

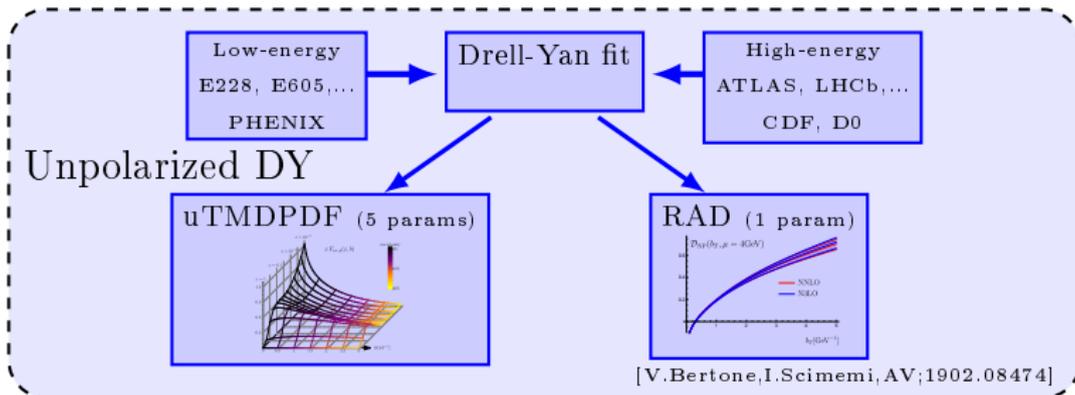
# Extraction of TMD distributions from data

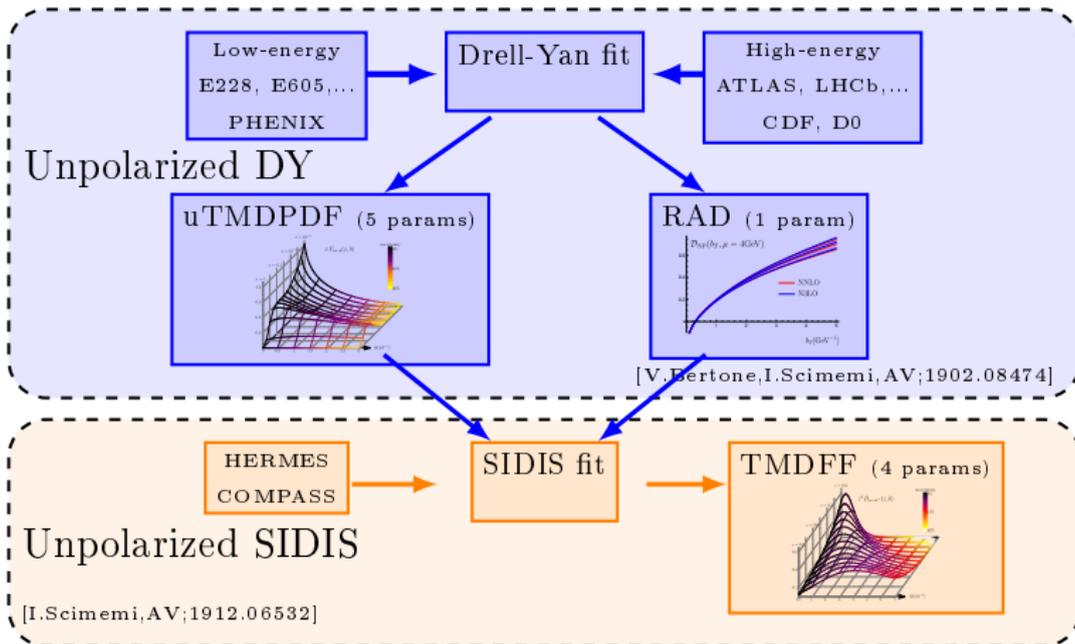
**Alexey Vladimirov**  
(Regensburg University)

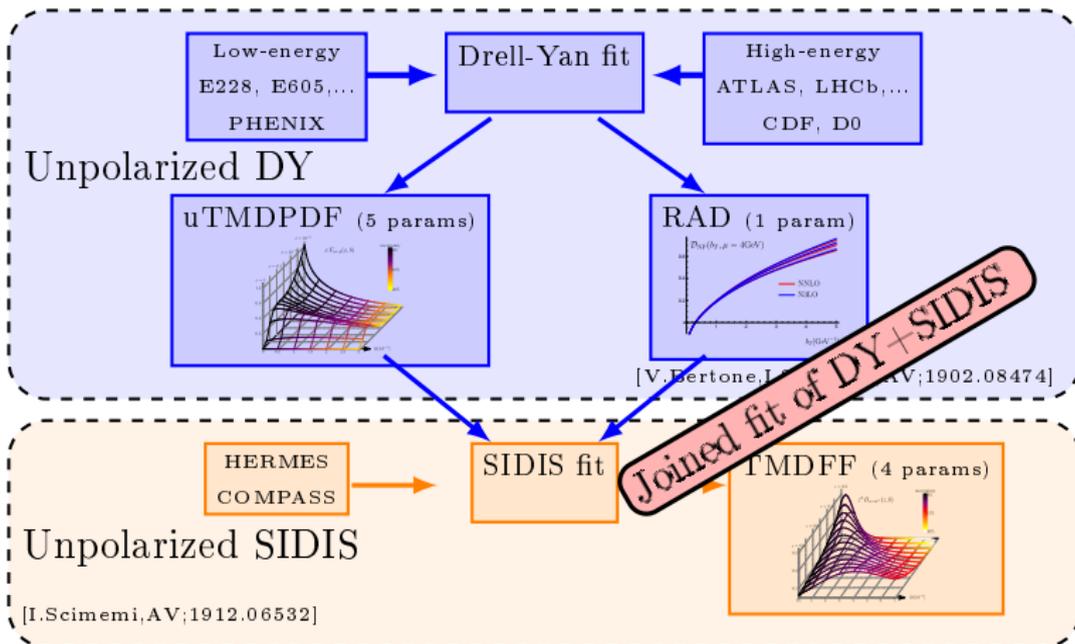
Workshop:  
**TMD Studies: from JLab to EIC**  
May 7, 2021

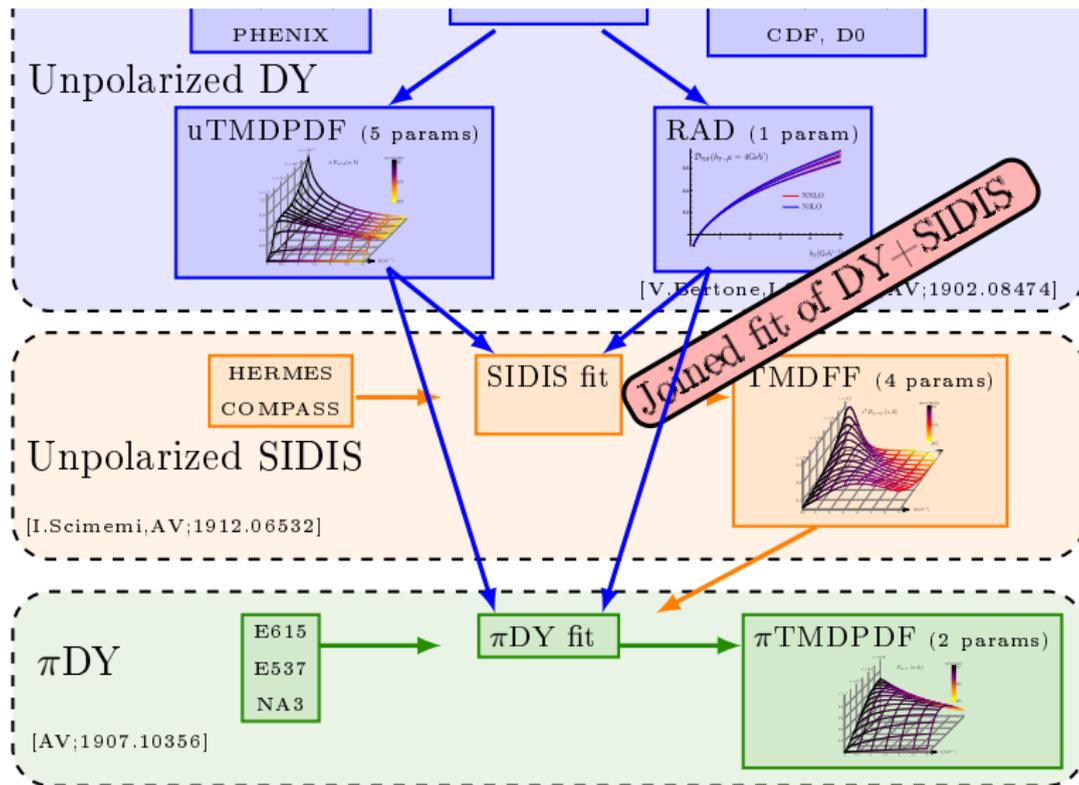
## Extraction of TMD with artemide

- ▶ Main features
  - ▶ Position space
  - ▶  $\zeta$ -prescription
  - ▶ (maximum) Perturbative matching
  - ▶ Strict data cuts
- ▶ Code
  - ▶ **artemide**
    - ▶ Fortran 95
    - ▶ <https://github.com/VladimirovAlexey/artemide-public>
  - ▶ **artemide-DataProcessor**
    - ▶ Python
    - ▶ <https://github.com/VladimirovAlexey/artemide-DataProcessor>

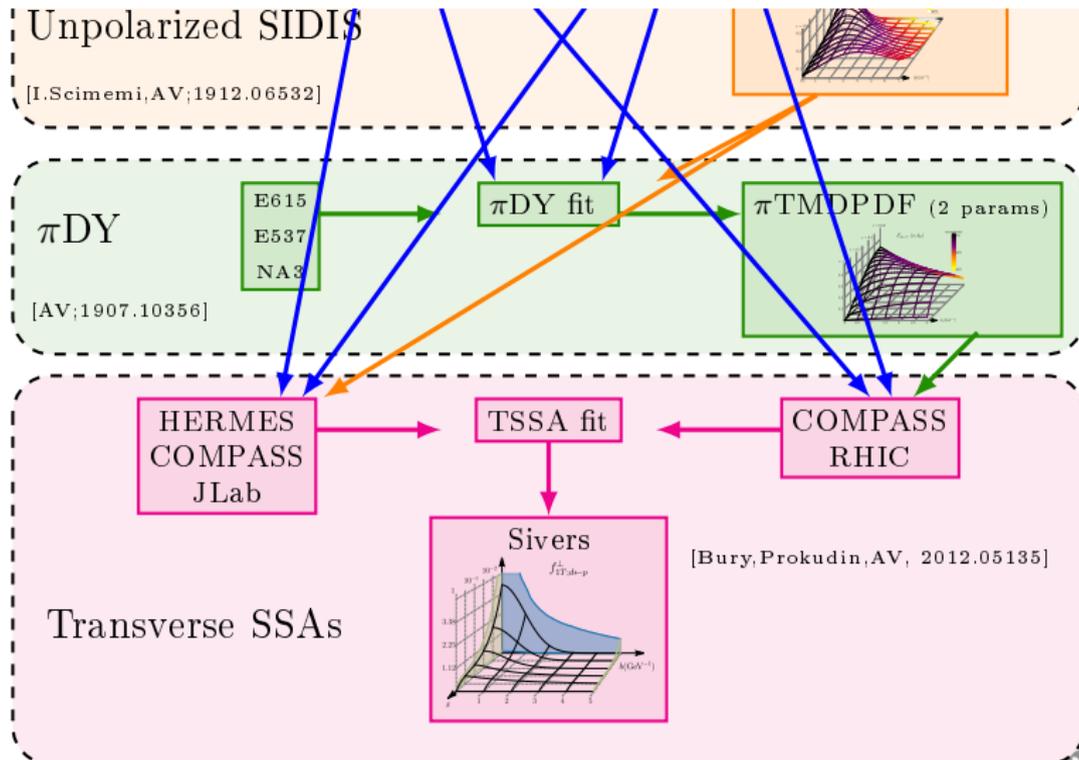




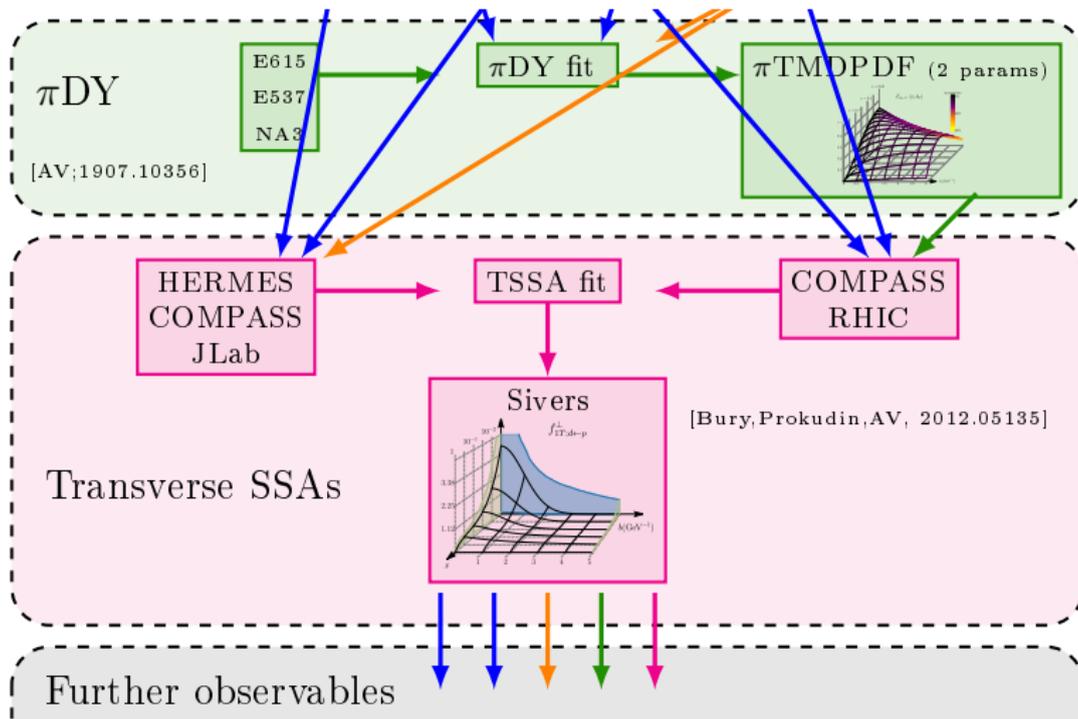




# Universality & the chain of extractions



# Universality & the chain of extractions



In my talk I am going to focus on particular aspects of  
TMD phenomenology

## Data selection for TMD factorization

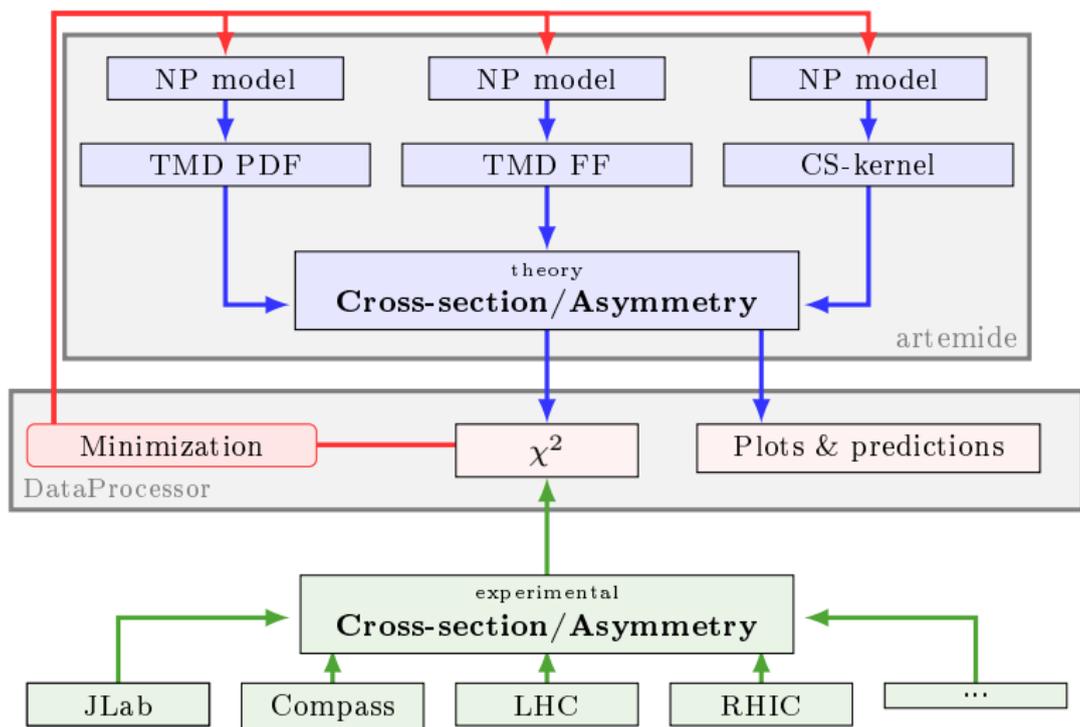
- ▶ Theoretical justification
- ▶ Experimental evidence
- ▶ Impact on present and future data-sets

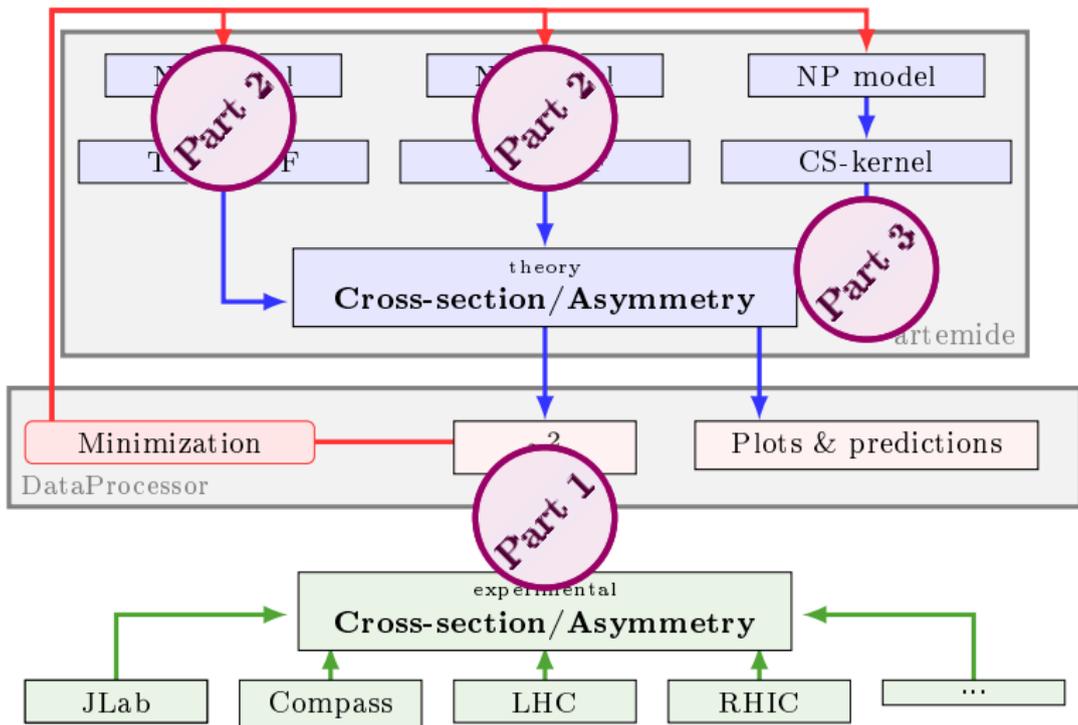
## PDF-bias and problem of TMD modeling

- ▶ Structure of (modern) TMD models
- ▶ PDF-bias
- ▶ Flavor dependence (**work in progress**)

## Problem with determination of CS-kernel

- ▶ Correlation between TMDs and CS-kernel
- ▶  $\zeta$ -prescription





# Part I

## Data cuts

TMD factorization is a systematic expansion of **hadronic tensor** in power of  $\frac{q_T}{Q}$

$$W^{\mu\nu} = \underbrace{W_0^{\mu\nu}}_{F \otimes D} + \overbrace{\frac{q_T}{Q} W_1^{\mu\nu}}^{\text{not present in } F_{UU}, F_{UT}^{\text{sin}}} + \underbrace{\frac{q_T^2}{Q^2} W_2^{\mu\nu}}_{\substack{\partial^2 F \otimes D \\ \partial F \otimes \partial D \\ \Upsilon \otimes \Delta \\ \partial F \otimes \Delta \\ \dots \\ \text{twist-4} \\ \dots}} + \dots$$

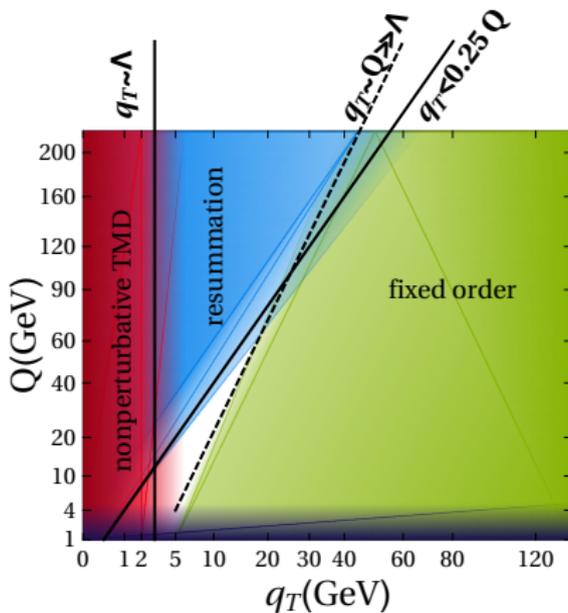
- ▶ Leading-power term is well-investigated
- ▶ Something is known about NLP  $W_1$  (factorization is not proven)
- ▶ **Nothing** is known about NNLP  $W_2$

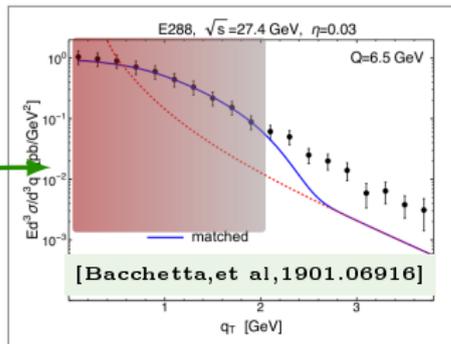
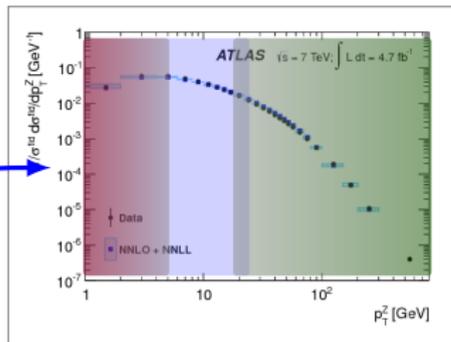
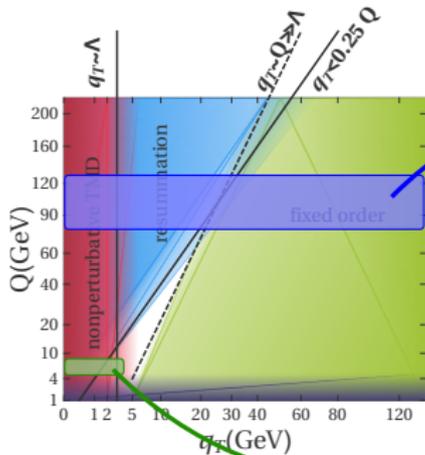
**Important:** TMD factorization is  $Q \rightarrow \infty$ ,  $q_T = \text{fixed}$ .  
 Could be corrections  $\sim \frac{\Lambda}{Q}$  even at  $q_T \rightarrow 0$ .

## Factorization regions

$$q_T \lesssim \delta Q \quad \text{TMD factorization} \quad = \begin{cases} q_T \lesssim \Lambda & \text{nonperturbative regime} \\ q_T \gg \Lambda & \text{"resummation" regime} \end{cases}$$

$$q_T \sim Q \gg \Lambda \quad \text{collinear factorization}$$



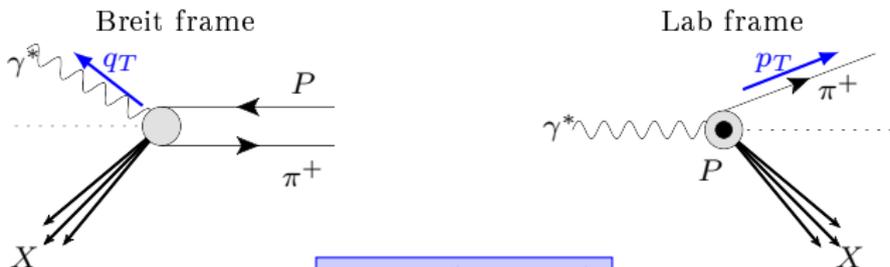


Here I draw  $\delta = 0.25$ , how to justify this choice?

$$\delta^2 = \frac{q_T^2}{Q^2} \sim 0.06 \ll 1$$

## $q_T$ for SIDIS

The factorization for SIDIS is done in the Breit frame



$$q_T^2 = \frac{p_T^2}{z^2} \frac{1 + \gamma^2}{1 - \varsigma^2}$$

$$x_1 = -x \frac{z}{\gamma^2} \left( 1 - \sqrt{1 + \gamma^2 \left( 1 - \frac{q_T^2}{Q^2} \right)} \right), \quad z_1 = z \frac{x_1}{x} \frac{1 + \sqrt{1 - \varsigma^2}}{2 \left( 1 - \frac{q_T^2}{Q^2} \right)}$$

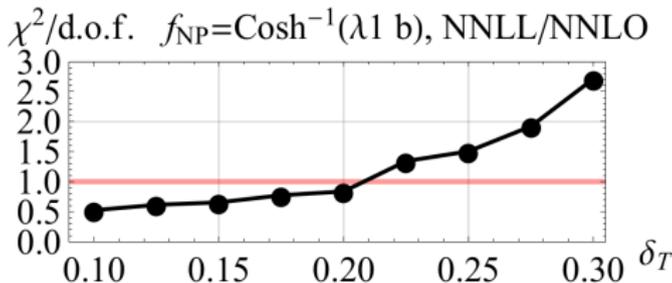
$$\gamma = \frac{2Mx}{Q}$$

$$\varsigma = \gamma \frac{m}{zQ}$$

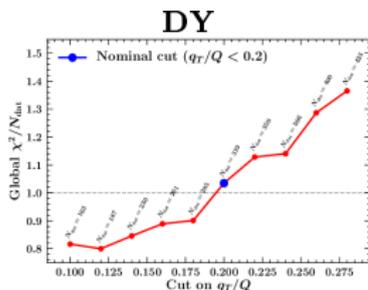
- Only in the Breit frame one can justify the parton model

## Value of $\delta$ from the data

- 1 Fit data at some small  $\delta$  ( $\chi/N_{pt} \sim 1$ )
- 2 Increase  $\delta$  and fit again, starting from the previous minimum (repeat)
- 3 At some moment the  $\chi^2/N_{pt}$  blows up

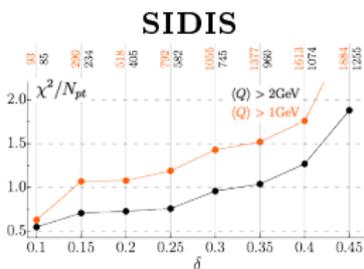


[Scimemi,AV, 1706.01473]



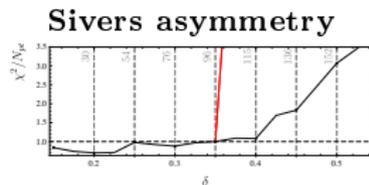
$\delta \sim 0.2$

[Bacchetta,Bertone,etal,1912.07550]



$\delta \sim 0.25$

[Scimemi,AV,1912.06532]



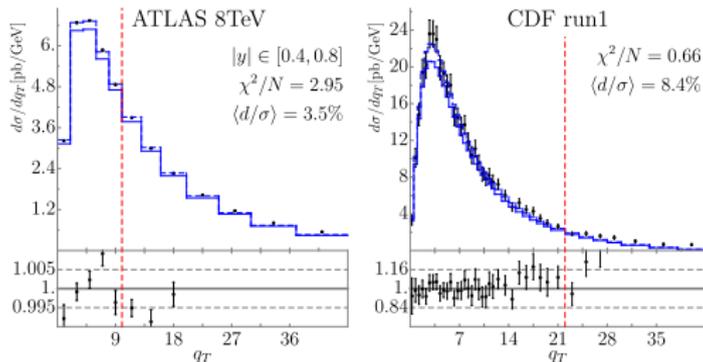
$\delta \sim 0.3$

[Bury,Prokudin,AV,2103.03270]

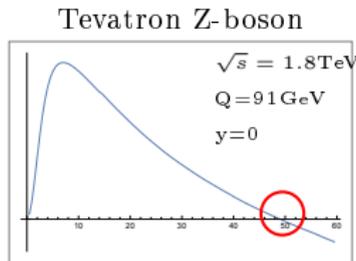
## Data indicates that $\delta \sim 0.2 - 0.25$

### But...

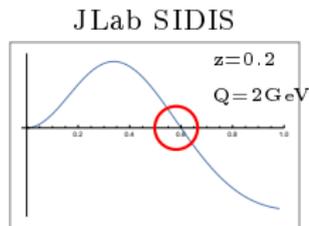
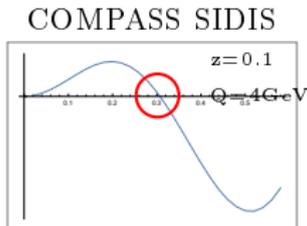
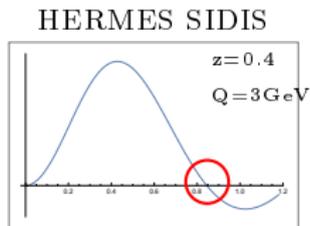
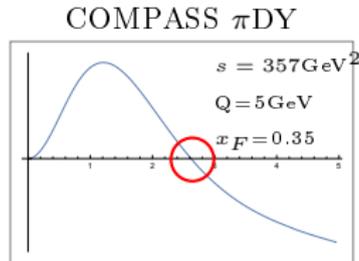
- ▶ For asymmetries (ratios of cross-section)  $\delta$  could be larger
- ▶ The method outcome strongly depends on precision of the data
  - ▶ Precise data  $\rightarrow$  more sensitivity to small effects, e.g. power corrections
  - ▶ E.g. for ATLAS ( $\sim 0.5\%$  accuracy) at  $\delta \sim 0.2$  deviation is  $\sim 2 - 3\%$  :(
  - ▶ E.g. for CDF ( $\sim 5\%$  accuracy) at  $\delta \sim 0.2$  deviation is  $\sim 2 - 3\%$  :)
- ▶ There could be models which **incorporates** power corrections to factorization into NP-behavior of TMDs
  - ▶ The result of extraction is not a TMD distribution (although it could perfectly describe the data), e.g. it violates universality
  - ▶ Anyway, at some moment TMD factorization fails ( $\rightarrow$  next slide)



The cross-section with LP TMD factorization eventually became negative.  
It happens at large  $q_T$



$$\frac{q_T}{Q} \sim 0.5 - 0.7$$



- ▶ The position of node depends also on process and  $x$
- ▶ For large- $Q$  bins the node can go down to  $q_T/Q \sim 0.35$  ( $\pi\text{DY}$  at COMPASS!)

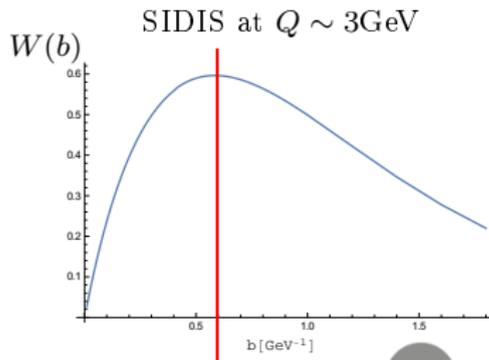
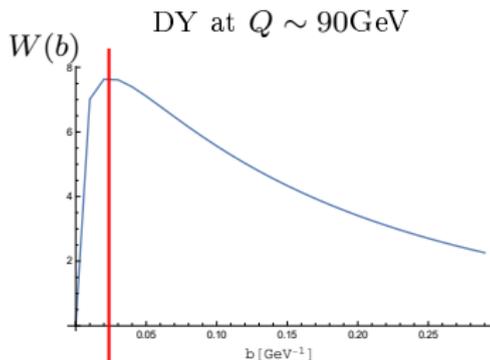


The negative cross-section is a small-b problem

$$d\sigma \simeq \int d^2b e^{-ibq_T} W(b) \simeq \int db b J_0(bq_T) W(b)$$

**When 2D Fourier is positive definite?**

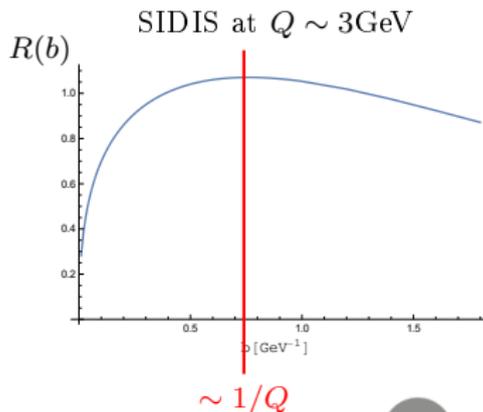
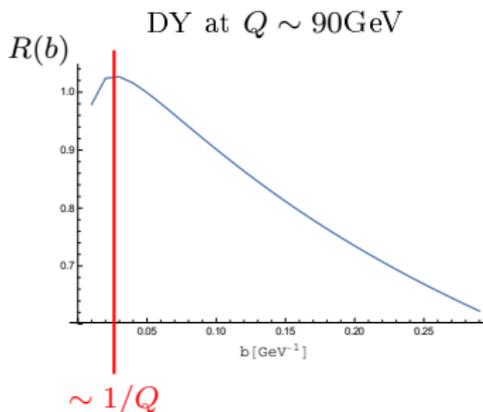
- ▶ **1D cos-transformation** → Bochner's theorem → "non-growing function"
- ▶ **Generally, it is a complicated question** see e.g. [Giraud, Peschanski, 1405.3155]
- ▶ **The first requirement** (but not sufficient):  $W(b)$  has maximum at  $b = 0$ .



This fall down is (mainly) due to TMD perturbative evolution !

$$W(b, Q; x, z) \simeq R(b, Q)^2 F(x, b) D(z, b)$$

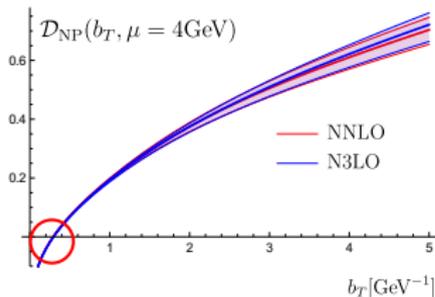
- ▶ In  $\zeta$ -prescription  $F$  and  $D$  are (almost) monotonous functions
- ▶  $R(b, Q) = \exp(-\mathcal{D}(b, Q) \ln(Q^2/\zeta_Q(b)))$



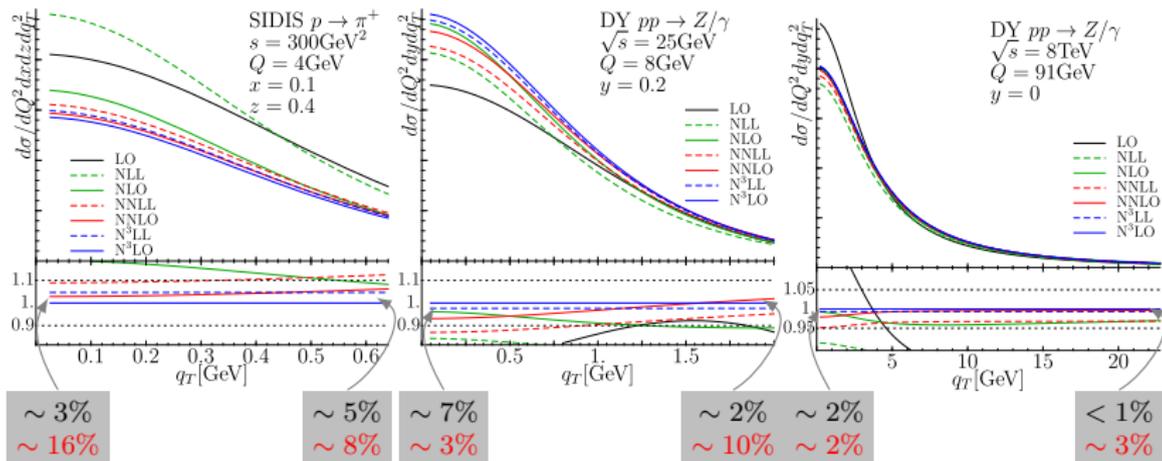
This fall down is (mainly) due to TMD perturbative evolution !

$$W(b, Q; x, z) \simeq R(b, Q)^2 F(x, b) D(z, b)$$

- ▶ In  $\zeta$ -prescription  $F$  and  $D$  are (almost) monotonous functions
- ▶  $R(b, Q) = \exp(-\mathcal{D}(b, Q) \ln(Q^2/\zeta_Q(b)))$



## Perturbation theory = predictive power



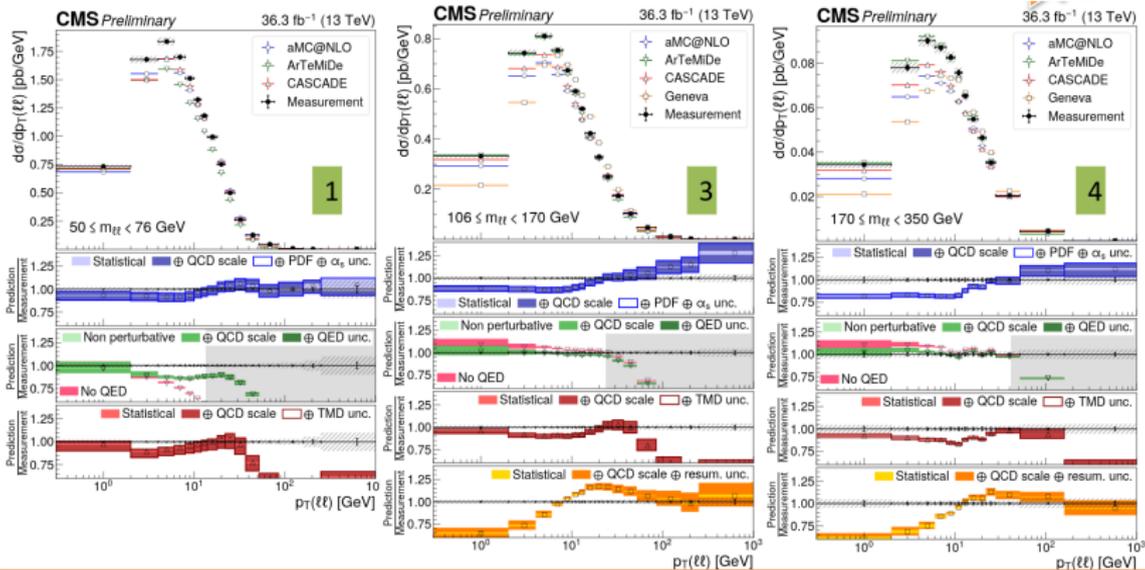
Difference between NNLO and N<sup>3</sup>LO is not that important

Difference between NLO and NNLO is important (**especially at low-energy!**)



# Perturbation theory = predictive power

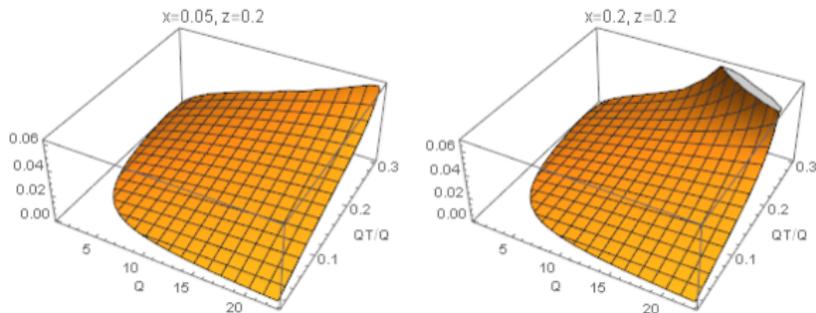
B. Bilin, DIS2021



## Principal problem for asymmetries

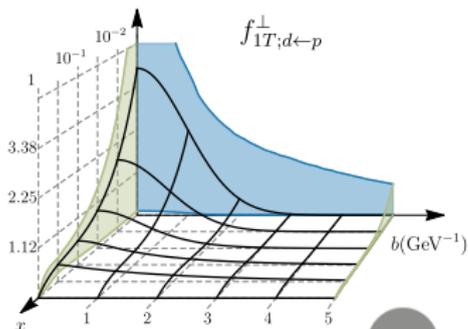
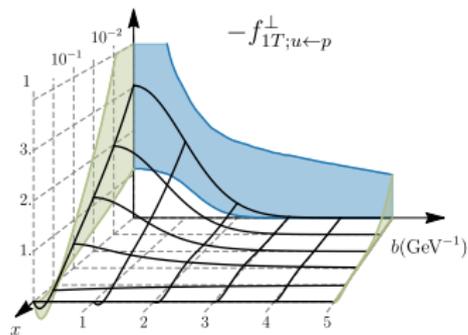
$$\text{asymmetry} \simeq \frac{\text{something}}{F_{UU}}$$

- ▶ Eventually  $F_{UU} = 0$  (in TMD factorization)



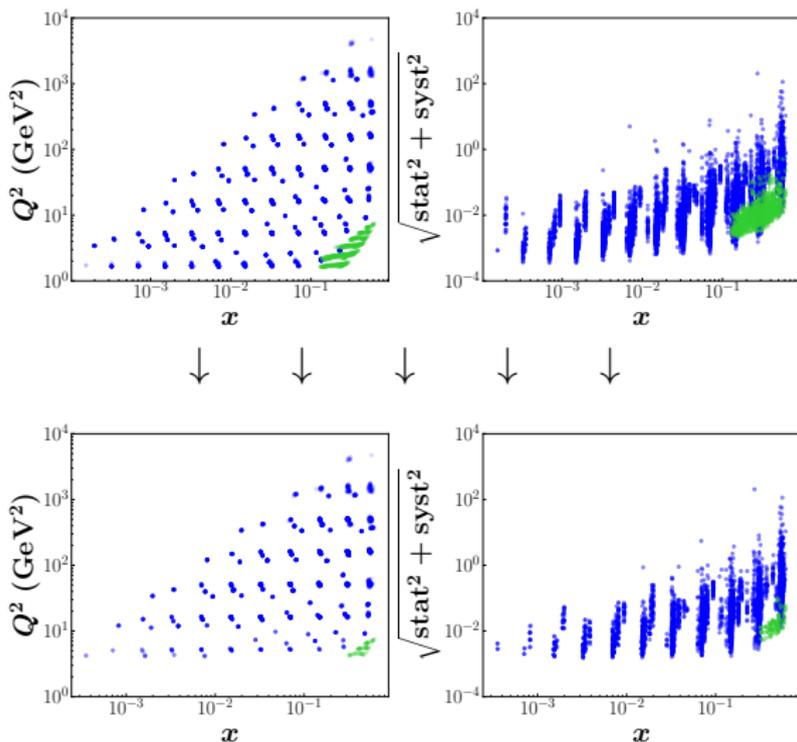
## Cutting present data : Siverts asymmetry

Dataset name	Ref.	Reaction	# Points
Compass08	[36]	$d^\uparrow + \gamma^* \rightarrow \pi^+$	1 / 9
		$d^\uparrow + \gamma^* \rightarrow \pi^-$	1 / 9
		$d^\uparrow + \gamma^* \rightarrow K^+$	1 / 9
		$d^\uparrow + \gamma^* \rightarrow K^-$	1 / 9
Compass16	[39]	$p^\uparrow + \gamma^* \rightarrow h^+$	5 / 40
		$p^\uparrow + \gamma^* \rightarrow h^-$	5 / 40
Hermes	[35]	$p^\uparrow + \gamma^* \rightarrow \pi^+$	11 / 64
		$p^\uparrow + \gamma^* \rightarrow \pi^-$	11 / 64
		$p^\uparrow + \gamma^* \rightarrow K^+$	12 / 64
		$p^\uparrow + \gamma^* \rightarrow K^-$	12 / 64
JLab	[41, 42]	$p^\uparrow + \gamma^* \rightarrow \pi^+$	1 / 4
		$p^\uparrow + \gamma^* \rightarrow \pi^-$	1 / 4
		$p^\uparrow + \gamma^* \rightarrow K^+$	1 / 4
		$p^\uparrow + \gamma^* \rightarrow K^-$	0 / 4
SIDIS total			63
CompassDY	[40]	$\pi^- + d^\uparrow \rightarrow \gamma^*$	2 / 3
Star.W+	[43]	$p^\uparrow + p \rightarrow W^+$	5 / 5
Star.W-		$p^\uparrow + p \rightarrow W^-$	5 / 5
Star.Z		$p^\uparrow + p \rightarrow \gamma^*/Z$	1 / 1
DY total			13
Total			76



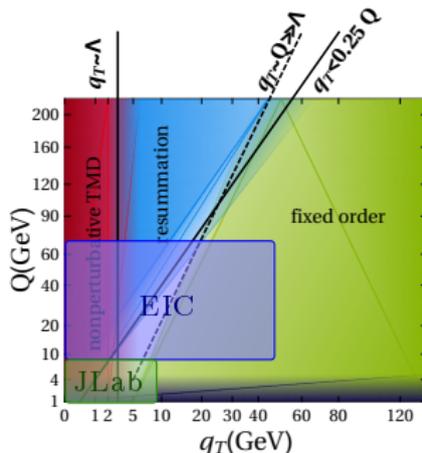
# Cutting future data: EIC and SoLID

plots by A.Prokudin



## Conclusion for part 1

- ▶ **There is a natural limit of TMD factorization**  $q_T < (0.2 - 0.3)Q$ 
  - ▶ This limit is required from by theory
  - ▶ This limit is also seen in the data
- ▶ Pushing this limit higher does not help practically
  - ▶ At  $q_T \sim 0.5Q$  cross-section become negative
  - ▶ It is pure perturbative effect
- ▶ Ways out:
  - ▶ Interpolate to fix order (works only at large  $Q$ )
  - ▶ Introduce  $b_{\min}$
  - ▶ Go to power corrections
  - ▶ ...
- ▶ **A lot of stuff to explore** especially at lower energy JLab  $\rightarrow$  EIC



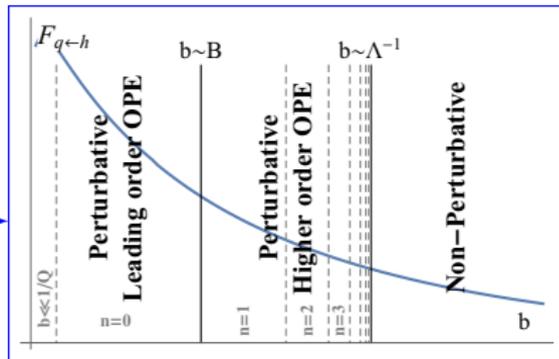
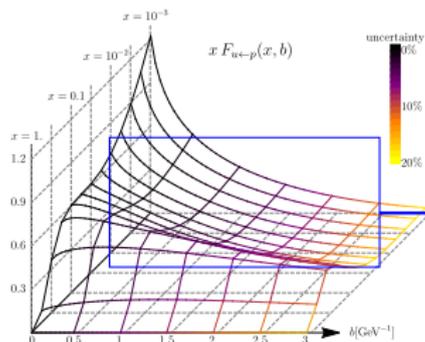
# Part II

## PDF-bias

TMD distributions are independent 3D function for each flavor

$$\text{TMDPDF} = F(x, b), \quad \text{TMDFF} = D(z, b)$$

**Too much freedom!**



$$F(x, b) = [q(x) + \alpha_s (p(x) \ln(b^2 \mu^2) + \dots) + \alpha_s^2 \dots] + b^2 \dots + \dots$$

Lead.power OPE  
up  $N^3\text{LO}$

Higher power OPE  
e.g. [V.Moos, AV, 2008.01744]

$$F(x, b) = C(x, b) \otimes q(x) f_{NP}(x, b)$$



## Current status of the small- $b$ matching

refs. are defined in [V.Moos,AV,2008.01744]

Name	Function	Twist of leading matching	Twist-2 distributions in matching	Twist-3 distributions in matching	Order of leading power coef.function	Ref.
unpolarized	$f_1(x, b)$	tw-2	$f_1(x)$	-	N <sup>3</sup> LO ( $\alpha_s^3$ )	[21, 22]
Sivers	$f_{1T}^\perp(x, b)$	tw-3	-	$T(-x, 0, x)$	NLO ( $\alpha_s^1$ )	[23]
helicity	$g_{1L}(x, b)$	tw-2	$g_1(x)$	$\mathcal{T}_g(x)$	NLO ( $\alpha_s^1$ )	[16, 17]
worm-gear T	$g_{1T}(x, b)$	tw-2/3	$g_1(x)$	$\mathcal{T}_g(x)$	LO ( $\alpha_s^0$ )	[13, 14]
transversity	$h_1(x, b)$	tw-2	$h_1(x)$	$\mathcal{T}_h(x)$	NNLO ( $\alpha_s^2$ )	[19]
Boer-Mulders	$h_1^\perp(x, b)$	tw-3	-	$\delta T_e(-x, 0, x)$	LO ( $\alpha_s^0$ )	[14]
worm-gear L	$h_{1L}^\perp(x, b)$	tw-2/3	$h_1(x)$	$\mathcal{T}_h(x)$	LO ( $\alpha_s^0$ )	[13, 14]
pretzelocity	$h_{1T}^\perp$	tw-3/4	-	$\mathcal{T}_h(x)$	LO ( $\alpha_s^0$ )	eq.(4.8)

► Twist-2 and twist-3 contributions at all powers of  $b^2$  (tree)

► Typical expression (here for Sivers function):

$$f_{1T}^\perp(x, b) = \pm \pi \left\{ T_q(x) + \sum_{n=1}^{\infty} \left( \frac{x^2 b^2 M^2}{4} \right)^n \int_0^1 du \int dy \frac{\delta(x - uy)}{(n+1)!(n-1)!} \left( \frac{\bar{u}}{u} \right)^n \frac{1 + (n-1)u + u^2}{1-u} T_q(y) \right\}$$

►  $T(x) = T(-x, 0, x)$  is Qiu-Sterman function

►  $\text{TMDs}_{proton} \sim \text{TMDs}_{nuclei}$

► Non-trivial matching for pretzelocity

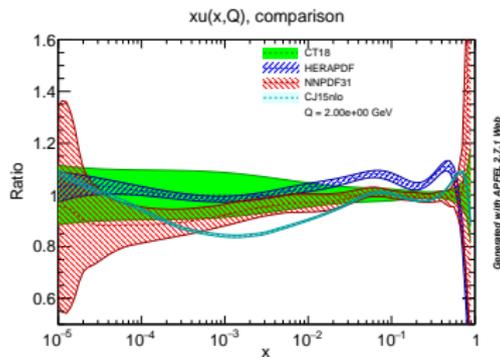
► Leading term:  $h_{1T}^\perp(x, b) = -x^2 \int_x^1 \frac{du}{u} \frac{1-u^2}{u} \mathcal{T}_h\left(\frac{x}{u}\right) + \text{tw-4}$



## Matching is essential part on nowadays TMD phenomenology

- ▶ The region  $5\text{GeV} \lesssim q_T < 0.25Q$  is accurately described by  $f_{NP} \sim 1$ 
  - ▶ LHC, Tevatron, RHIC, ( $\rightarrow$  EIC)
- ▶ It is observed that  $q(x)$  carries the most part of  $x$ -dependence. I.e.  $f_{NP}(x, b) \sim f_{NP}(b)$ 
  - ▶ Greatly reduces the parametric freedom
- ▶ It is observed that  $q_f(x)$  carries the most part of the flavor dependence, i.e.  $f_{NP}(x, b) \sim \text{flavor-independent}$
- ▶ In fact, the simplest model  $F_f(x, b) \sim C \otimes q_f(x) f_{NP}(b)$  capable to describe the most part of the data rather accurately

Matching to PDF leads to high predictive power, but is also a pitfall  $\rightarrow$



**Result of a TMD fit is 100% dependent on PDF in use!**

- ▶ Different PDF set are different
  - ▶ Especially in a “TMD-important” region  $x \sim 0.1 - 0.5$
  - ▶ Different flavor decomposition
- ▶ As the result:

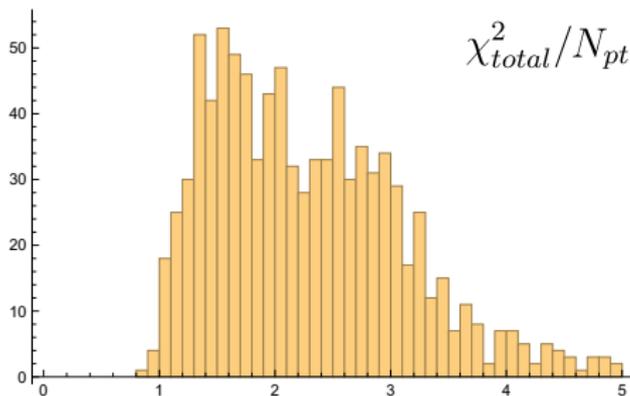
PDF & FF sets	$\chi^2/N_{pt}$
HERA20 & DSS	0.76
HERA20 & JAM19	0.93
NNPDF31 & DSS	1.00
NNPDF31 & JAM19	1.65
HERA20 & DSS ( $N^3LO$ )	0.88
NNPDF31 & DSS ( $N^3LO$ )	1.31

SIDIS+DY fit [SV19]

Obviously, one must include PDF uncertainty into the fit

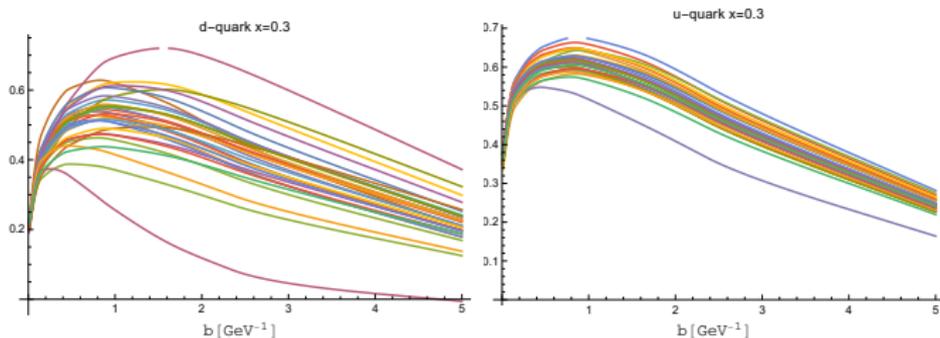
## Including PDF uncertainty “straightforwardly”

- ▶ PDF uncertainty is **larger** than the experimental precision
  - ▶ LHC 5-7% vs. 1%
  - ▶ Low energy DY 10-50% vs. 10%
  - ▶ SIDIS 10-50% vs.
- ▶ TMD physics (in comparison to DIS) is sensitive to different  $x$ -domain
- ▶ **Strongly depends on the set**



← SV19 fit  
NNPDF+DSS  
SIDIS+DY  
 $N_{pt} = 1039$   
fit made for central replica  
← distribution of  $\chi^2$  for  
1000 replicas of NNPDF

PDF essentially changes behavior of TMD = **PDF-bias**



SV19 fit, 40 random replicas of NNPDF3.1

In fact, each PDF replica must be equipped by its own  $f_{NP}$   
It will partially compensate PDF-bias  
So, together they form a TMD distribution

Inclusion of PDF uncertainty into TMD fit (**work in progress**)

## Computationally intensive work

- ▶ Represent PDF uncertainty as MC replicas (1000 replicas)
- ▶ Make a fit of TMD distribution, based on each replica

One could expect that result would be less dependent on PDF

Inclusion of PDF uncertainty into TMD fit (**work in progress**)

## Computationally intensive work

- ▶ Represent PDF uncertainty as MC replicas (1000 replicas)
- ▶ Make a fit of TMD distribution, based on each replica

One could expect that result would be less dependent on PDF

**NO**

We have observed that “simple” (5 params!) model from SV19 does not fit all PDF sets equally well

**Reason:** absence of flavor dependence

**Solution:** add flavor dependence  $f_{NP} \rightarrow f_{NP}^{u,d,\bar{u},\bar{d},rest}$

DY only (457 points)

SV19 model

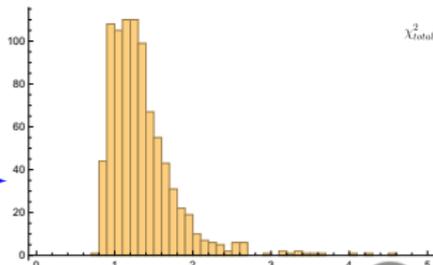
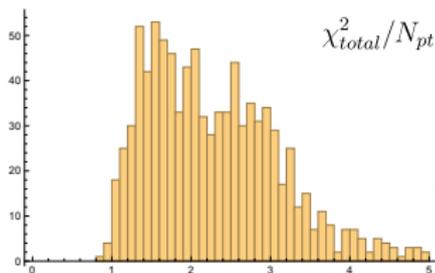
input PDF .	$\chi^2/N_{pt}$
HERA20	0.97
NNPDF31	1.14
CT18	1.26
MSHT20	1.39



DY only (457 points)

flavor-dependent model

input PDF .	$\chi^2/N_{pt}$
HERA20	0.90
NNPDF31	0.97
CT18	0.98
MSHT20	0.89

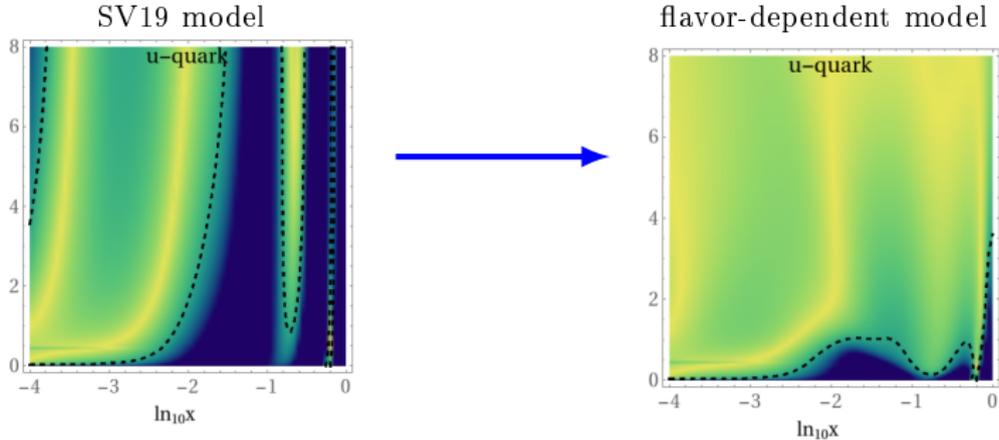


Plot for NNPDF3.1, but similar picture for other PDFs

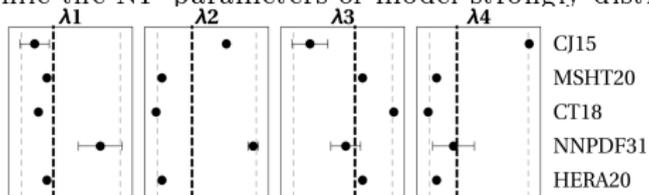


Now, TMDs based on different PDFs are in agreement

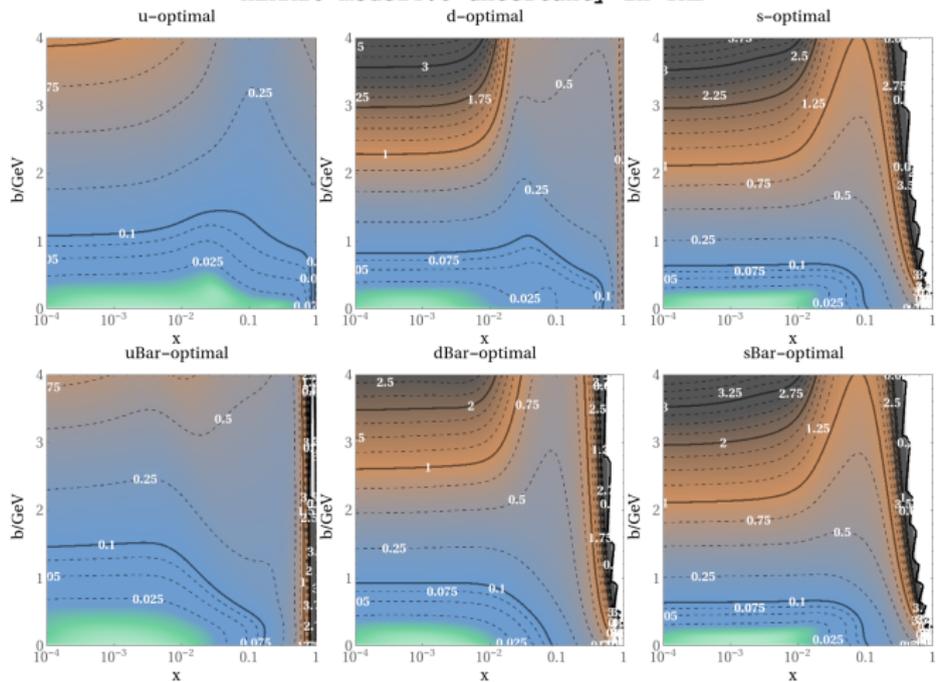
Intersection of TMDs based on HERA and CT18 within uncertainties



Meanwhile the NP parameters of model strongly distributed

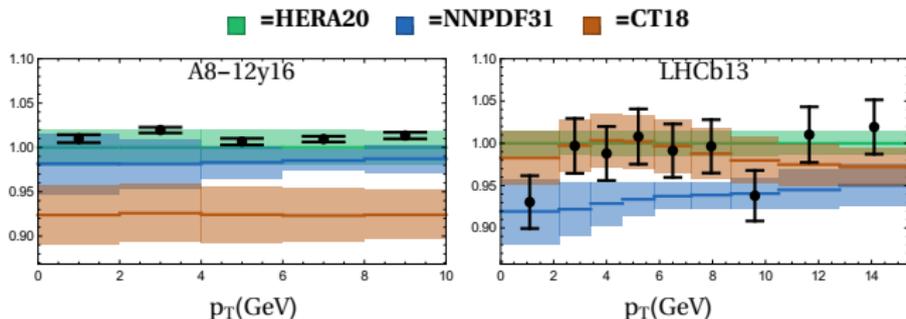


# HERA20 model4.0 uncertainty in TMD



## Conclusion for part 2

- ▶ **Matching of TMD to PDF is important!**
- ▶ TMD totally dependent on the PDF in use
  - ▶ There is no agreement between PDF sets
  - ▶ Uncertainty in PDF lead to crazy TMDs → fit each PDF replica
- ▶ To compensate PDF-bias one needs flavor dependence
  - ▶ Results for different PDFs are in agreement
  - ▶ TMDs are in agreement
  - ▶ Uncertainty on TMD is much larger
- ▶ **Work in progress**
- ▶ Future: one needs joint fit of PDF + TMD



# Part III

## Decorrelation of TMD evolution

TMD evolution depends on non-perturbative CS-kernel

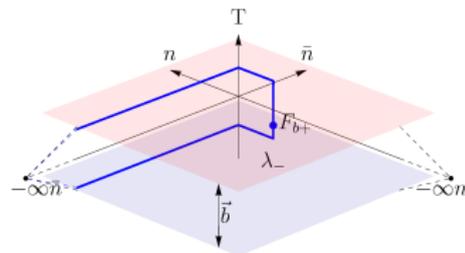
$$\mathcal{D} = -\frac{1}{2}K = \frac{1}{2}F_{q\bar{q}} = -\frac{1}{2}\gamma_\nu^{f\perp}$$

$$\begin{aligned}\mu^2 \frac{d}{d\mu^2} F_{f\leftarrow h}(x, b; \mu, \zeta) &= \frac{\gamma_F^f(\mu, \zeta)}{2} F_{f\leftarrow h}(x, b; \mu, \zeta) \\ \zeta \frac{d}{d\zeta} F_{f\leftarrow h}(x, b; \mu, \zeta) &= -\mathcal{D}^f(b, \mu) F_{f\leftarrow h}(x, b; \mu, \zeta)\end{aligned}$$

CS is fundamental QCD function

$$\mathcal{D}(b, \mu) = \lambda_- \frac{ig \operatorname{Tr} \int_0^1 d\beta \langle 0 | F_{b+}(-\lambda_- n + b\beta) W_{C'} | 0 \rangle}{\operatorname{Tr} \langle 0 | W_{C'} | 0 \rangle} + Z_{\mathcal{D}}(\mu)$$

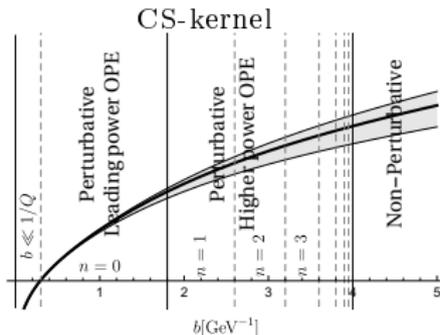
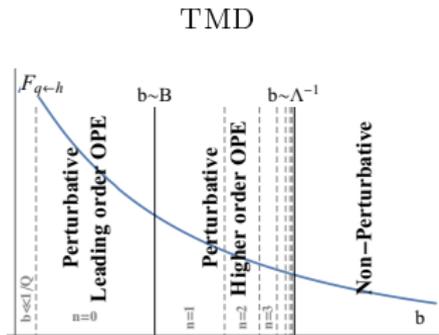
- ▶ Independent observable
- ▶ Measures QCD-vacuum



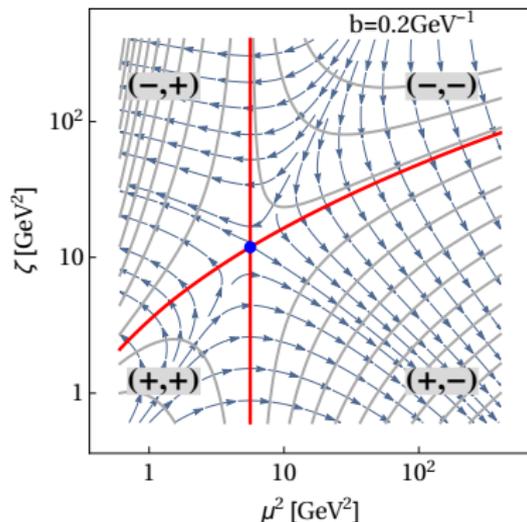
Thus, TMD distributions are functionals of CS-kernel

$$f_1[\mathcal{D}](x, b; \mu, \zeta)$$

- ▶ Is it a problem? **YES**, because we extract simultaneously  $\mathcal{D}$  and TMDs
  - ▶ Extraction is not universal!
    - ▶ E.g. one cannot use  $\mathcal{D}$  from lattice, together with TMDs from pheno.
    - ▶ In principle, very large/broad pull of data will reduce correlation
    - ▶ Problem of comparison/interpretation of result
- ▶ One makes situation **worse by splitting to perturbative and NP parts**
  - ▶ **Keep CS-kernel a whole function**



There are several solutions for this problem  
my-preferred is  $\zeta$ -prescription



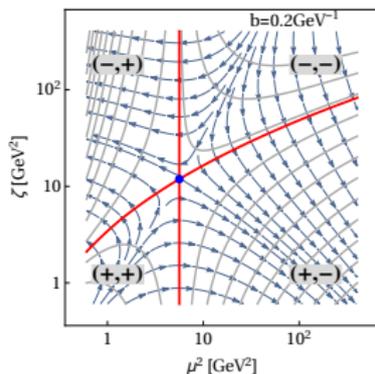
- ▶ In a nutshell: define TMD at  $\zeta(b, \mu)$ 
  - ▶  $\zeta(b, \mu)$  = equi-evolution line (NP!)
- ▶ Equivalent to fixed-point definition
  - ▶ TMDs on the same equi-evolution line are the same (by definition!)
- ▶ Generally: does not matter which line use as reference
- ▶ But there is one **very special line** = which passes through the saddle point
  - ▶ TMD on this line = optimal TMD

## Why optimal TMD is optimal?

- 1 At saddle point  $\mathcal{D} = 0$

$$(\text{optimal}) f_1(x, b) = f_1[\mathcal{D} = 0](x, b; (\mu, \zeta)_{\text{saddle}})$$

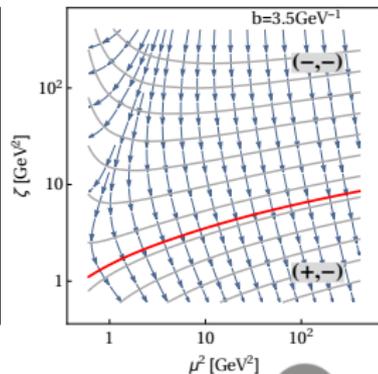
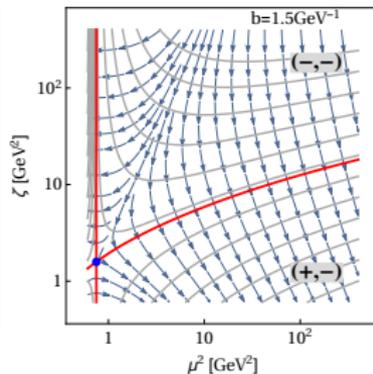
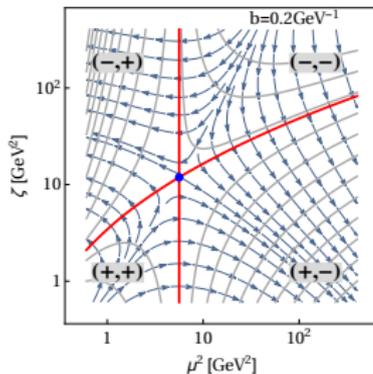
- 2 Optimal equi-potential line is continuous (important for small-b matching)
- 3 Greatly simplifies all equations



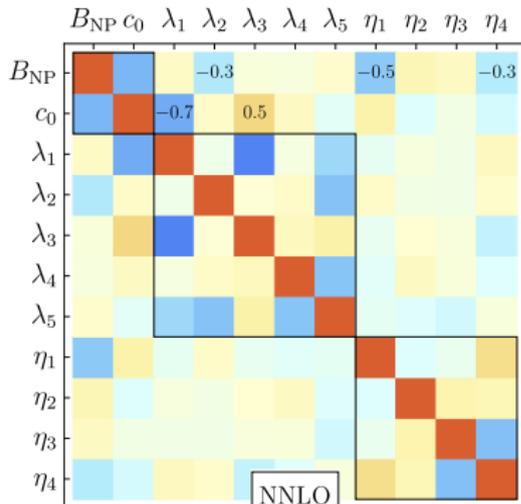
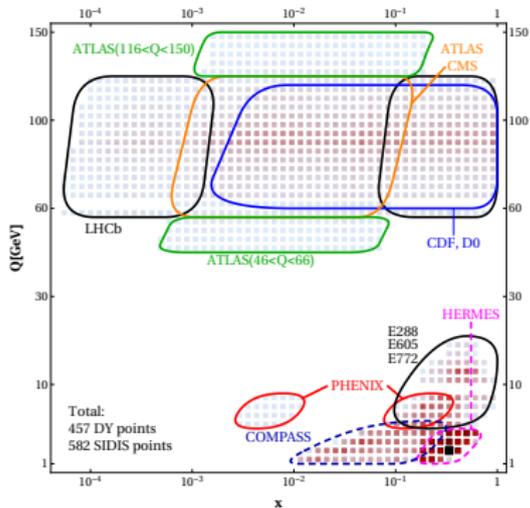
Everything is nonperturbative!  
 Position of saddle-point,  $\zeta$ -line,...  
 Solve all equations in terms of  $\mathcal{D}$ !

- ▶ At large- $b$  saddle-point goes below  $\Lambda_{QCD}$ .
  - ▶ Not possible to build perturbative-like solution
  - ▶ But there is an **exact solution!** (see [Scimemi,AV,1912.06532,app.C])

$$2\mathcal{D} + 2\beta(a_s) \frac{\partial g(a_s, \mathcal{D})}{\partial a_s} - \Gamma_{\text{cusp}}(a_s) \frac{\partial g(a_s, \mathcal{D})}{\partial \mathcal{D}} + \gamma_V(a_s) = 0, \quad , \quad g(a_s, 0) = 0$$

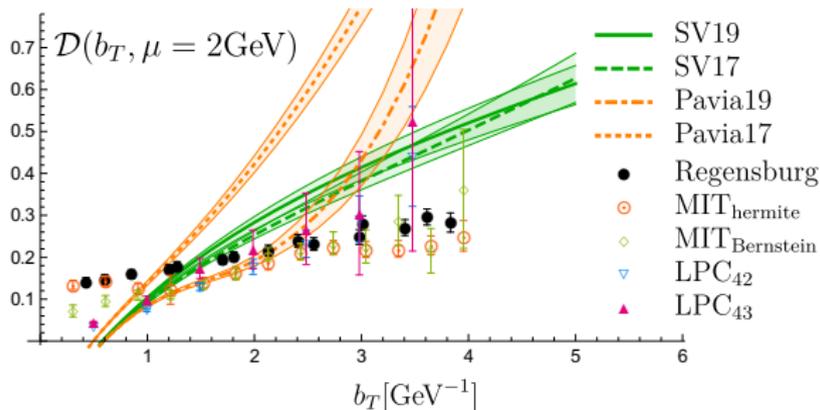


## CS-kernel still correlated with the TMDs



## Conclusion for part 3

- ▶ Independent extraction of evolution and TMDs is cumbersome task
- ▶ **Keep CS-kernel as a whole function!**
- ▶ Fixed-scale schemes are preferable
  - ▶  $\zeta$ -prescription



## Conclusion

- ▶ Extraction of TMDs is a very peculiar task
  - ▶ Involves several NP functions
  - ▶ Requires strict data-cuts
  - ▶ Perturbative input is important
  - ▶ **Many open theoretical questions**
- ▶ JLab
  - ▶ There will be not much **pure** TMD-factorizable data
    - ▶ Large- $x$
  - ▶ Paradise to study power corrections
    - ▶ Higher-twist TMDs
    - ▶ Mass/kinematic corrections
    - ▶ Interesting and weakly studied field
    - ▶ current/next frontier of QCD
  - ▶ Going to be the challenge for theoreticians
- ▶ EIC
  - ▶ Will happen in 10+ years (I doubt that our understanding will remain the same)

Always take into account uncertainties!

