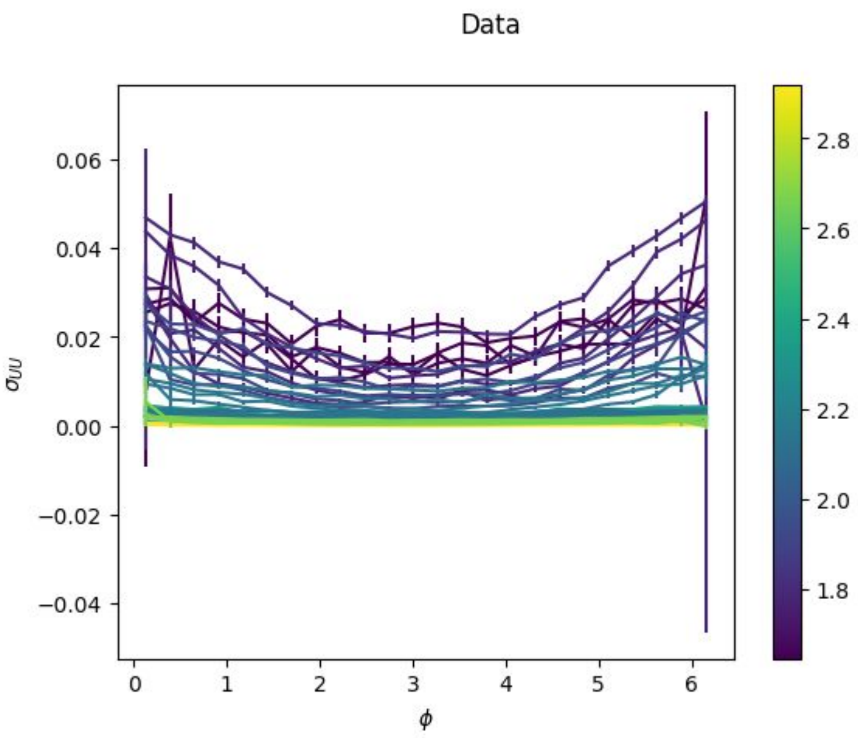
Data: BSS-Hall-A-18 (used from gepard)

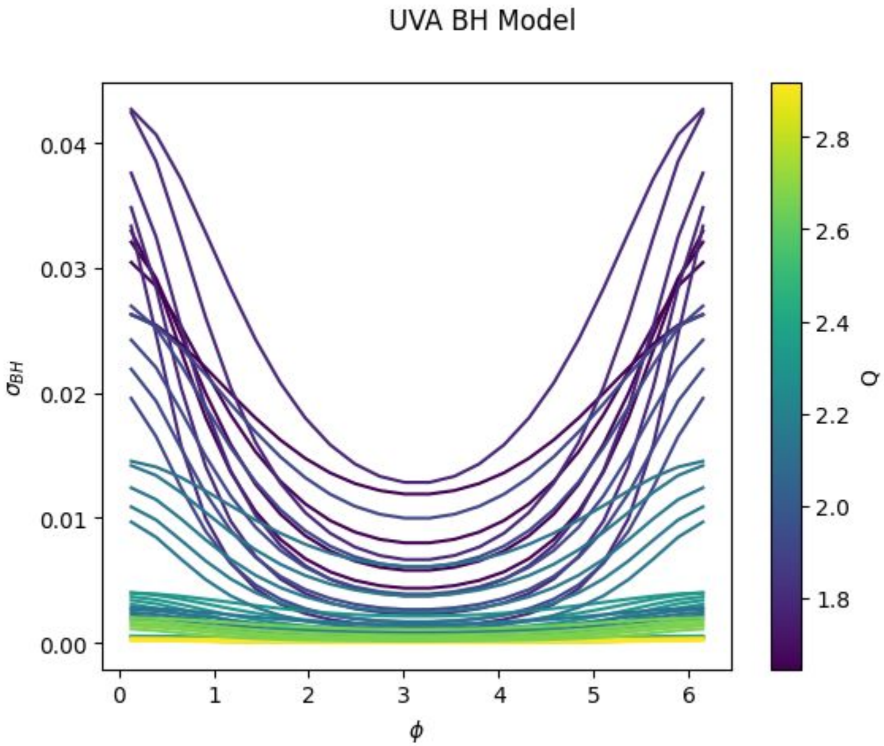
	Eb	x	Q	t	phi	XUU	errstat	XBH	XUU-XBH
384	4.487	0.483	1.646208	-0.3906	0.130900	0.02549	0.0035	0.030466	-0.004976
385	4.487	0.483	1.646208	-0.3906	0.392699	0.02750	0.0032	0.028554	-0.001054
386	4.487	0.483	1.646208	-0.3906	0.654498	0.02570	0.0028	0.025363	0.000337
387	4.487	0.483	1.646208	-0.3906	0.916298	0.02224	0.0025	0.021723	0.000517
388	4.487	0.483	1.646208	-0.3906	1.178097	0.02092	0.0022	0.018266	0.002654
389	4.487	0.483	1.646208	-0.3906	1.439897	0.02043	0.0021	0.015309	0.005121
390	4.487	0.483	1.646208	-0.3906	1.701696	0.01553	0.0020	0.012941	0.002589
391	4.487	0.483	1.646208	-0.3906	1.963495	0.01764	0.0020	0.011128	0.006512
392	4.487	0.483	1.646208	-0.3906	2.225295	0.01629	0.0018	0.009797	0.006493
393	4.487	0.483	1.646208	-0.3906	2.487094	0.01423	0.0017	0.008871	0.005359
394	4.487	0.483	1.646208	-0.3906	2.748894	0.01430	0.0018	0.008288	0.006012
395	4.487	0.483	1.646208	-0.3906	3.010693	0.01194	0.0017	0.008007	0.003933
396	4.487	0.483	1.646208	-0.3906	3.272492	0.01639	0.0018	0.008007	0.008383
397	4.487	0.483	1.646208	-0.3906	3.534292	0.01862	0.0019	0.008288	0.010332
398	4.487	0.483	1.646208	-0.3906	3.796091	0.01804	0.0019	0.008871	0.009169
399	4.487	0.483	1.646208	-0.3906	4.057891	0.01375	0.0019	0.009797	0.003953
400	4.487	0.483	1.646208	-0.3906	4.319690	0.01540	0.0020	0.011128	0.004272
401	4.487	0.483	1.646208	-0.3906	4.581489	0.02347	0.0022	0.012941	0.010529
402	4.487	0.483	1.646208	-0.3906	4.843289	0.02409	0.0024	0.015309	0.008781
403	4.487	0.483	1.646208	-0.3906	5.105088	0.02080	0.0024	0.018266	0.002534
404	4.487	0.483	1.646208	-0.3906	5.366887	0.02711	0.0026	0.021723	0.005387
405	4.487	0.483	1.646208	-0.3906	5.628687	0.02834	0.0029	0.025363	0.002977
406	4.487	0.483	1.646208	-0.3906	5.890486	0.02480	0.0031	0.028554	-0.003754
407	4.487	0.483	1.646208	-0.3906	6.152286	0.02869	0.0034	0.030466	-0.001776

Same kinematics

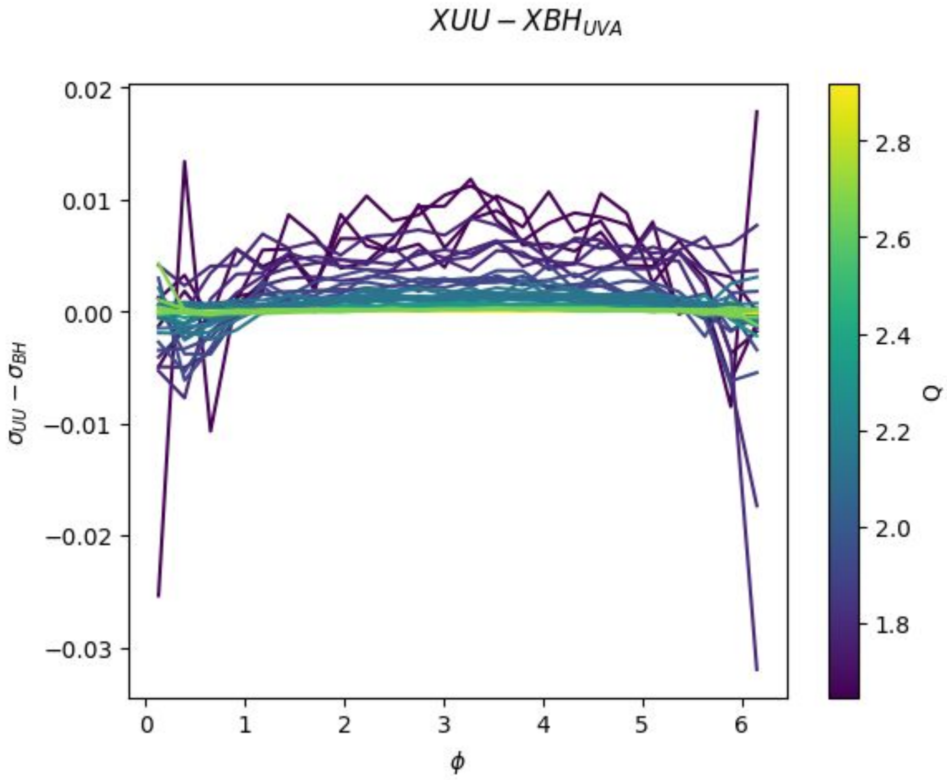
Different phi's and cross sections

UVA Model: (vs Cross Section Data)





UVA Model: Residual Non BH Contribution

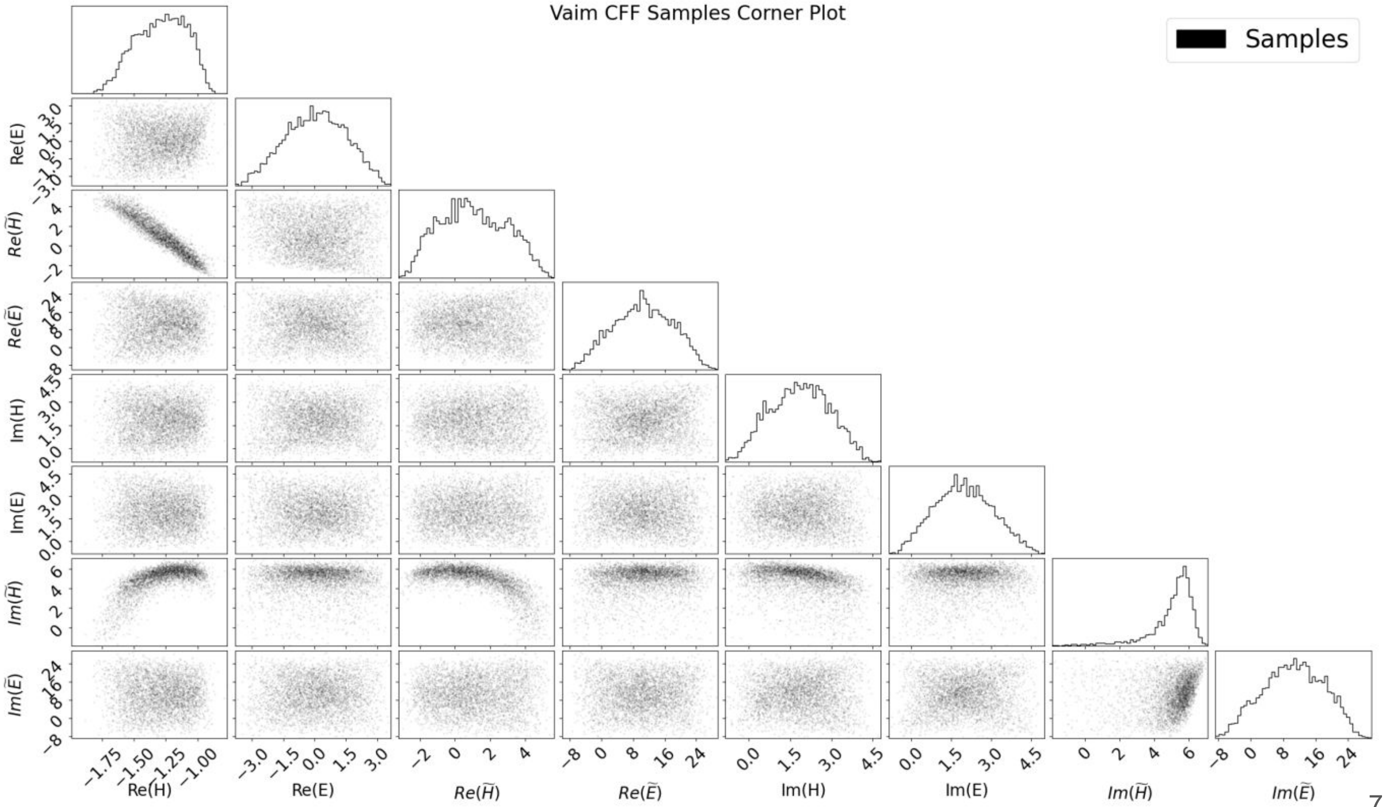


VAIM Result Using Prior for CFFs

- Apply cross section equation as constraint with observed data

- Include a prior

- Generate random but viable CFFs which try to satisfy the constraint



https://arxiv.org/pdf/2405.05826

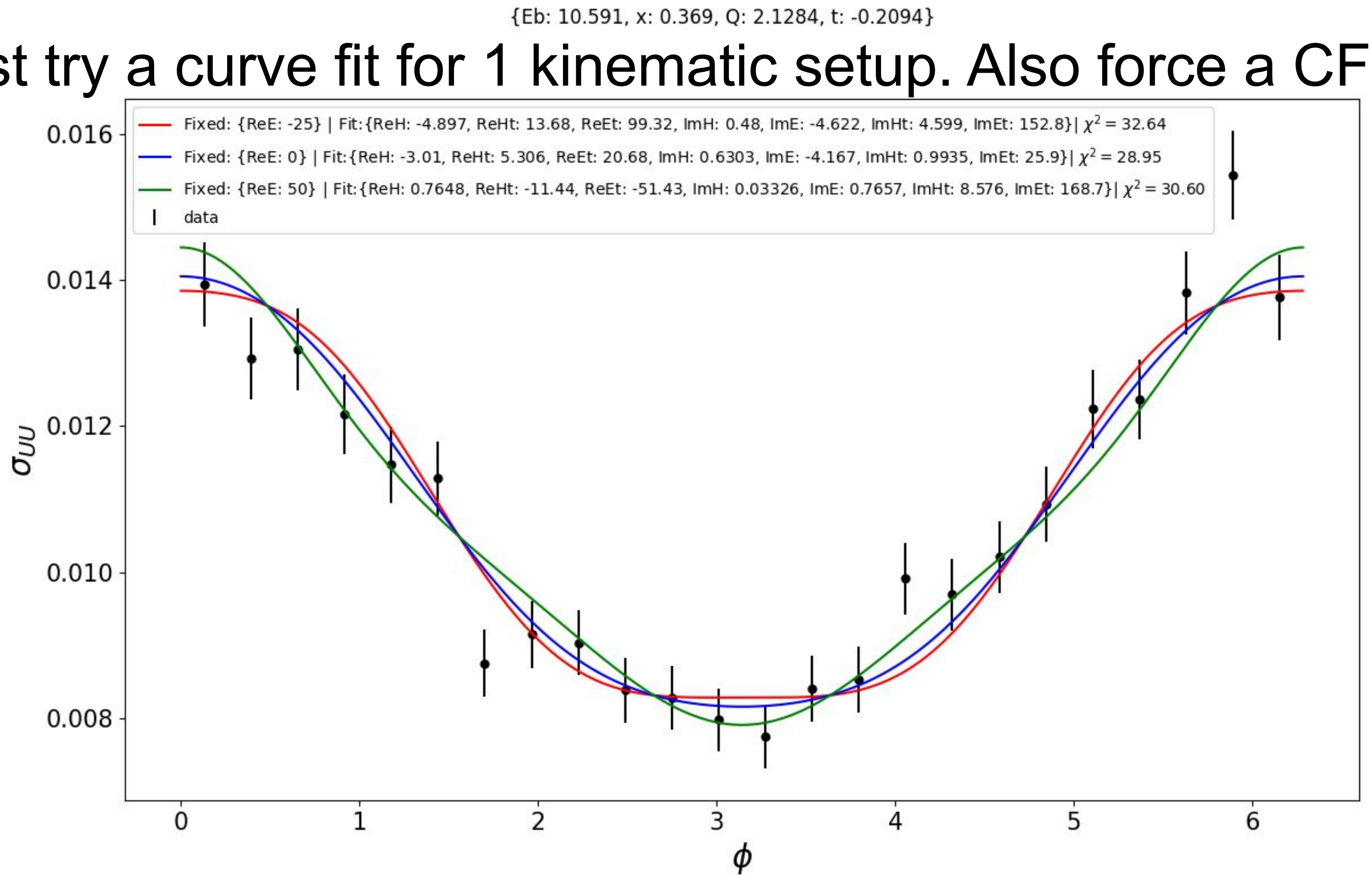
VAIM CAVEATS -> Motivate a likelihood analysis

- Requires a prior for the CFFs
- Assume the same CFFs work for many different kinematic bins
- Ignore the error bars on the data

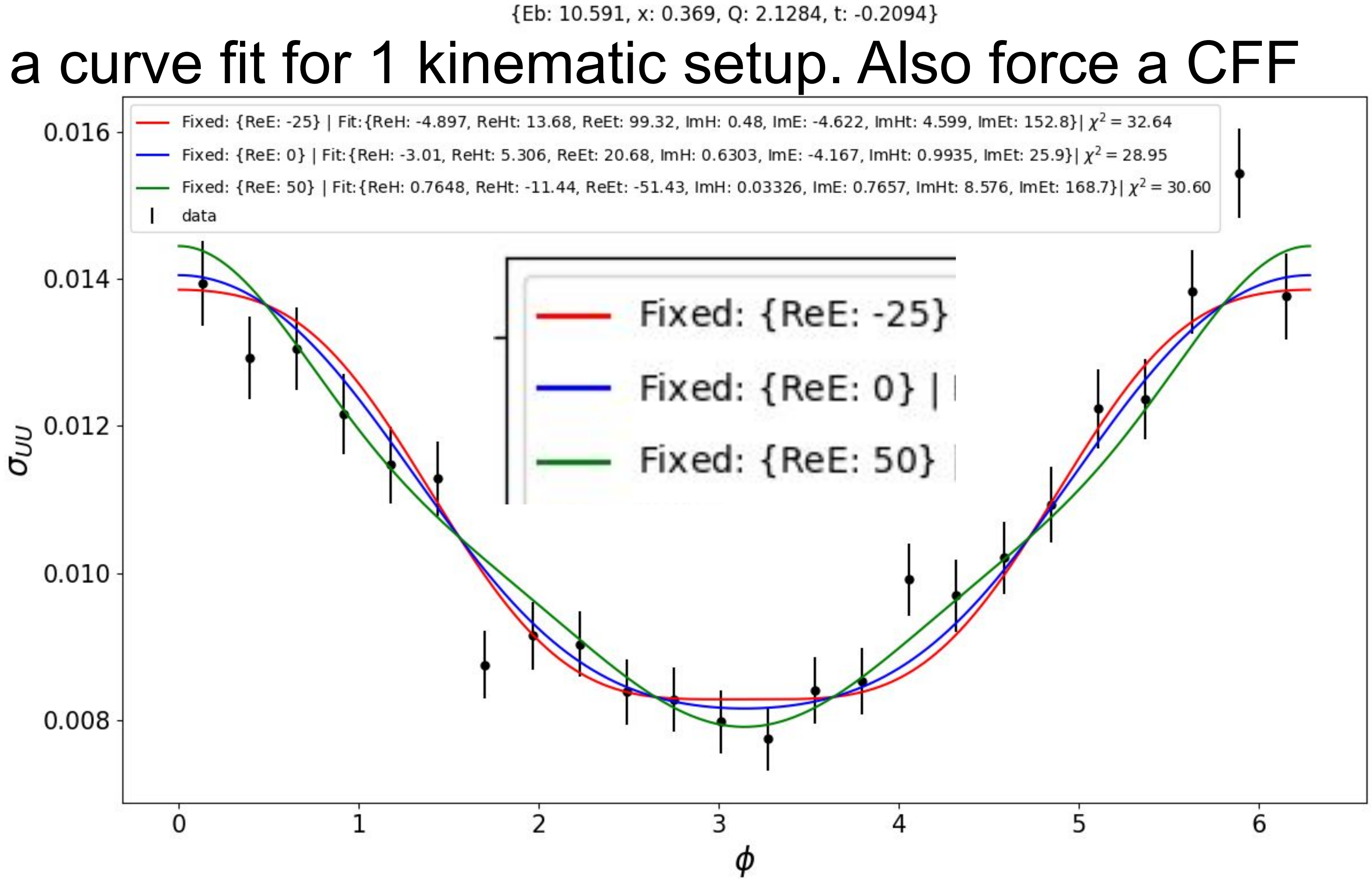
It would be nice to not have these caveats

https://arxiv.org/pdf/2405.05826

First try a curve fit for 1 kinematic setup. Also force a CFF



Try a curve fit for 1 kinematic setup. Also force a CFF



Likelihood function: A careful wordy definition

The likelihood function is

a conditional probability density function

which describes the probability of a set of parameters being correct

given some observations of data from any number of experiments

while assuming that a given model is true,

and that all experiments measurement error of data is known

Likelihood function: Use bayes law

$$\underbrace{\underbrace{\boxed{\vec{X}_{all}\&\vec{\Theta}}_{pdf}(\vec{v}_{x\,all},\vec{v}_{\theta})}_{\underbrace{Joint}_{\text{Omited In Textbooks}}} = \underbrace{\boxed{\vec{\Theta}|\vec{X}_{all}}_{pdf}(\vec{v}_{\Theta}|\vec{v}_{x\,all})}_{=\underbrace{Posterior}_{\text{(all data)}}} \times \underbrace{\boxed{\vec{X}_{all}}_{pdf}(\vec{v}_{x\,all})}_{\times\underbrace{Evidence}_{(=1)}} = \underbrace{\boxed{\vec{X}_{all}|\vec{\Theta}}_{pdf}(\vec{v}_{x\,all},\vec{v}_{\Theta})}_{=\underbrace{Likelihood}_{\text{(all data)}}} \times \underbrace{\boxed{\vec{\Theta}}_{pdf}(\vec{v}_{\Theta})}_{\times\underbrace{Prior}_{\text{(no data)}}}}$$

For frequentists: Prior = 1

$$\text{Likelihood} = \boxed{\vec{X}_{all}|\vec{\Theta}}_{pdf} = \prod_i \boxed{\vec{X}_{single}|\vec{\Theta}}_{pdf}(\vec{v}_{xi},\vec{v}_{\Theta}) = \text{ is model and experiment error determined}$$

$$\text{Evidence} = \boxed{\vec{X}_{all}}_{pdf} = \prod_i \boxed{\vec{X}_{single}}_{pdf}(\vec{v}_{xi}) = 1 \text{ (there is one universe)}$$

Canonical Likelihood Derivation

$$\mathcal{L}_{canonical}(\text{parameters}) = \prod_i \text{Gaussian}(x = \sigma_{obs}(\phi_i), \mu = \sigma_{model}(\phi_i), \sigma = Err(\sigma_{obs}))$$

Each error bar:

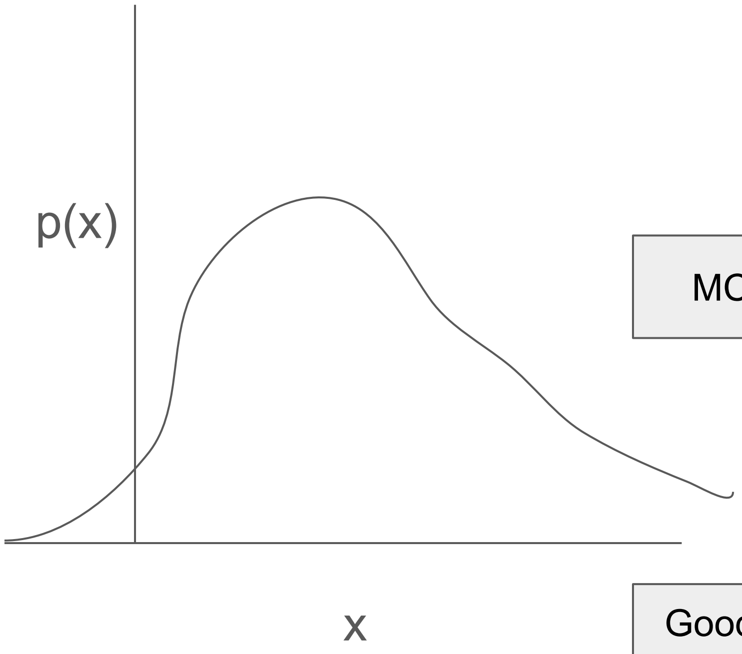
- defines a gaussian
- should explain why the data does not match the mode exactly
- multiplies to derive a total likelihood function

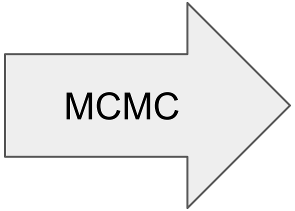
The total likelihood function:

- uniquely defines a posterior probability density function
- can be used to generate samples (MCMC)

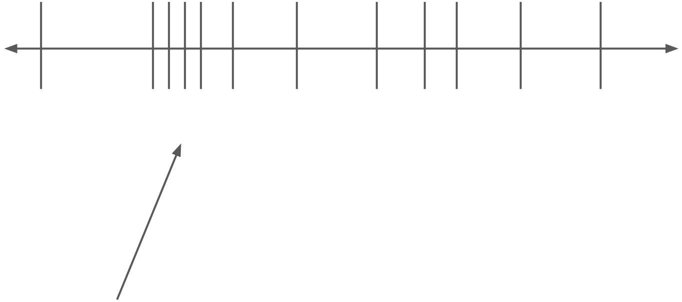
Reminder: What is MCMC?

Start with a probability distribution





Generate samples which represent that distribution



Good MCMC algorithms generate samples which would reproduce the distribution as a histogram

DVCS total cross section parameters (twist 2)

σ_{DVCS} = f { x, Q, t, Eb, ReH, ReE, ReEt, ReHt, ImH, ImEt, ImE, ImHt}

σ_{BH} = f { x, Q, t, phi, Eb}

σ_{INT} = f { x, Q, t, phi, Eb, ReH, ReE, ReHt }

Kinematic Setup

CFFs

Total Cross Section

$$\sigma_{TOT} = \sigma_{DVCS} + \sigma_{BH} + \sigma_{INT}$$

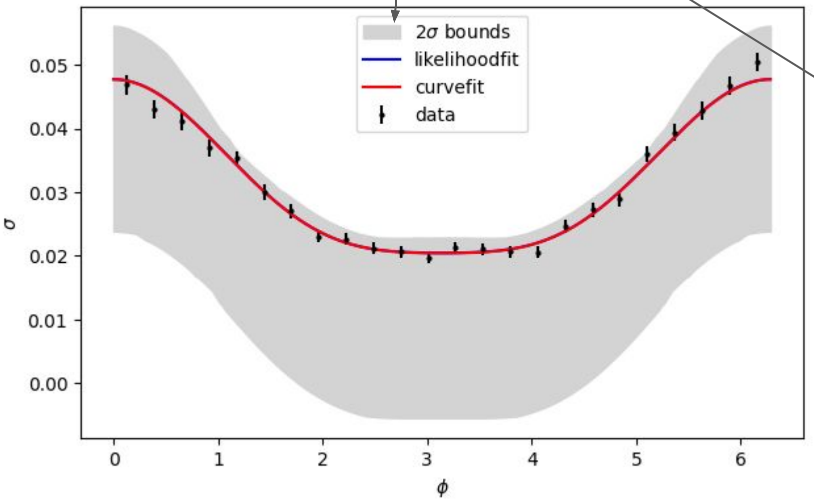
Naive MCMC 1

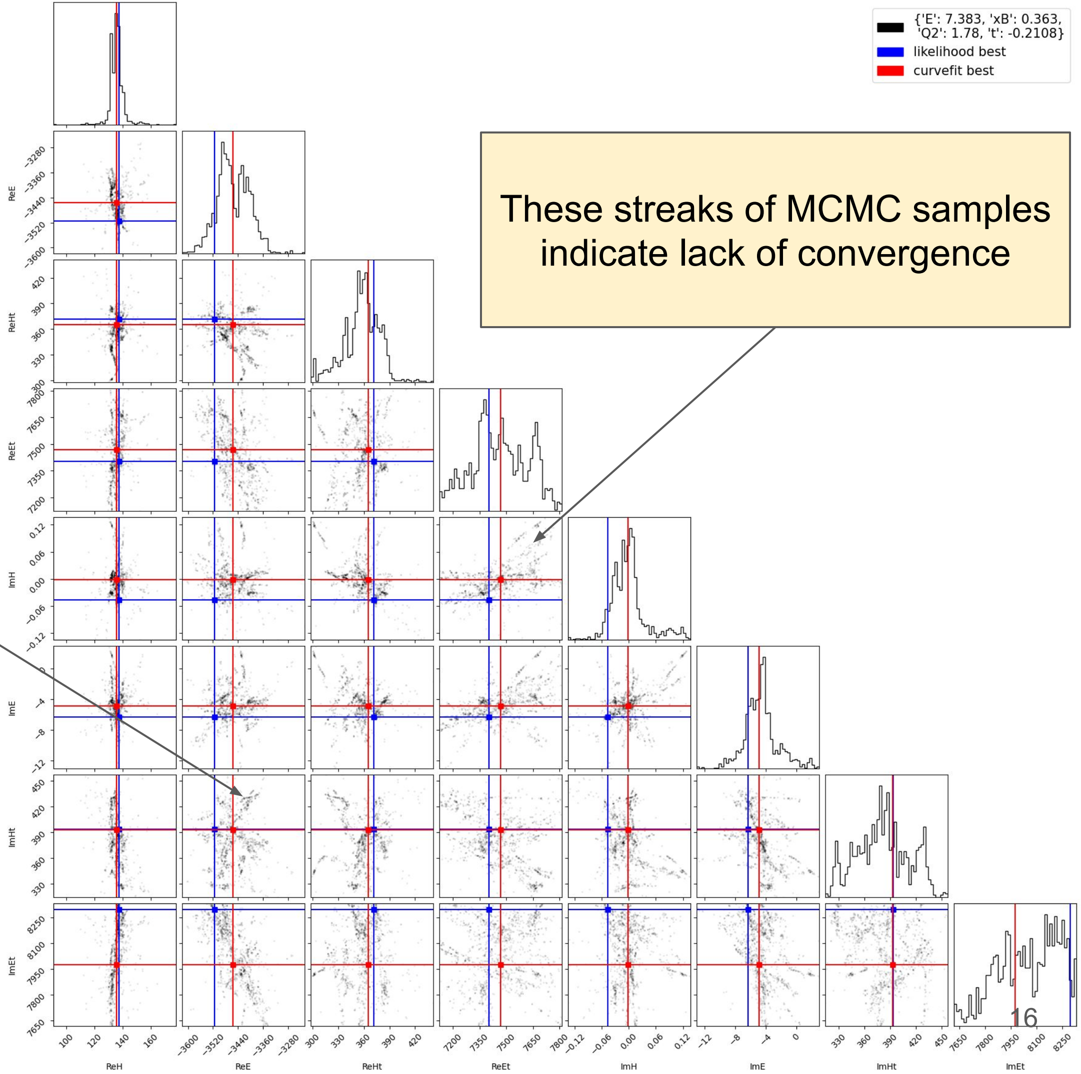
Fitting $\sigma_{TOT}(\phi_A)$ directly with all 8 CFFs

Provides a highly degenerate result

(nonsense bounds)

```
Kinematics:{'E': 7.383, 'xB': 0.363, 'Q2': 1.78, 't': -0.2108}
  CFFs Free  : [ReH, ReE, ReHt, ReEt, ImH, ImE, ImHt, ImEt]
  CFFs Fixed : []
```





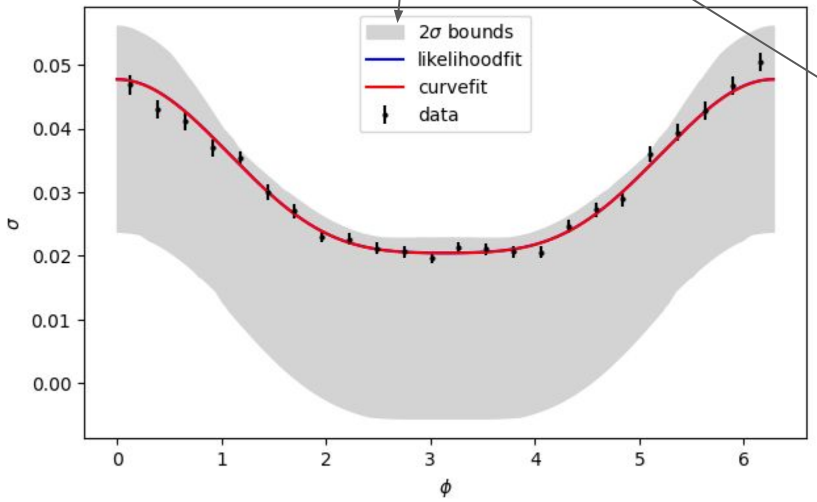
These streaks of MCMC samples indicate lack of convergence

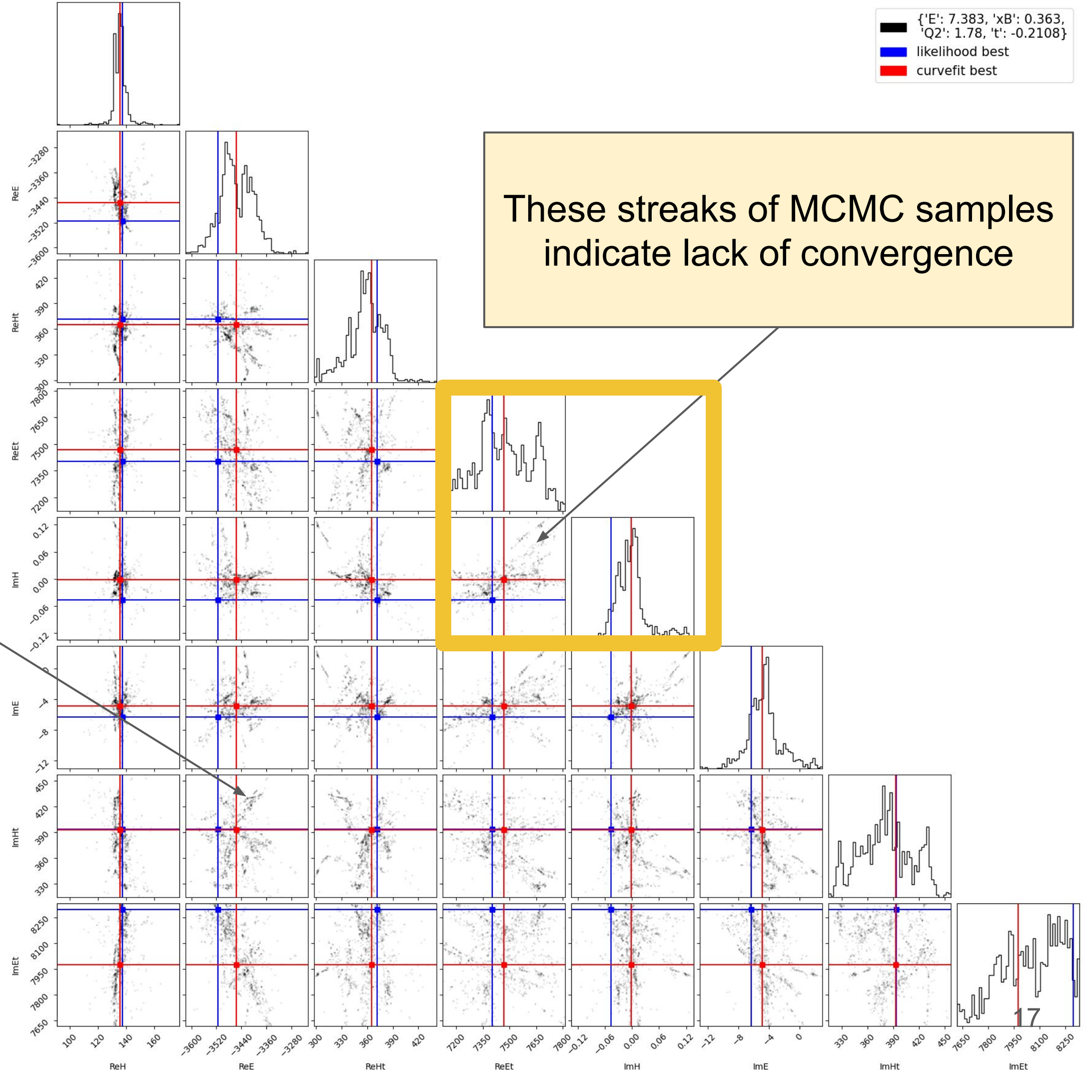
Naive MCMC 2

Fitting $\sigma_{TOT}(\phi_A)$ directly with all 8 CFFs

Provides a highly degenerate result (nonsense bounds)

```
Kinematics:{'E': 7.383, 'xB': 0.363, 'Q2': 1.78, 't': -0.2108}
  CFFs Free  : [ReH, ReE, ReHt, ReEt, ImH, ImE, ImHt, ImEt]
  CFFs Fixed : []
```

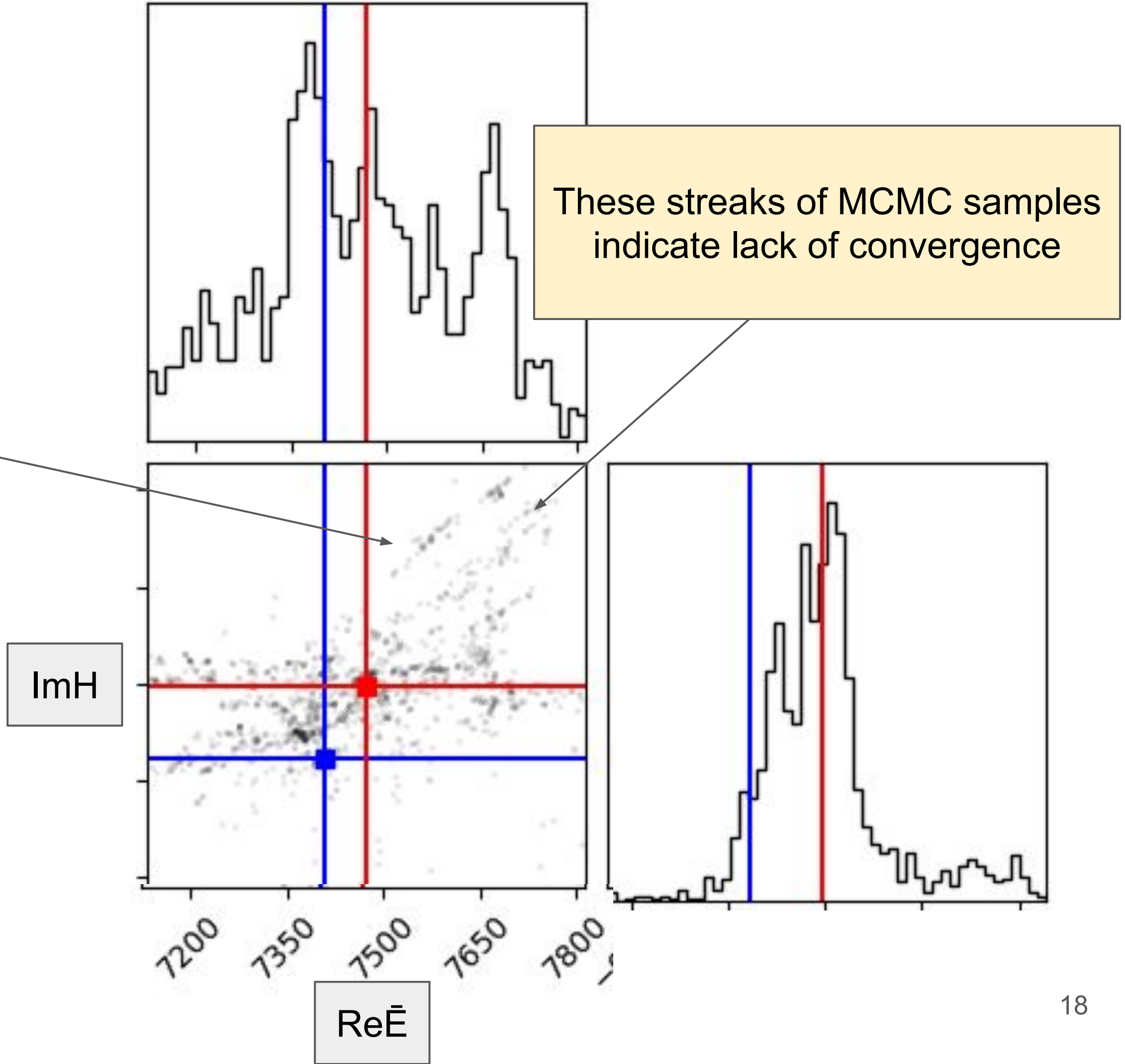




These streaks of MCMC samples indicate lack of convergence

Naive MCMC 3

Fitting $\sigma_{TOT}(\phi_A)$ directly with all 8 CFFs
Provides a highly degenerate result
(nonsense bounds)



Cross Section Dependence with constant σ_{DVCS}
(Gia-Wei & Simonetta's idea)

const

σ_{DVCS} = ~~f {ReE, ReEt, ReHt, x, Q, ImH, ImEt, t, ReH, Eb, ImE, ImHt}~~

No ϕ dependence

σ_{BH} = f {x, Q, t, phi, Eb}

σ_{INT} = f {x, Q, t, phi, Eb, ReH, ReE, ReHt}

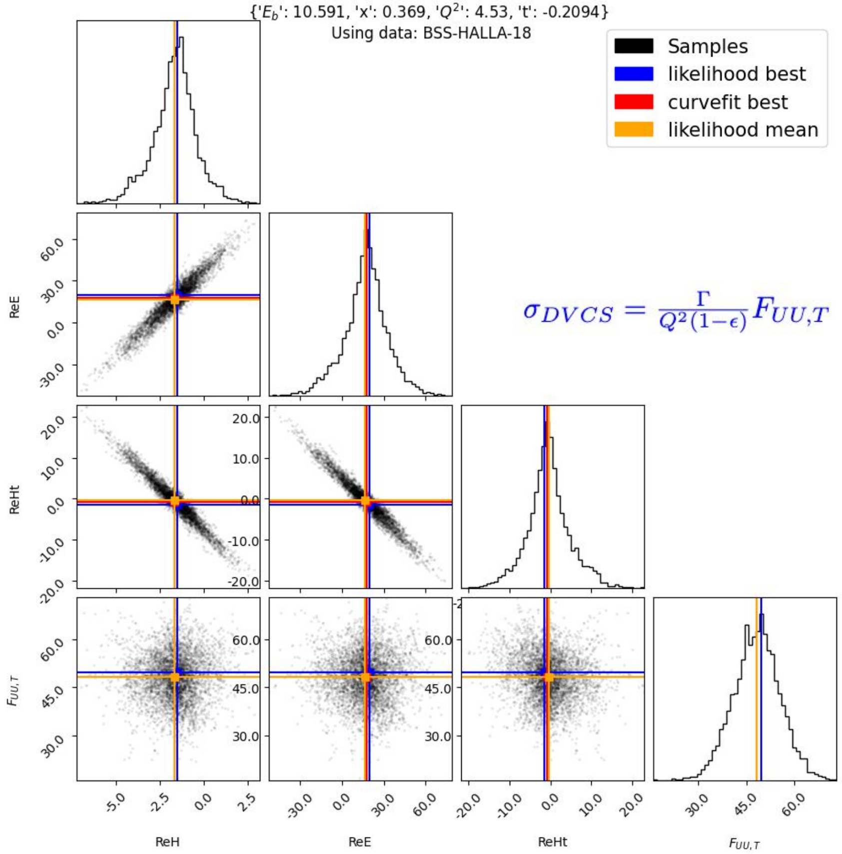
We can fully constrain these 3 CFFs and treat σ_{DVCS} as a constant

Total Cross Section:

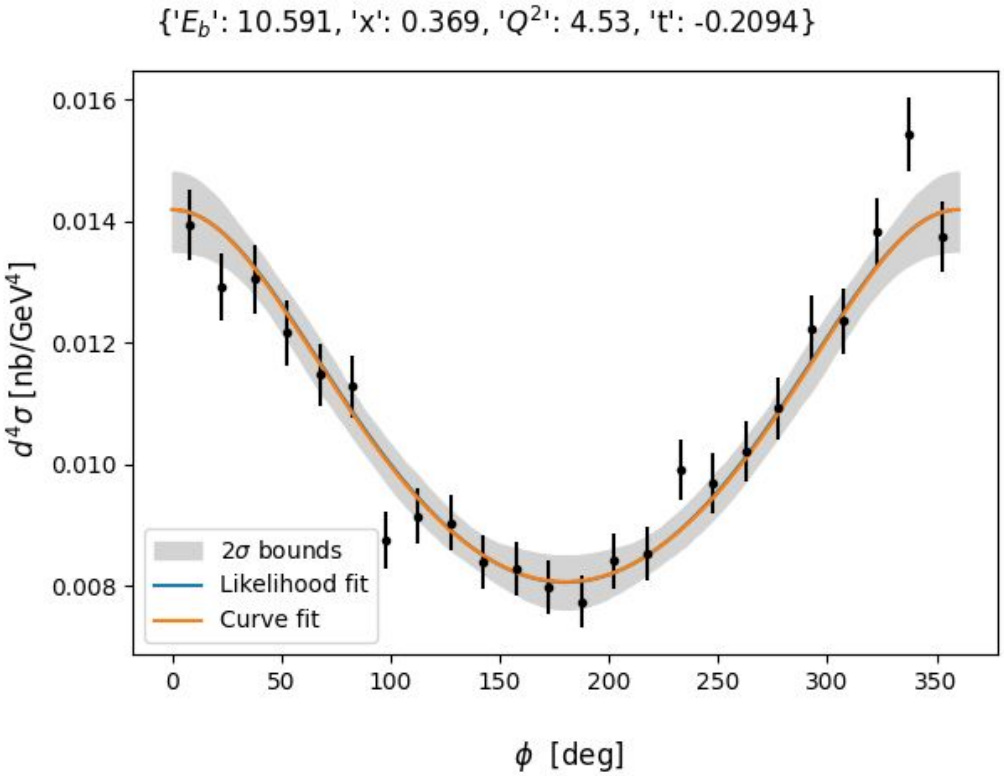
$$\sigma_{TOT} = \sigma_{DVCS} + \sigma_{BH} + \sigma_{INT}$$

Here the maximum likelihood is achieved allowing 3CFFs and σ_{DVCS} to vary as a constant.
Only single angles are used

Canonical Likelihood Result with parameter σ_{DVCS}

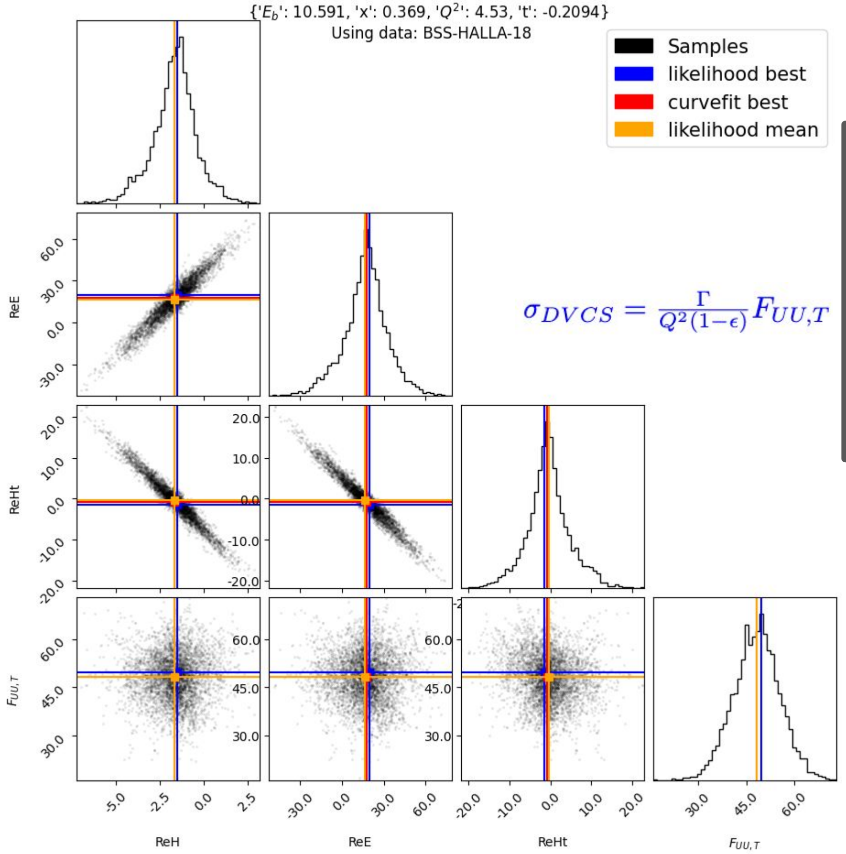


$$\sigma_{DVCS} = \frac{\Gamma}{Q^2(1-\epsilon)} F_{UU,T}$$

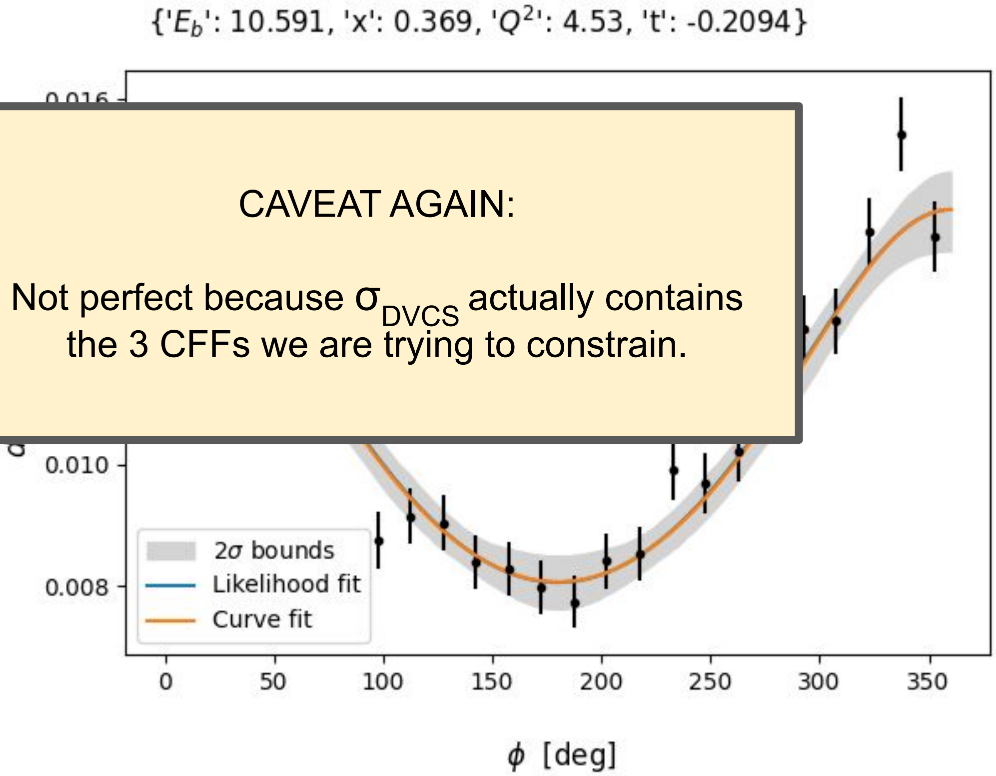


Here the maximum likelihood is achieved allowing 3CFFs and σ_{DVCS} to vary as a constant.
Only single angles are used

Canonical Likelihood Result with parameter σ_{DVCS}



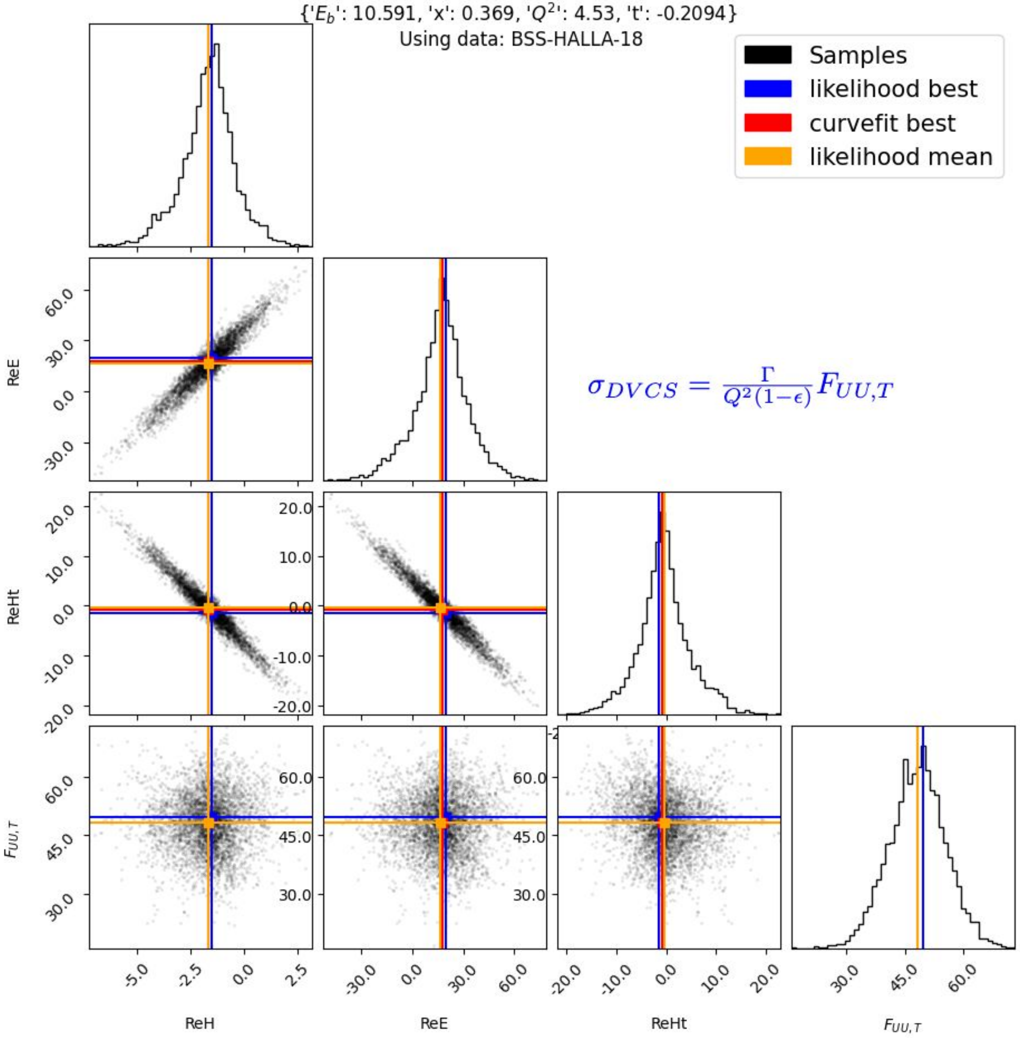
$$\sigma_{DVCS} = \frac{\Gamma}{Q^2(1-\epsilon)} F_{UU,T}$$



CAVEAT AGAIN:

Not perfect because σ_{DVCS} actually contains the 3 CFFs we are trying to constrain.

Canonical Likelihood corner with parameter σ_{DVCS} max size plot



Cross Section Dependence & Difference Method
(Gia-Wei & Simonetta's idea)

σ_{DVCS} = f {ReE, ReEt, ReHt, x, Q, ImH, ImEt, t, ReH, Eb, ImE, ImHt}

No ϕ dependence

σ_{BH} = f {x, Q, t, phi, Eb}

σ_{INT} = f {x, Q, t, phi, Eb, ReH, ReE, ReHt}

We can fully constrain these 3 CFFs

Total Cross Section = $\sigma_{TOT} = \sigma_{DVCS} + \sigma_{BH} + \sigma_{INT}$

Measured

No CFF dependence

At fixed [x, Q, t, Eb] we can perform the following subtraction:
$\sigma_{TOT}(\phi_A) - \sigma_{TOT}(\phi_B) = [\sigma_{BH}(\phi_A) + \sigma_{INT}(\phi_B) + \sigma_{DVCS}] - [\sigma_{BH}(\phi_A) + \sigma_{INT}(\phi_B) + \sigma_{DVCS}]$
$[\sigma_{TOT}(\phi_A) - \sigma_{BH}(\phi_A)] - [\sigma_{TOT}(\phi_B) - \sigma_{BH}(\phi_A)] = \sigma_{INT}(\phi_B, 3CFFs) - \sigma_{INT}(\phi_B, 3CFFs)$

Defining a 3CFF Likelihood with 1 kinematic Bin

Eb	x	Q	t
4.487	0.483	1.646208	-0.3906

	phi	XUU	errstat	XBH	XUU-XBH
384	0.130900	0.02549	0.0035	0.030466	-0.004976
385	0.392699	0.02750	0.0032	0.028554	-0.001054
386	0.654498	0.02570	0.0028	0.025363	0.000337
387	0.916298	0.02224	0.0025	0.021723	0.000517
388	1.178097	0.02092	0.0022	0.018266	0.002654
389	1.439897	0.02043	0.0021	0.015309	0.005121
390	1.701696	0.01553	0.0020	0.012941	0.002589
391	1.963495	0.01764	0.0020	0.011128	0.006512
392	2.225295	0.01629	0.0018	0.009797	0.006493
393	2.487094	0.01423	0.0017	0.008871	0.005359
394	2.748894	0.01430	0.0018	0.008288	0.006012
395	3.010693	0.01194	0.0017	0.008007	0.003933
396	3.272492	0.01639	0.0018	0.008007	0.008383
397	3.534292	0.01862	0.0019	0.008288	0.010332
398	3.796091	0.01804	0.0019	0.008871	0.009169
399	4.057891	0.01375	0.0019	0.009797	0.003953
400	4.319690	0.01540	0.0020	0.011128	0.004272
401	4.581489	0.02347	0.0022	0.012941	0.010529
402	4.843289	0.02409	0.0024	0.015309	0.008781
403	5.105088	0.02080	0.0024	0.018266	0.002534
404	5.366887	0.02711	0.0026	0.021723	[illegible]
405	5.628687	0.02834	0.0029	0.025363	0.002977
406	5.890486	0.02480	0.0031	0.028554	-0.003754
407	6.152286	0.02869	0.0034	0.030466	-0.001776

Plug each combination of two rows
into the interference term difference equation

$$\sigma_{TOT}(\phi_A) - \sigma_{BH}(\phi_A) - \sigma_{TOT}(\phi_B) + \sigma_{BH}(\phi_A)$$

$$= \sigma_{INT}(\phi_B, 3CFFs) - \sigma_{INT}(\phi_B, 3CFFs)$$

Without measurement error: the left side and the right side of the equation should equal for every combination of rows.

Instead they will not equal, but difference should be described by measurement error

$\sigma_{TOT}(\phi_A)$, $\sigma_{TOT}(\phi_B)$ are measured with error

Likelihood assuming Dependent Difference Gaussians 1

Not all differences of two cross section measurements are independent.

There is redundant information if we include every difference of 2 angles.

The errors we have are still gaussian, so we can use results for adding the dependent random variables involved

$$X + Y \sim N(\mu_X + \mu_Y,\ \sigma_X^2 + \sigma_Y^2 + 2\sigma_{X,Y})$$
$$aX + bY \sim N(a\mu_X + b\mu_Y,\ a^2\sigma_X^2 + b^2\sigma_Y^2 + 2ab\sigma_{X,Y})$$

https://stats.stackexchange.com/questions/19948/what-is-the-distribution-of-the-sum-of-non-i-i-d-gaussian-variates

Likelihood assuming Dependent Difference Gaussians 2

$$\mathcal{L} = \text{Gaussian}(x = \Delta\sigma_{obs,A,B}, \mu = [\Delta\sigma_{model,A,B} \forall A, B], cov = cov(\Delta\sigma_{obs,A_i,B_i}, \Delta\sigma_{obs,A_j,B_j}))$$

$$\text{Gaussian}(x, \mu, \Sigma) = \frac{1}{2\pi|\Sigma|} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu) \right)$$

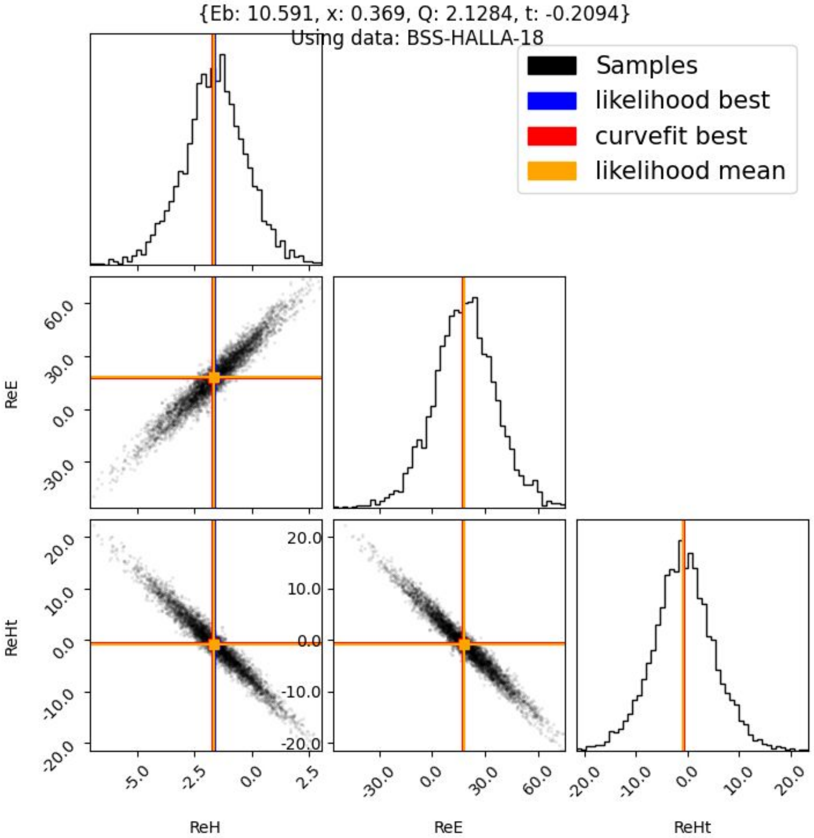
$$\begin{aligned} cov(\Delta\sigma_{obs,A_i,B_i}, \Delta\sigma_{obs,A_j,B_j}) &= cov(\sigma_{obs,A_i} - \sigma_{obs,B_i}, \sigma_{obs,A_j} - \sigma_{obs,B_j}) \\ &= cov(\sigma_{obs,A_i}, \sigma_{obs,A_j}) - cov(\sigma_{obs,A_i}, \sigma_{obs,B_j}) \\ &\quad - cov(\sigma_{obs,B_i}, \sigma_{obs,A_j}) + cov(\sigma_{obs,B_i}, \sigma_{obs,B_j}) \end{aligned}$$

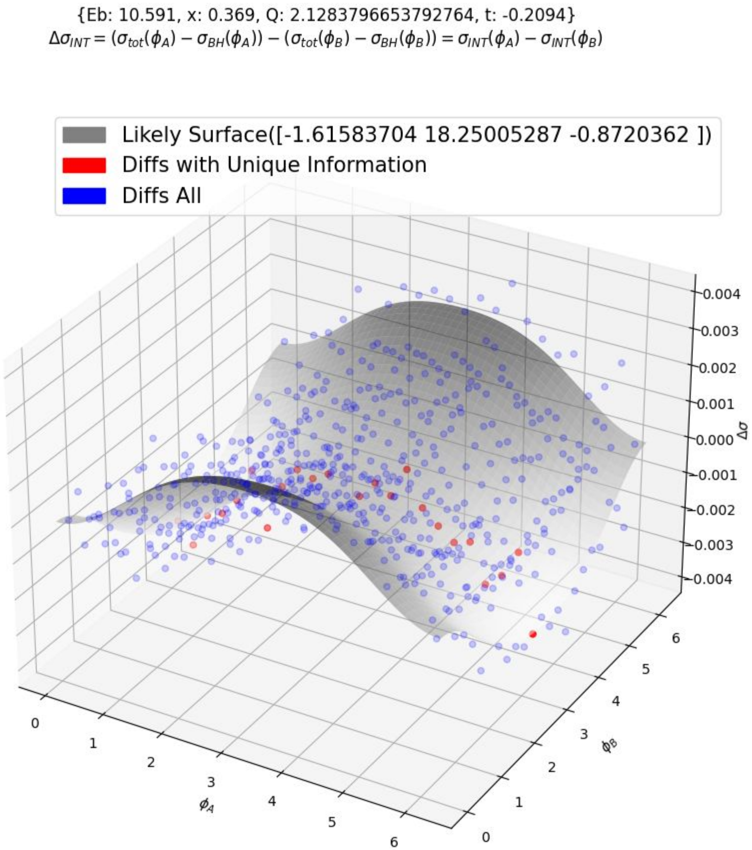
$$x_{\Delta\sigma_{obs,A,B}} = \sigma_{obs,A} - \sigma_{obs,B} \qquad \mu_{\Delta\sigma_{model,A,B}} = \sigma_{model,A} - \sigma_{model,B}$$

$$\mathbf{x} = \vec{x} = \begin{bmatrix} \sigma_{obs,\phi 0} - \sigma_{obs,\phi 23} \\ \sigma_{obs,\phi 1} - \sigma_{obs,\phi 23} \\ \dots \\ \sigma_{obs,\phi 22} - \sigma_{obs,\phi 23} \end{bmatrix} \qquad \mu = \vec{\mu} = \begin{bmatrix} \sigma_{model,\phi 0} - \sigma_{model,\phi 23} \\ \sigma_{model,\phi 1} - \sigma_{model,\phi 23} \\ \dots \\ \sigma_{model,\phi 22} - \sigma_{model,\phi 23} \end{bmatrix}$$

Here the maximum likelihood is achieved allowing 3CFFs to vary.
Only 23 combinations of 2 angles are used.

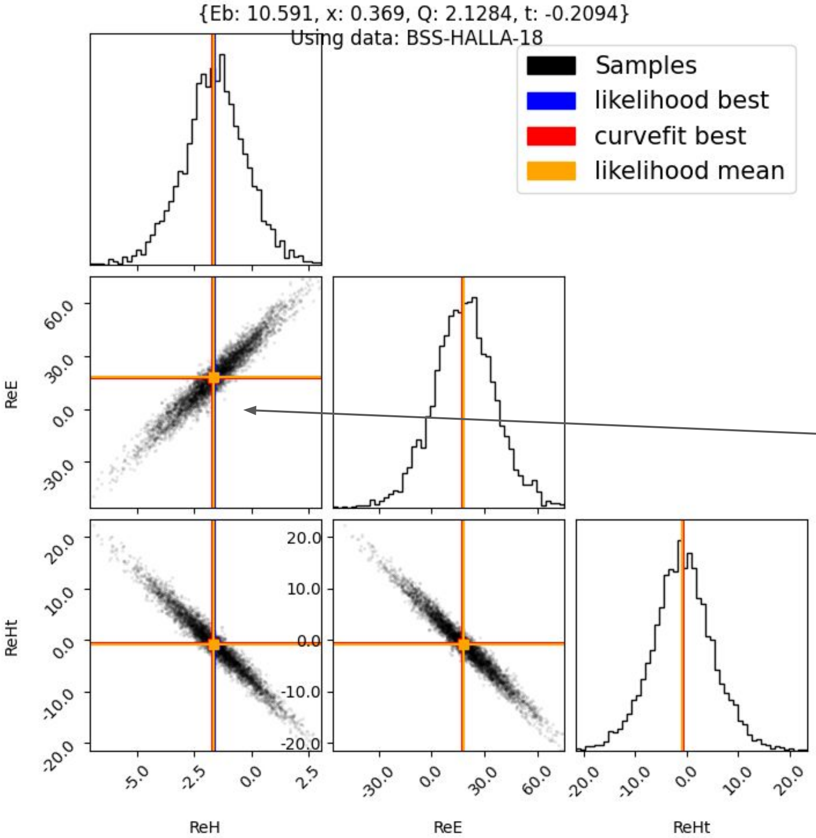
Interference Difference Likelihood Result

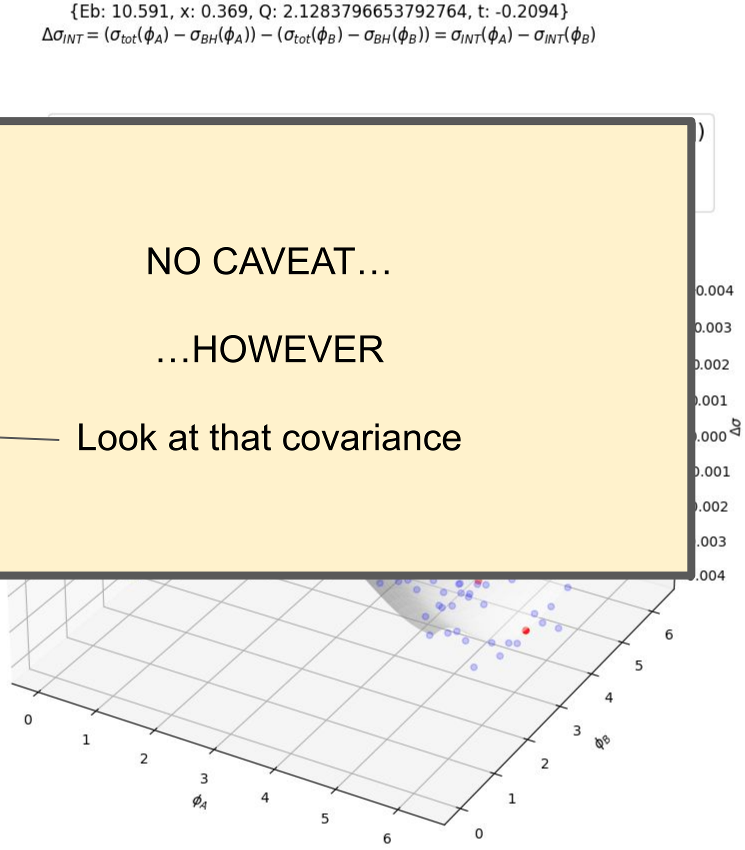




Here the maximum likelihood is achieved allowing 3CFFs to vary.
Only 23 combinations of 2 angles are used.

Interference Difference Likelihood Result



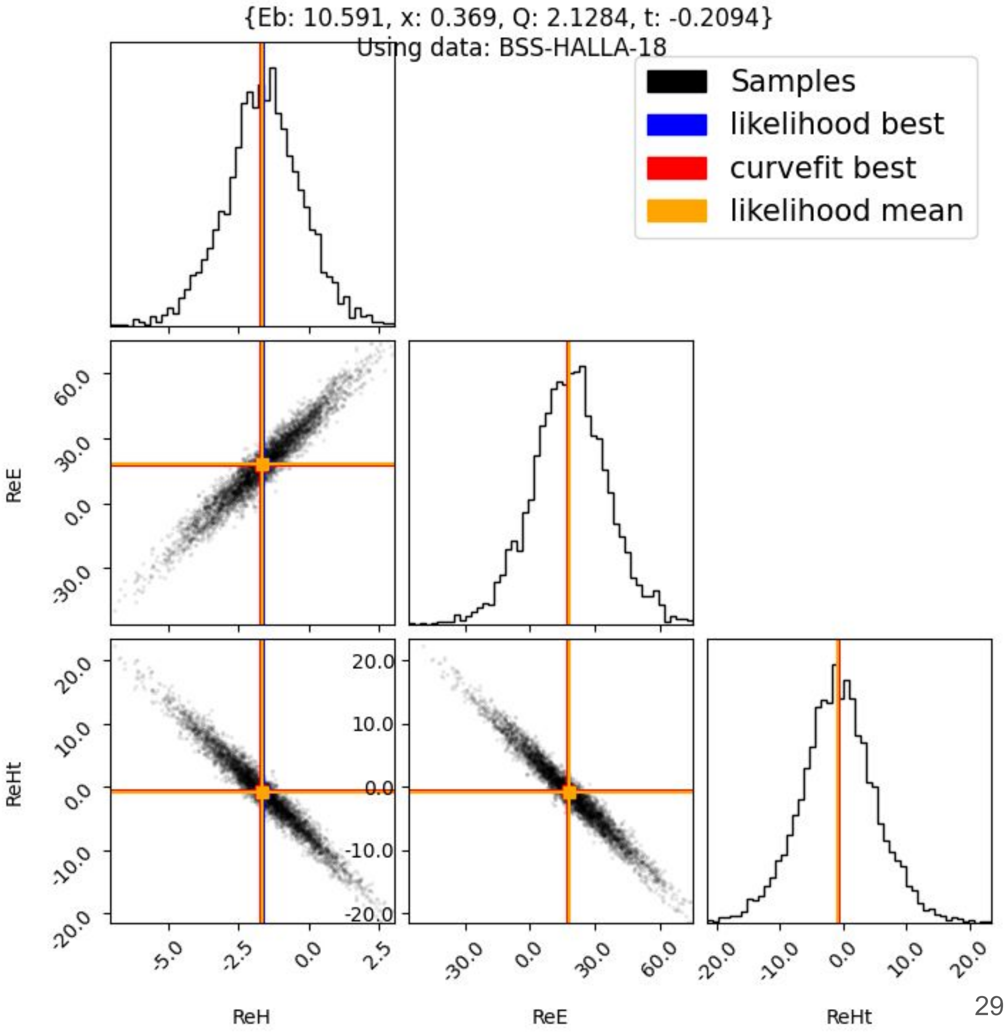


NO CAVEAT…

…HOWEVER

Look at that covariance

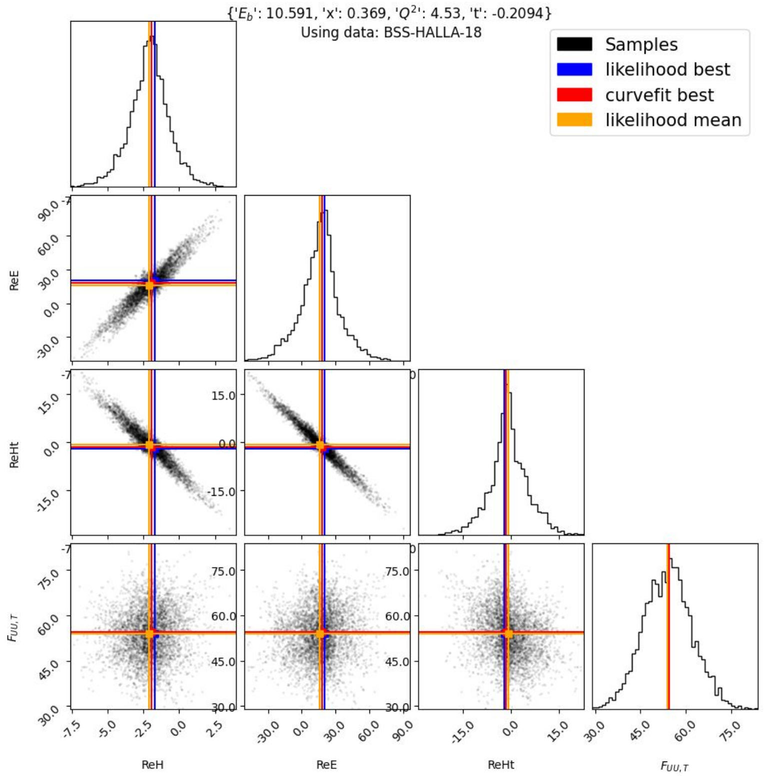
Difference Likelihood corner max size plot

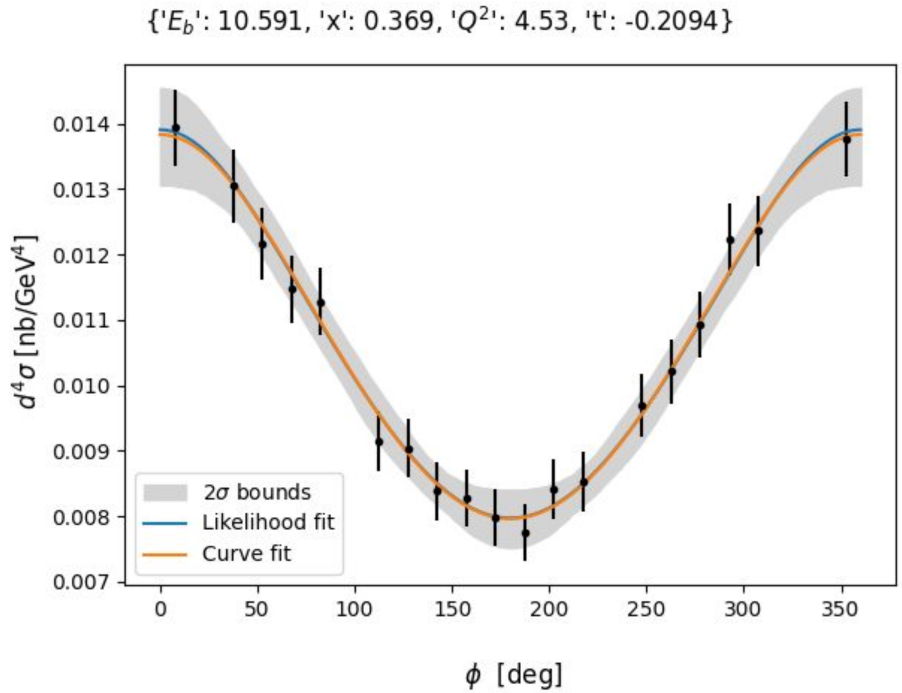


Ideas to try and fix the covariance:

1) Try fake data with smaller error bars (yes this helps a little)
2) Try removing outliers (cherry picking shows yes this helps a little)

Cherry Pick Data: (admittedly not the best practice…)





Conclusion & Caveats

- Using the UVA DVCS twist 2 unpolarized cross section model ($\sigma_{TOT,\ UVA,\ UU}$)
    - Assume the model is True
    - Assume the model has 8 CFFs only
    - Assume each CFF is independent of phi
- Using Gepard Hall A DVCS Data from Georges thesis
    - doi:10.1038/s41567-019-0774-3
    - each kinematic bin has 24 rows of (ϕ, σ_{TOT}) data
- Naively one would assume we can use the model to produce 24 equations and 8 unknowns to fully constrain the unknowns (as an overdetermined system).
    - However 5 CFFs are degenerate because σ_{DVCS} has no phi dependence.
    - Thus only the other 3 CFFs can be fully constrained using σ_{INT}
- We produced a table of CFF results for 45 kinematic bins (on the archive soon)

END

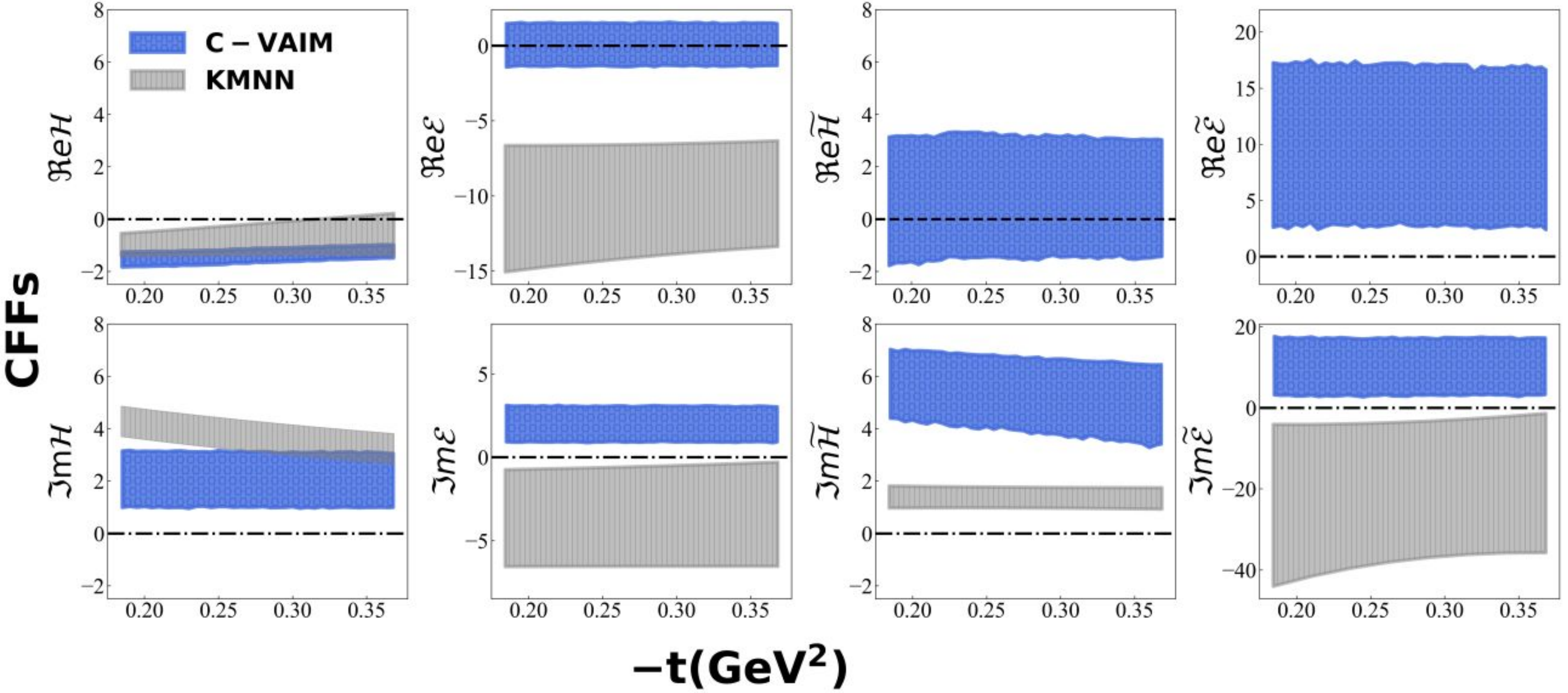


FIG. 6. Prediction of all eight CFFs as a function of t for a fixed kinematics: $x_{Bj} = 0.35$ and $Q^2 = 1.9$ GeV^2, and initial electron energy of 6 GeV [32] Results are compared with the NN based extraction of Ref. [31]. Predictions are generated according to the training prior defined in Sec. III E.

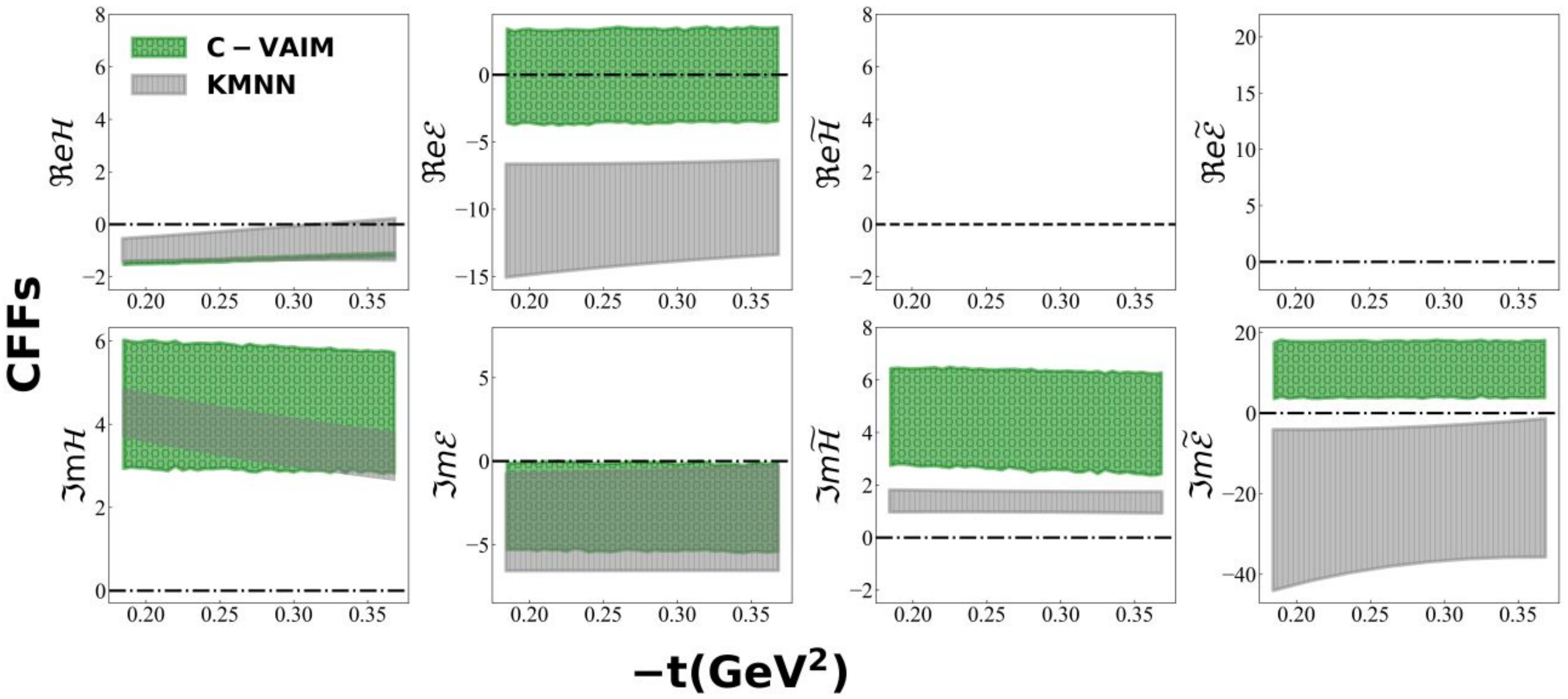


FIG. 7. Prediction of six CFFs as a function of t for a fixed kinematics $x_{Bj} = 0.35$ and $Q^2 = 1.9\ \mathrm{GeV}^2$ for 6 GeV initial electron energy [32] fixing $\Re e\widetilde{H}$ and $\Re e\widetilde{E} = 0$. The choice of the six CFFs matches the one in the analysis of Ref. [31], therefore allowing for a more consistent comparison of results. Predictions are generated according to the training prior defined in Sec. III E.

Appendix: Unique Kinematic Bins

	Eb	x	Q	t
0	4.487	0.483	1.646208	-0.3906
1	4.487	0.483	1.646208	-0.3481
2	4.487	0.484	1.646208	-0.4350
3	4.487	0.485	1.646208	-0.4797
4	4.487	0.485	1.649242	-0.5399
5	7.383	0.363	1.780449	-0.2966
6	7.383	0.363	1.780449	-0.2108
7	7.383	0.364	1.783255	-0.5858
8	7.383	0.365	1.783255	-0.4709
9	7.383	0.365	1.783255	-0.3849
10	8.521	0.367	1.910497	-0.2658
11	8.521	0.367	1.910497	-0.2046
12	8.521	0.369	1.915724	-0.3303
13	8.521	0.370	1.918333	-0.4805
14	8.521	0.370	1.918333	-0.3925
15	8.521	0.610	2.366432	-0.7645
16	8.521	0.612	2.370654	-0.9042
17	8.521	0.615	2.374868	-1.0482
18	8.521	0.616	2.379075	-1.3732
19	8.521	0.617	2.376973	-1.1896
20	8.847	0.482	2.308679	-0.3854
21	8.847	0.483	2.310844	-0.4456
22	8.847	0.485	2.315167	-0.5085
23	8.847	0.486	2.317326	-0.5696
24	8.847	0.486	2.319483	-0.6547
25	8.851	0.497	2.121320	-0.4098
26	8.851	0.501	2.128380	-0.4796
27	8.851	0.504	2.135416	-0.5475
28	8.851	0.506	2.137756	-0.6138
29	8.851	0.508	2.142429	-0.7073
30	10.591	0.369	2.128380	-0.2094
31	10.591	0.370	2.133073	-0.2719
32	10.591	0.371	2.137756	-0.4834
33	10.591	0.372	2.137756	-0.3356
34	10.591	0.373	2.140093	-0.3988
35	10.591	0.608	2.905168	-0.7911
36	10.591	0.609	2.906888	-0.9120
37	10.591	0.611	2.912044	-1.0371
38	10.591	0.613	2.915476	-1.1631
39	10.591	0.613	2.917190	-1.3282
40	10.992	0.494	2.653300	-0.4325
41	10.992	0.498	2.662705	-0.5256
42	10.992	0.498	2.664583	-0.8556
43	10.992	0.499	2.666458	-0.6979
44	10.992	0.499	2.668333	-0.6129

Hall A Data DVCS

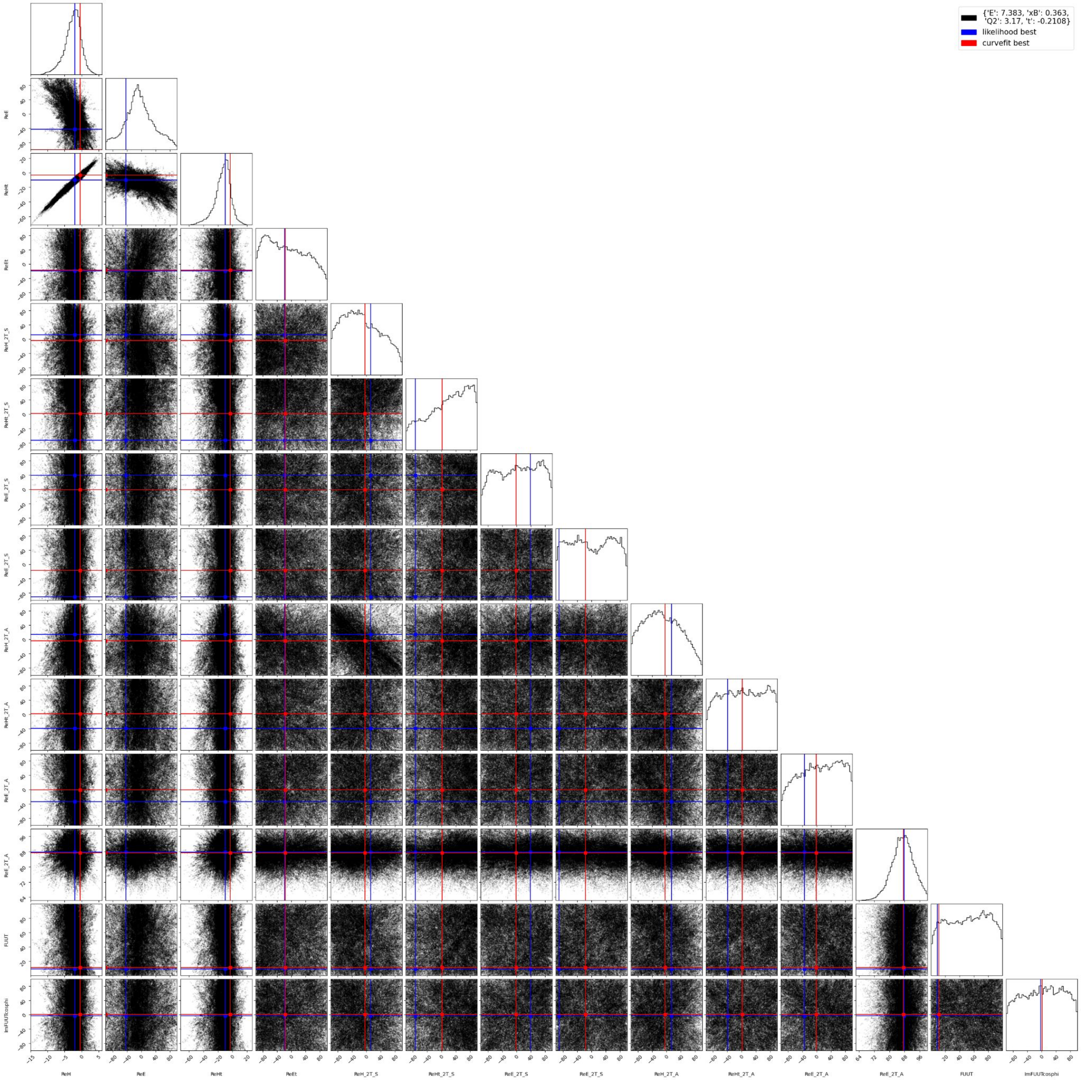
https://github.com/kkumer/gepard/blob/master/src/gepard/datasets/ep2epgamma/BSS-HALLA-18.dat
doi:10.1038/s41567-019-0774-3

```
#E      xB      Q2        t          phi    XUU      errstat
###############################################################################
7.383   0.363   3.17      -0.2108    7.5    0.04685 0.0015
7.383   0.363   3.17      -0.2108    22.5   0.043   0.0015
7.383   0.363   3.17      -0.2108    37.5   0.0412  0.0014
7.383   0.363   3.17      -0.2108    52.5   0.03692 0.0013
7.383   0.363   3.17      -0.2108    67.5   0.03529 0.0012
7.383   0.363   3.17      -0.2108    82.5   0.02995 0.0012
7.383   0.363   3.17      -0.2108    97.5   0.02703 0.0011
7.383   0.363   3.17      -0.2108    112.5  0.02304 0.00097
....
```

Twist 3 Plot

There is degeneracy

we need to investigate



Symbols Scratch

σ_{DVCS}
σ_{INT}
σ_{BH}
σ_{TOT}
ϕ_{A}
ϕ_{B}