Extractions of TMDs using neural networks

Valerio Bertone

IRFU, CEA, Université Paris-Saclay

(A MAP Collaboration effort, mostly in collaboration with C. Bissolotti, M. Cerutti, S. Rodini)

Partially based on:

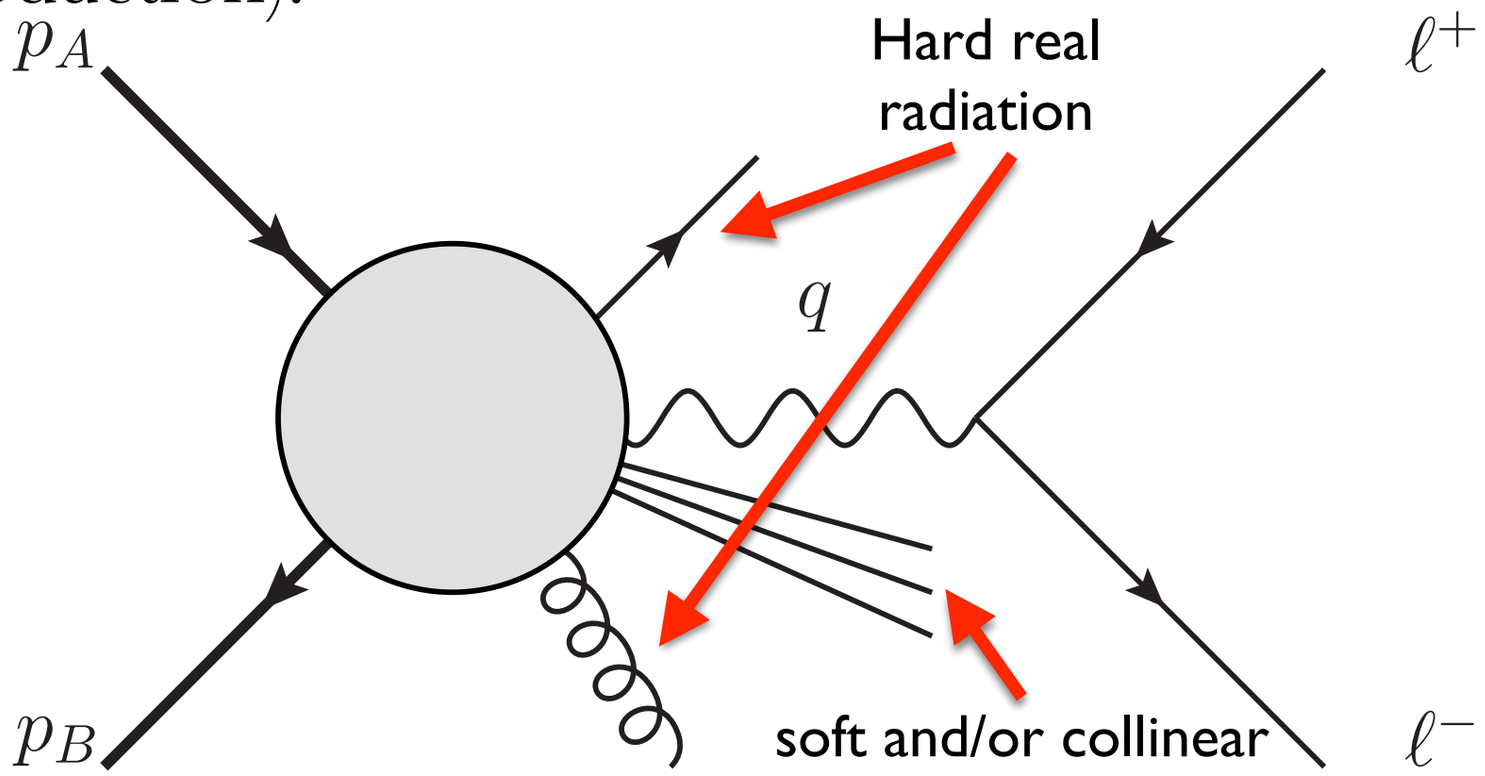
Phys. Rev. Lett. 135 (2025) 2, 021904

Phys. Rev. Lett. 136 (2026) 17, 171902

May 12, 2026, QCD Evolution, El Escorial (Madrid)

Analysis of a q_T distribution

- Take the inclusive production of a lepton pair in *pp* collisions (a.k.a. Drell-Yan production):



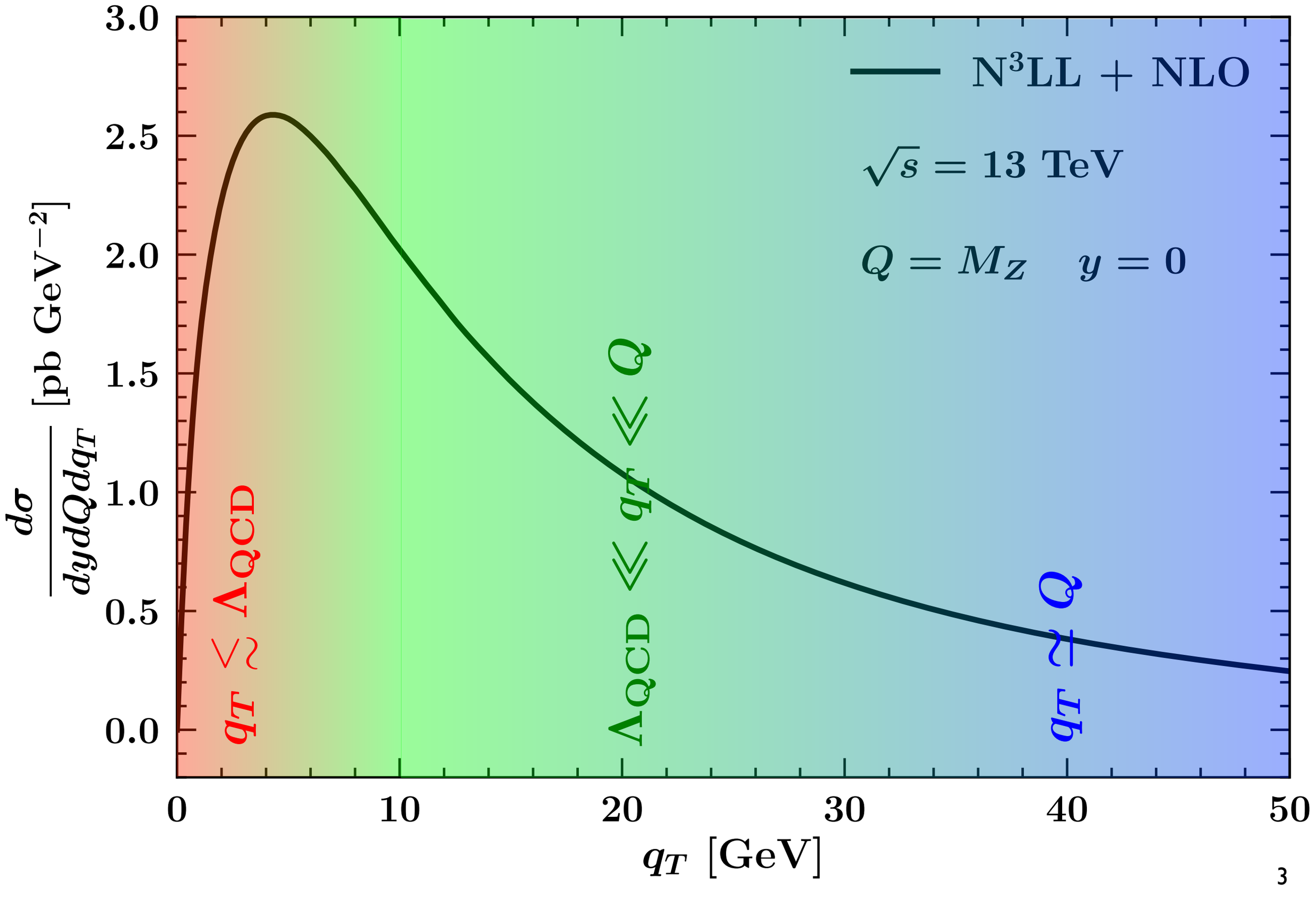
- **Three** relevant scales (four, in fact, considering also $\sqrt{s}$, but we assume $\sqrt{s} \lesssim Q$):

$$Q = \sqrt{q^2}\,, \qquad \Lambda_{\rm QCD}\,, \qquad q_T = \sqrt{q_x^2 + q_y^2}$$

- We must require $Q \gg \Lambda_{\rm QCD}$ to guarantee applicability of pQCD.
- q_T acts as a needle of the scale defining **three possible regimes**:

$$q_T \lesssim \Lambda_{\rm QCD}\,, \qquad \Lambda_{\rm QCD} \ll q_T \ll Q\,, \qquad q_T \simeq Q$$

Analysis of a q_T distribution



Analysis of a q_T distribution

- In each regime, **ratios of scales** enter both *powers* and *logarithms*:
  - size of powers actually define the different regimes.
  - Logarithms must often be *resummed* to all orders in α_s because otherwise they spoil the perturbative series (corrections scale like $\alpha_s^n L^m$).

	Power corrections			Logarithms		
	$\left(\frac{\Lambda}{q_T}\right)^{\alpha}$	$\left(\frac{\Lambda}{Q}\right)^{\alpha}$	$\left(\frac{q_T}{Q}\right)^{\alpha}$	$\ln\left(\frac{\Lambda}{q_T}\right)$	$\ln\left(\frac{\Lambda}{Q}\right)$	$\ln\left(\frac{q_T}{Q}\right)$
$q_T \simeq Q$	neglected		Order-by-order in pQCD: Coll. fact.	resummed through α_s		Order-by-order in pQCD: Coll. fact.
$\Lambda \ll q_T \ll Q$	neglected			resummed through α_s		Resummed through TMD fact.
$q_T \lesssim \Lambda$	parametrised	neglected		resummed through α_s		Resummed through TMD fact.

- The ultimate goal of a **TMD fit** is the determination of the Λ_{QCD}/q_T power corrections, which are directly related to non-pertubative effects.

TMD factorisation

- TMD factorisation (valid for $q_T \ll Q$) can be written as:

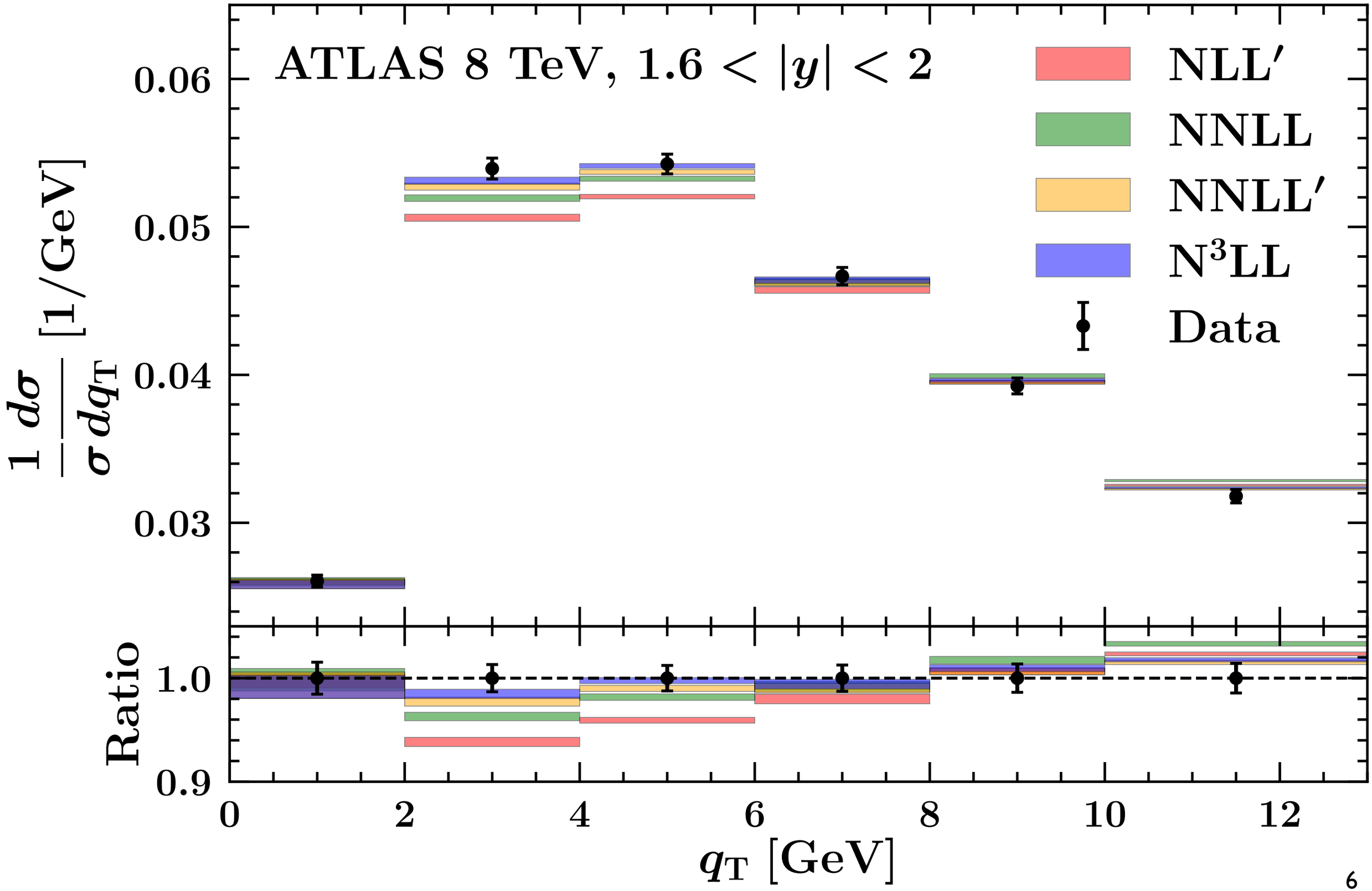
$$\frac{d\sigma}{dydQdq_T} = \sigma_0 H(Q) \int_0^\infty db_T\, b_T J_0(q_T b_T) F_1(x_1, b_T, Q, Q^2) F_2(x_2, b_T, Q, Q^2)$$

- **Evolution** and **matching** give us the single TMD distributions:

$$F = \exp\left\{ K \ln\frac{\sqrt{\zeta}}{\mu_b} + \int_{\mu_b}^{\mu} \frac{d\mu'}{\mu'} \left[\gamma_F - \gamma_K \ln\frac{\sqrt{\zeta}}{\mu'} \right] \right\} C \otimes f \qquad \mu_b = \frac{2e^{-\gamma_E}}{b_T}$$

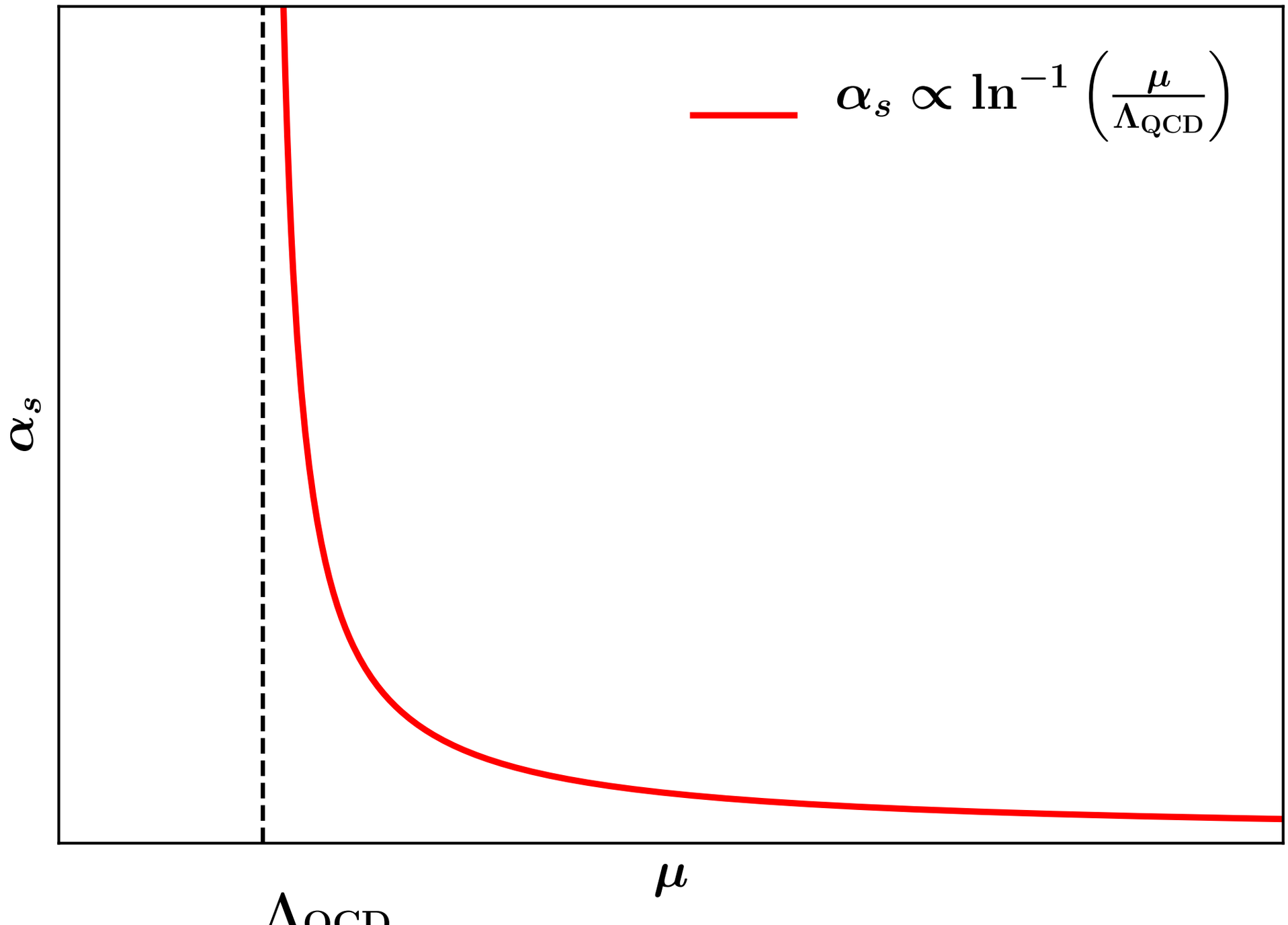
Accuracy	γ_K	γ_F	K	C	H	PDFs/FFs
LL	α_s	-	-	1	1	-
NLL	α_s^2	α_s	α_s	1	1	LO
NLL’	α_s^2	α_s	α_s	α_s	α_s	LO
N^{2}LL	α_s^3	α_s^2	α_s^2	α_s	α_s	NLO
N^{2}LL’	α_s^3	α_s^2	α_s^2	α_s^2	α_s^2	NLO
N^{3}LL	α_s^4	α_s^3	α_s^3	α_s^2	α_s^2	NNLO
N^{3}LL’	α_s^4	α_s^3	α_s^3	α_s^3	α_s^3	NNLO
N^{4}LL	α_s^5	α_s^4	α_s^3	α_s^3	α_s^3	N^{3}LO

TMD factorisation



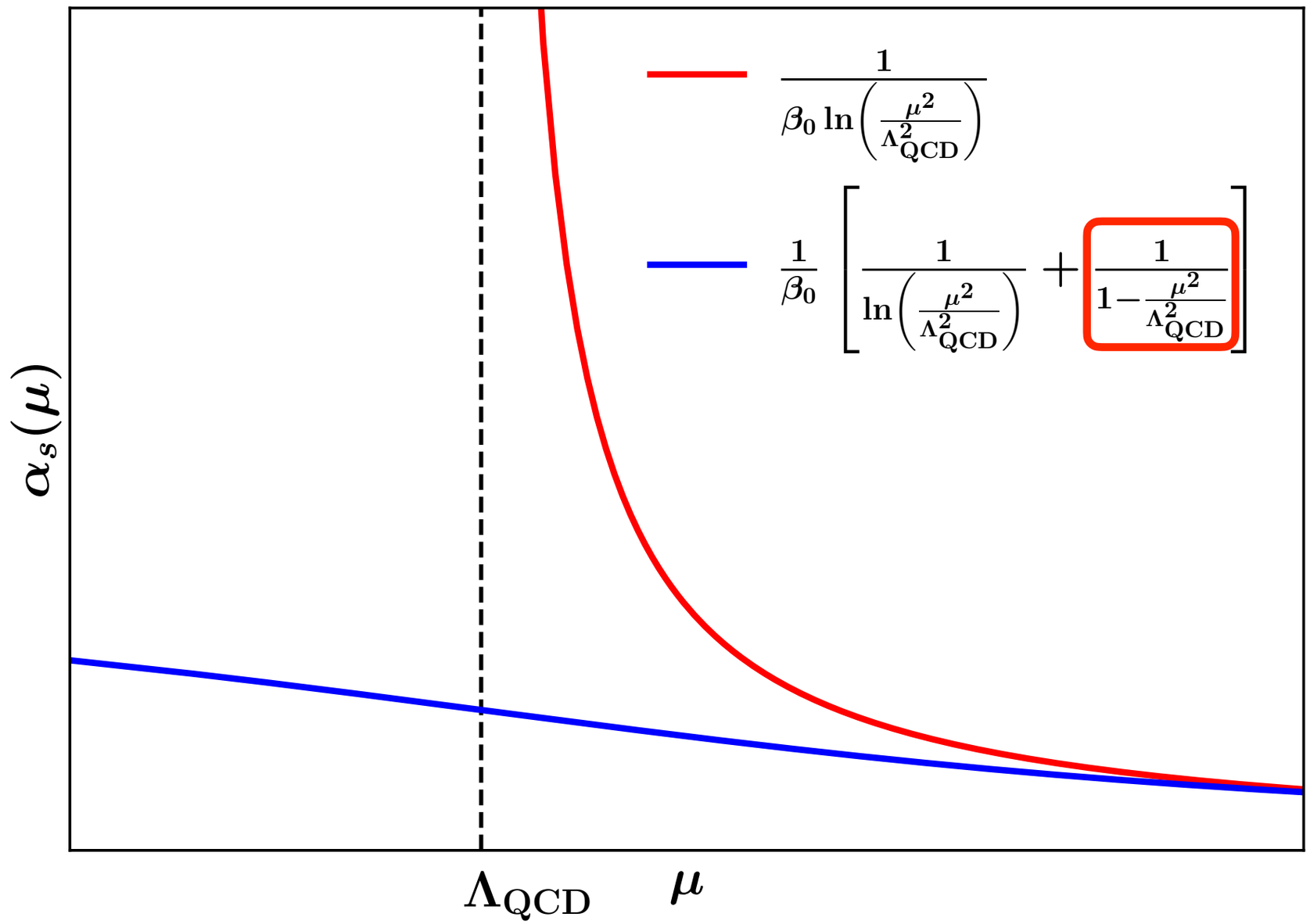
Origin of $\mathcal{O}(\Lambda_{\rm QCD}/q_T)$ corrections

$$\sigma \propto \int_0^\infty db_T\, \alpha_s^p\left(\mu_b\right) \cdots \simeq \int_0^\infty db_T\, \alpha_s^p\left(1/b_T\right) \ldots$$



Origin of $\mathcal{O}(\Lambda_{\rm QCD}/q_T)$ corrections

- **Prescriptions** to avoid integrating over the **Landau pole** introduce power corrections $\mathcal{O}(\Lambda_{\rm QCD}/q_T)$ of **non-perturbative origin**.
  - different prescriptions distribute these corrections differently.
- To see this, one can use a **dispersive** approach [Bogolyubov, Shirkov; Dokshitzer, Marchesini, Webber]:



Origin of $\mathcal{O}(\Lambda_{\rm QCD}/q_T)$ corrections

- A popular regularisation prescription is the so-called b_*-prescription:

$$b_*(b_{\rm T}) = \frac{b_{\rm T}}{\sqrt{1+b_{\rm T}^2/b_{\rm max}^2}} \qquad \Longrightarrow \qquad \mu_{b_*} = \frac{2e^{-\gamma_E}}{b_*(b_T)}$$

$$F(x,b_T;\mu,\zeta) = F(x,b_*(b_T);\mu,\zeta)\left[\frac{F(x,b_T;\mu,\zeta)}{F(x,b_*(b_T);\mu,\zeta)}\right] \equiv F(x,b_*(b_T);\mu,\zeta) f_{\rm NP}(x,b_T;\zeta)$$

- **Important**: $f_{\rm NP}$ is *not universal* as it depends on the specific choice of b_* or, more in general, on the strategy used to regularise the Landau pole.

- One finally has:

$$F(x,b_T,\mu,\zeta) = [C\otimes f](x,\mu_{b_*})$$

Perturbative matching on collinear distributions

$$\times \quad \exp\left\{K(\mu_{b_*})\ln\frac{\sqrt{\zeta}}{\mu_{b_*}} + \int_{\mu_{b_*}}^{\mu}\frac{d\mu'}{\mu'}\left[\gamma_F(\alpha_s(\mu')) - \gamma_K(\alpha_s(\mu'))\ln\frac{\sqrt{\zeta}}{\mu'}\right]\right\}$$

Perturbative evolution (resummation of $\ln(q_T/Q)$)

$$\times \quad \exp\left[g(x,b_T) + g_K(b_T)\ln\left(\frac{\zeta}{Q_0^2}\right)\right]$$

Non-perturbative contribution (powers of $\Lambda_{\rm QCD}/q_T$)

Determining TMDs

- Extracting TMDs boils down to determining:

$$f_{\mathrm{NP}}^{(i)}(x, b_T; \zeta) = \exp\left[g_i(x, b_T) + g_K(b_T) \ln\left(\frac{\zeta}{Q_0^2}\right)\right]$$

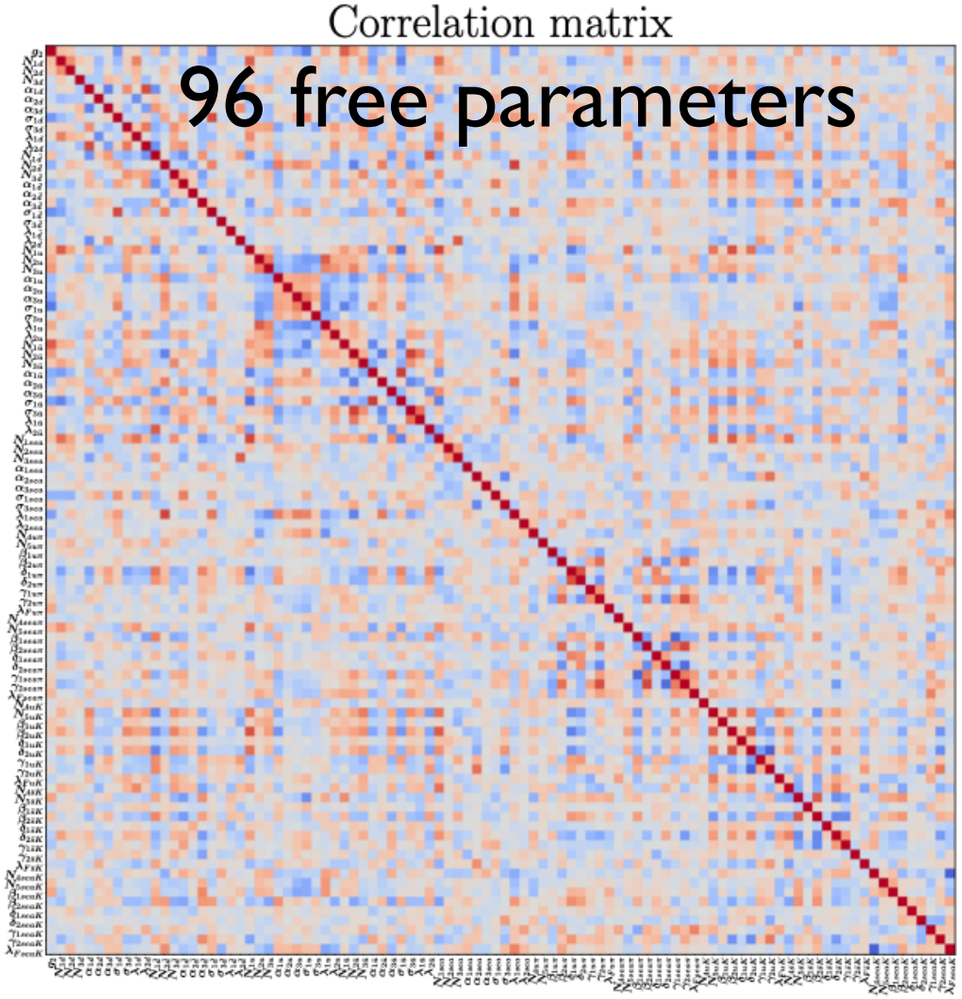
- Usually done through **fits to experimental data**.
- There is not a huge latitude in defining f_{NP}:
  - Possible **flavour dependence**.
  - g_i and g_K have to go to zero as b_T tends to zero, and become large (and negative) as b_T becomes large:
    - f_{NP} can be sensibly different from one only for $b_T \gtrsim \Lambda_{\mathrm{QCD}}^{-1}$.
  - g_K (non-perturbative contribution to the Collins-Soper kernel) can be determined faithfully having a **large lever arm** in $\zeta \propto Q$:
    - **(very) high invariant-mass data** necessary (but scarce at the moment).
    - **Lattice** simulations can help (see below).

Parametrising TMDs

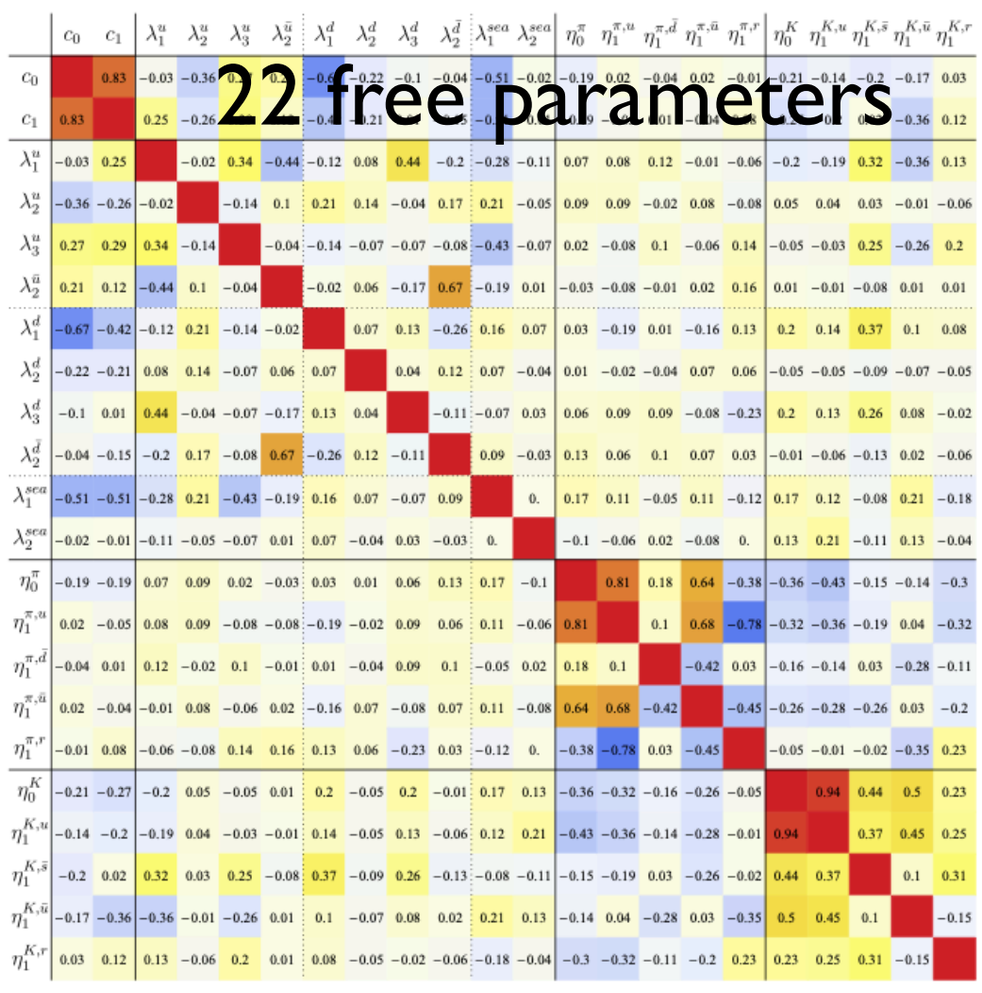
- TMD parametrisations were originally very simple:

$$f_{\rm NP}^{\rm DWS}(x, b_T; \zeta) = \exp\left[-g_1 b_T^2 - g_2 b_T^2 \ln\left(\frac{\zeta}{Q_0^2}\right)\right]$$ [*Nucl. Phys. B* 256 (1985) 413]

- They then became more involved as more data became available.



[MAP24, *JHEP* 08 (2024) 232]



[ART25, *JHEP* 11 (2025) 134]

- We are in a stage in which **TMD parametrisation** may significantly **bias** the extraction… as it happened for PDFs years ago.

Neural-Network TMDs

The parametrisation

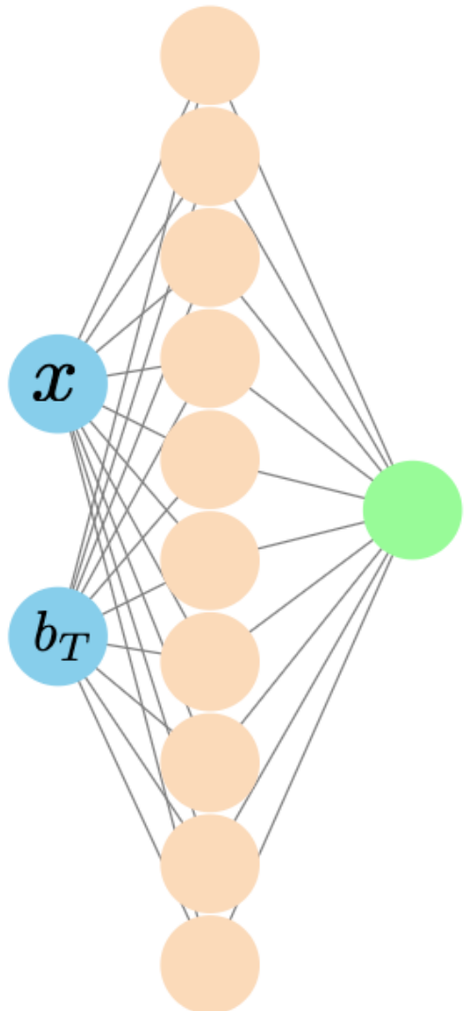
- Minimising the parametrisation bias calls for **neural networks**.
- In [*Phys. Rev. Lett.* 135 (2025) 2, 021904], we used:

$$f_{\rm NP}(x, b_T; \zeta) = \frac{\mathbb{NN}(x, b_T)}{\mathbb{NN}(x, 0)} \exp\left[-g_2^2 b_T^2 \log\left(\frac{\zeta}{Q_0^2}\right)\right]$$

- NN architecture [2, 10, 1]: inputs x and b_T, 10 hidden nodes, one output.
- Sigmoid-like activation function:

$$\sigma(z) = \frac{1}{2}\left(1 + \frac{z}{1+|z|}\right)$$

- $f_{\rm NP} \to 1$ as $b_T \to 0$ by construction.
- 42 free parameters: 41 NN + 1 for evolution (more on the CSK below).
- No flavour dependence (proof of concept).



Neural-Network TMDs

Fit setup

- χ^2 minimisation via Levenberg-Marquardt algorithm (Ceres-Solver).
  - account for experimental **correlations**:

$$\chi^2 = \sum_{ij} (d_i - t_i) C_{ij}^{-1} (d_j - t_j)$$

- **Analytic gradient** of χ^2 through analytic differentiation of the NN for fast and stable convergence.
- **Overfitting avoided** via cross-validation (training/validation split).
- Uncertainty propagated via **Monte Carlo** sampling: 250 data replicas.
- N^{3}LL perturbative accuracy:
  - kinematic cut $q_T/Q < 0.2$ to guarantee validity of **TMD factorisation**.
- MSHT2020 NNLO as a collinear PDF set:
  - **PDF uncertainty** fully accounted for.

Neural-Network TMDs

Dataset and fit quality

- Fit to Drell-Yan data only (proof of concept):
  - compare to MAP22 (DY only) to gauge the **impact of parametrisation**.

Experiment	$N_{\rm dat}$	$\chi^2/N_{\rm dat}$ $(\bar{\chi}^2_D/N_{\rm dat} + \bar{\chi}^2_\lambda/N_{\rm dat})$	
		NN	MAP22
Fixed-target	233	$1.08\,(0.98 + 0.10)$	$0.91\,(0.70 + 0.21)$
RHIC	7	$1.11\,(1.03 + 0.07)$	$1.45\,(1.37 + 0.08)$
Tevatron	71	$0.80\,(0.73 + 0.06)$	$1.20\,(1.17 + 0.04)$
LHCb	21	$0.98\,(0.88 + 0.10)$	$1.25\,(1.05 + 0.20)$
CMS	78	$0.40\,(0.38 + 0.02)$	$0.41\,(0.35 + 0.06)$
ATLAS	72	$1.38\,(1.09 + 0.29)$	$3.51\,(3.03 + 0.49)$
Total	482	$0.97\,(0.86 + 0.11)$	$1.28\,(1.09 + 0.20)$

$$\chi^2 = \sum_i \left(\frac{d_i - \bar{t}_i}{\sigma_{i,\rm unc}} \right)^2 + \sum_\alpha \lambda_\alpha^2 = \chi^2_D + \chi^2_\lambda$$

$$\bar{t}_i = t_i + \sum_\alpha \lambda_\alpha \sigma^{(\alpha)}_{i,\rm corr}$$

Neural-Network TMDs

Dataset and fit quality

- Fit to Drell-Yan data only (proof of concept):
  - compare to MAP22 (DY only) to gauge the **impact of parametrisation**.

- Globally, NN performs significantly better that MAP22:
  - significantly lower penalty $\chi^2_\lambda \Rightarrow$ **less reliance on systematic shifts**.

Experiment	$N_{\rm dat}$	$\chi^2/N_{\rm dat}$ $(\bar{\chi}^2_D/N_{\rm dat} + \bar{\chi}^2_\lambda/N_{\rm dat})$	
		NN	MAP22
Fixed-target	233	1.08 (0.98 + 0.10)	0.91 (0.70 + 0.21)
RHIC	7	1.11 (1.03 + 0.07)	1.45 (1.37 + 0.08)
Tevatron	71	0.80 (0.73 + 0.06)	1.20 (1.17 + 0.04)
LHCb	21	0.98 (0.88 + 0.10)	1.25 (1.05 + 0.20)
CMS	78	0.40 (0.38 + 0.02)	0.41 (0.35 + 0.06)
ATLAS	72	1.38 (1.09 + 0.29)	3.51 (3.03 + 0.49)
Total	482	0.97 (0.86 + 0.11)	1.28 (1.09 + 0.20)

$$\chi^2 = \sum_i \left(\frac{d_i - \bar{t}_i}{\sigma_{i,\rm unc}} \right)^2 + \sum_\alpha \lambda^2_\alpha = \chi^2_D + \chi^2_\lambda$$

$$\bar{t}_i = t_i + \sum_\alpha \lambda_\alpha \sigma^{(\alpha)}_{i,\rm corr}$$

Neural-Network TMDs

Dataset and fit quality

- Fit to Drell-Yan data only (proof of concept):
  - compare to MAP22 (DY only) to gauge the **impact of parametrisation**.

- Globally, NN performs significantly better that MAP22:
  - significantly lower penalty $\chi^2_\lambda \Rightarrow$ **less reliance on systematic shifts**.

- Dramatic improvement of ATLAS data description:
  - **most precise** data currently available.

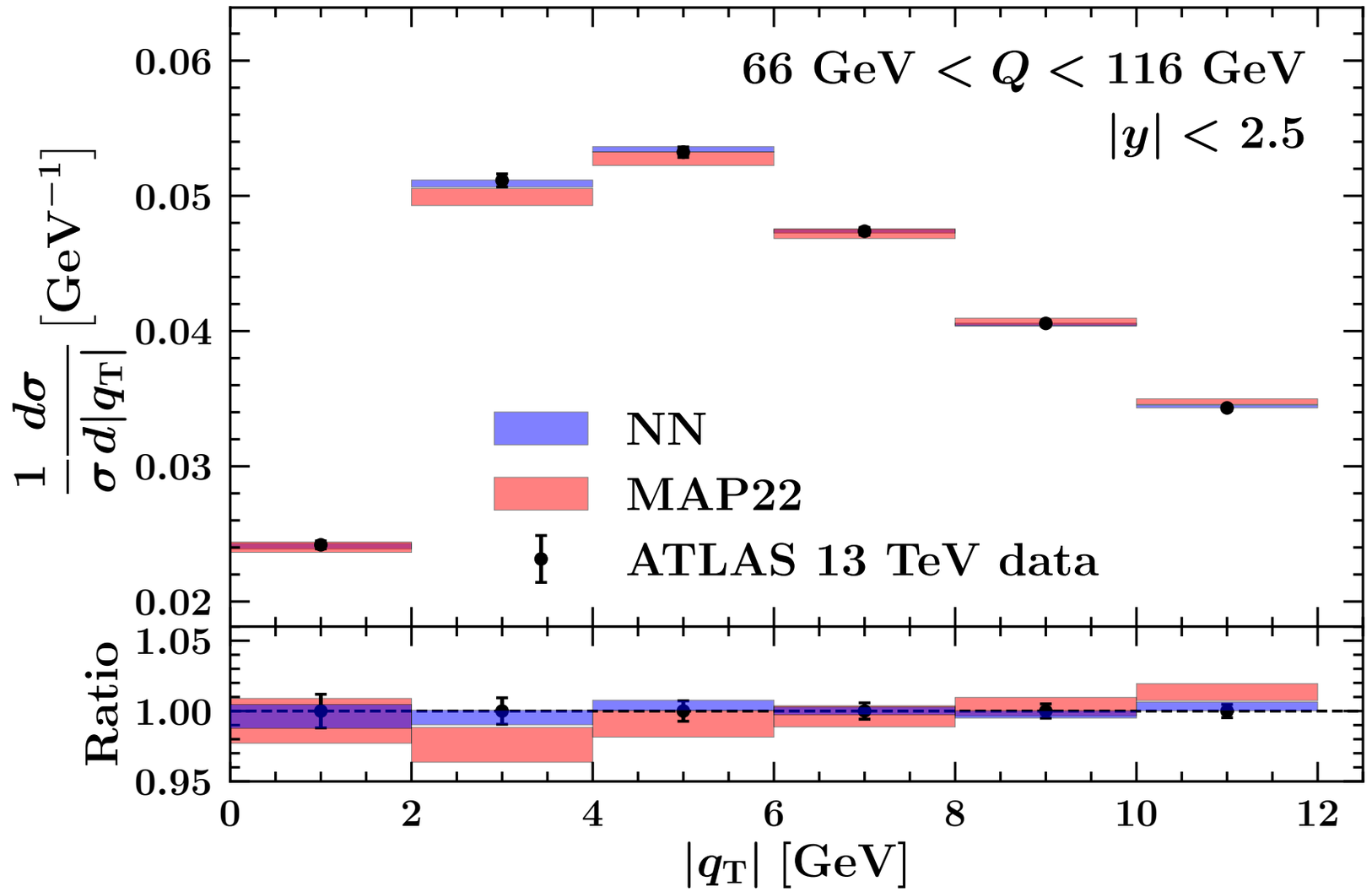
Experiment	$N_{\rm dat}$	$\chi^2/N_{\rm dat}$ ($\bar{\chi}^2_D/N_{\rm dat} + \bar{\chi}^2_\lambda/N_{\rm dat}$)	
		NN	MAP22
Fixed-target	233	1.08 (0.98 + 0.10)	0.91 (0.70 + 0.21)
RHIC	7	1.11 (1.03 + 0.07)	1.45 (1.37 + 0.08)
Tevatron	71	0.80 (0.73 + 0.06)	1.20 (1.17 + 0.04)
LHCb	21	0.98 (0.88 + 0.10)	1.25 (1.05 + 0.20)
CMS	78	0.40 (0.38 + 0.02)	0.41 (0.35 + 0.06)
ATLAS	72	1.38 (1.09 + 0.29)	3.51 (3.03 + 0.49)
Total	482	0.97 (0.86 + 0.11)	1.28 (1.09 + 0.20)

$$\chi^2 = \sum_i \left(\frac{d_i - \bar{t}_i}{\sigma_{i,\rm unc}} \right)^2 + \sum_\alpha \lambda^2_\alpha = \chi^2_D + \chi^2_\lambda$$

$$\bar{t}_i = t_i + \sum_\alpha \lambda_\alpha \sigma^{(\alpha)}_{i,\rm corr}$$

Neural-Network TMDs

Data-theory comparison

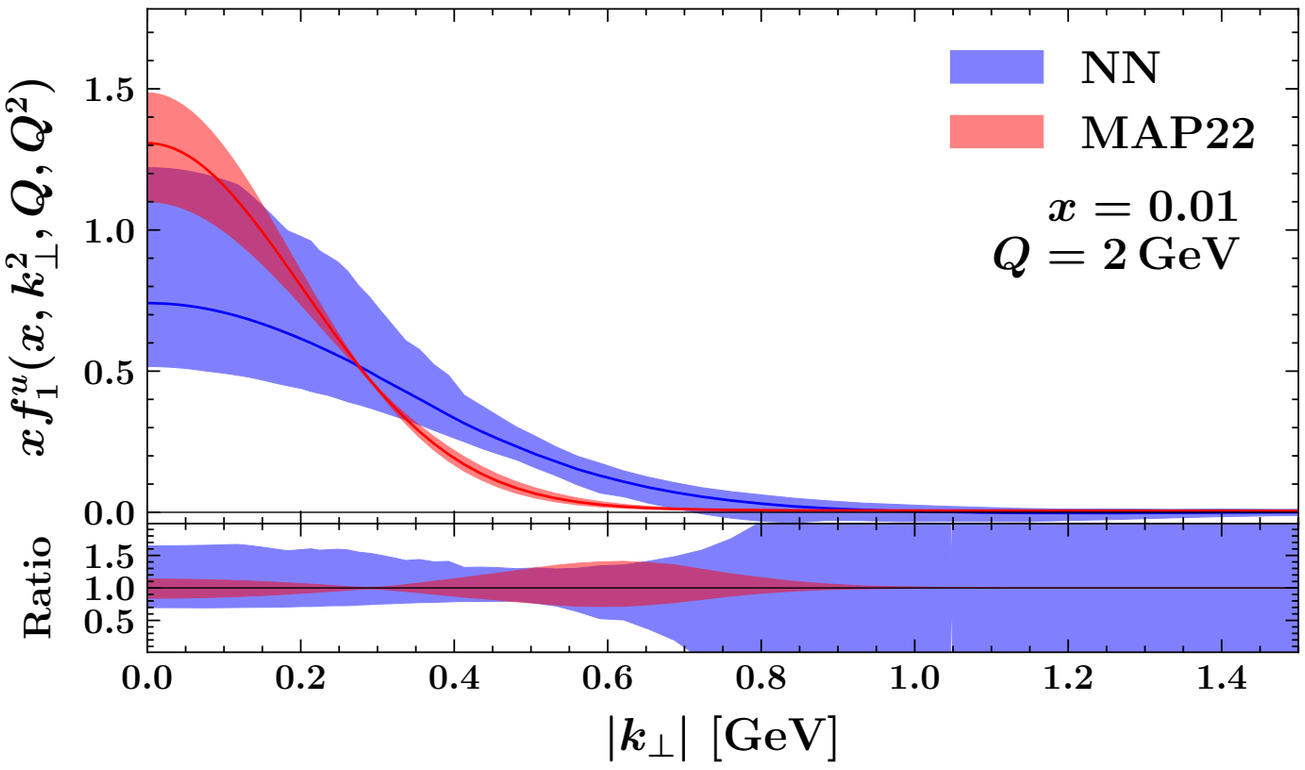


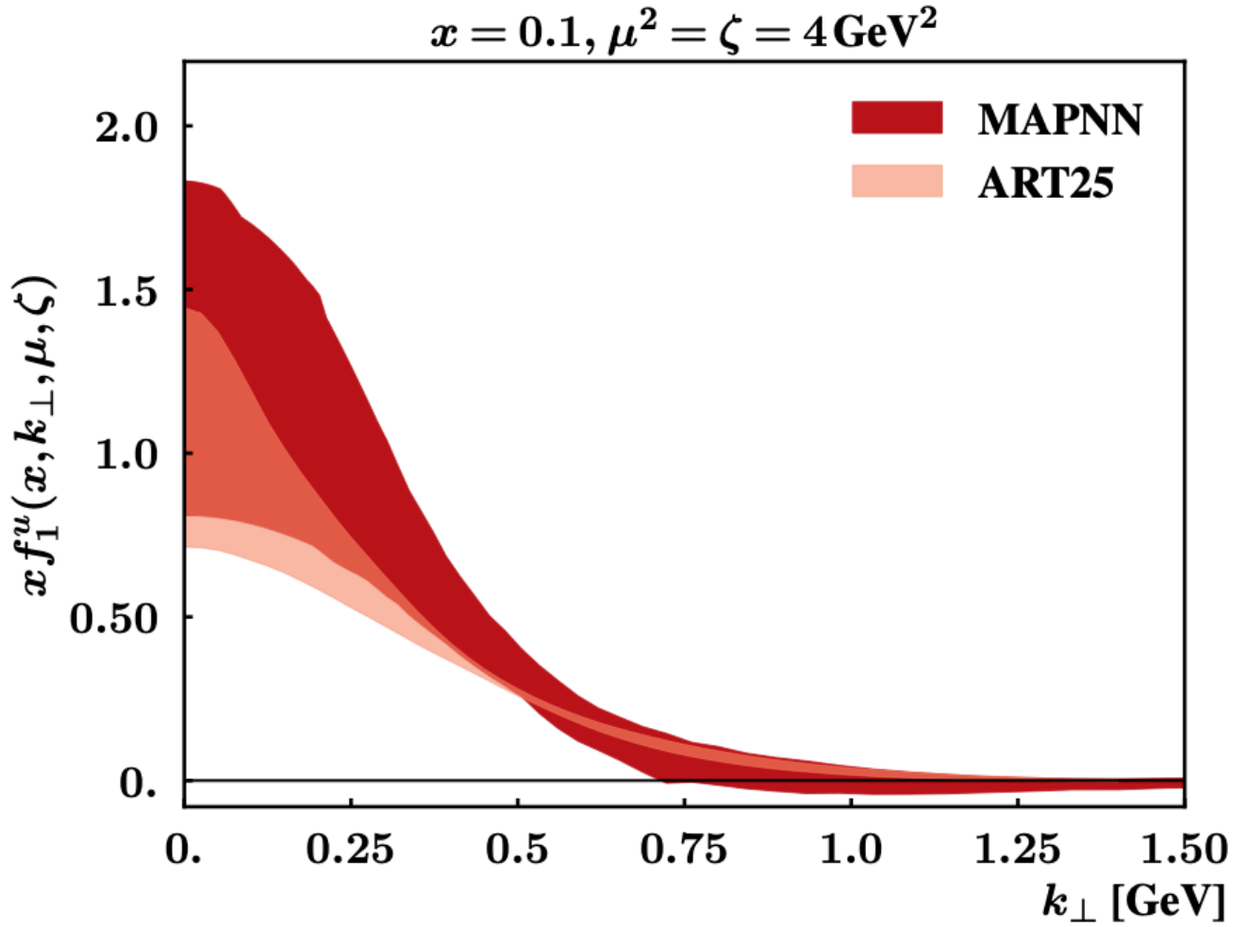
- NN performs significantly better on ATLAS data than MAP22:
  - **shape** is better described by NN,
  - Counterintuitively, NN uncertainties on predictions are smaller than MAP22 ones because the latter are driven by **systematic shifts**.

Neural-Network TMDs

TMD comparison

[C. Lorcé, A. Metz, B. Pasquini, P. Shweitzer, arXiv:2507.12664]



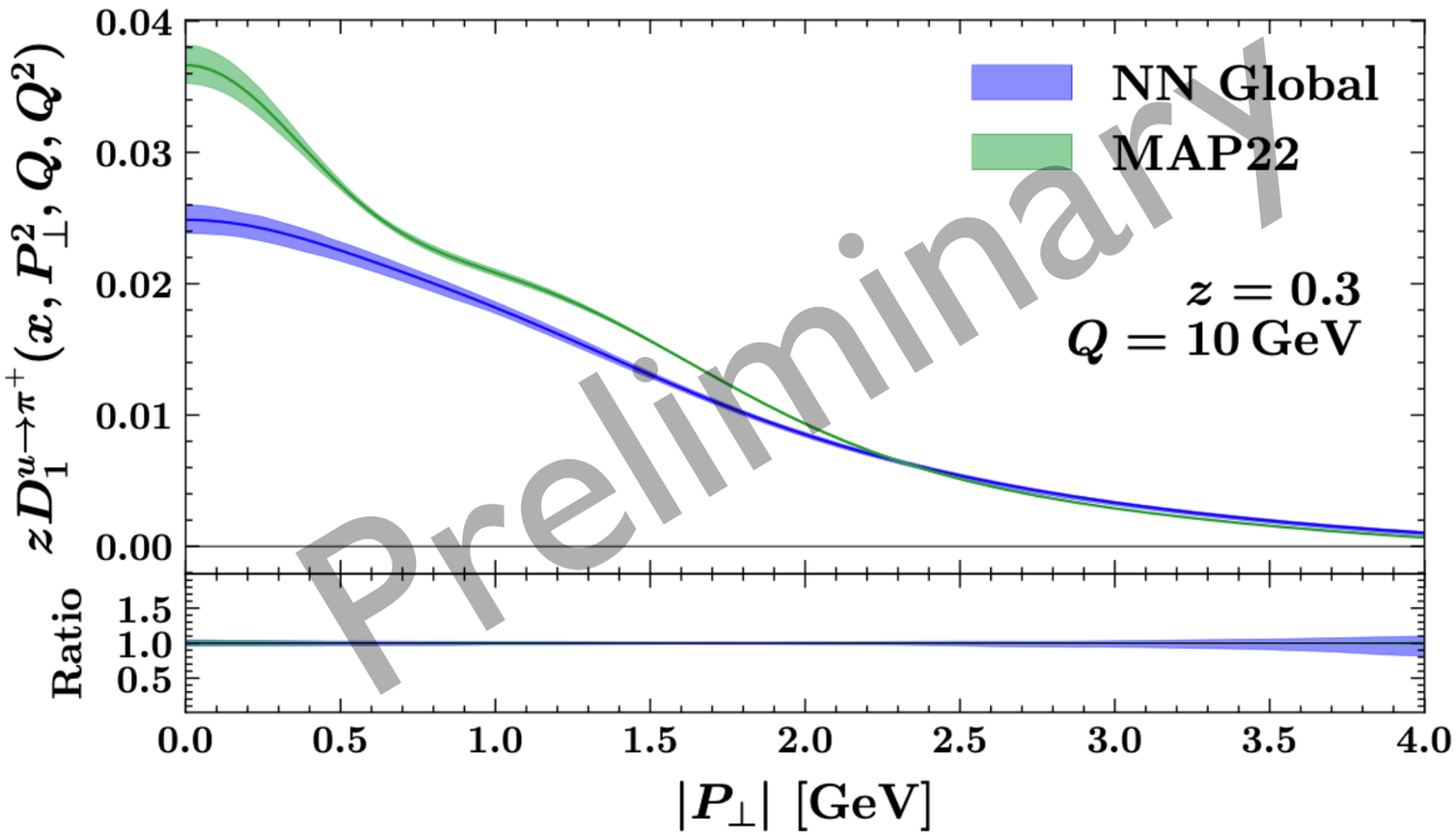


- NN TMDs tend to feature **larger uncertainties** than determinations carried out using "standard" parametrisations (MAP22, ART25):
  - reduction of **functional bias**.
  - Effect particularly evident in the **large-$k_\perp$ region**.

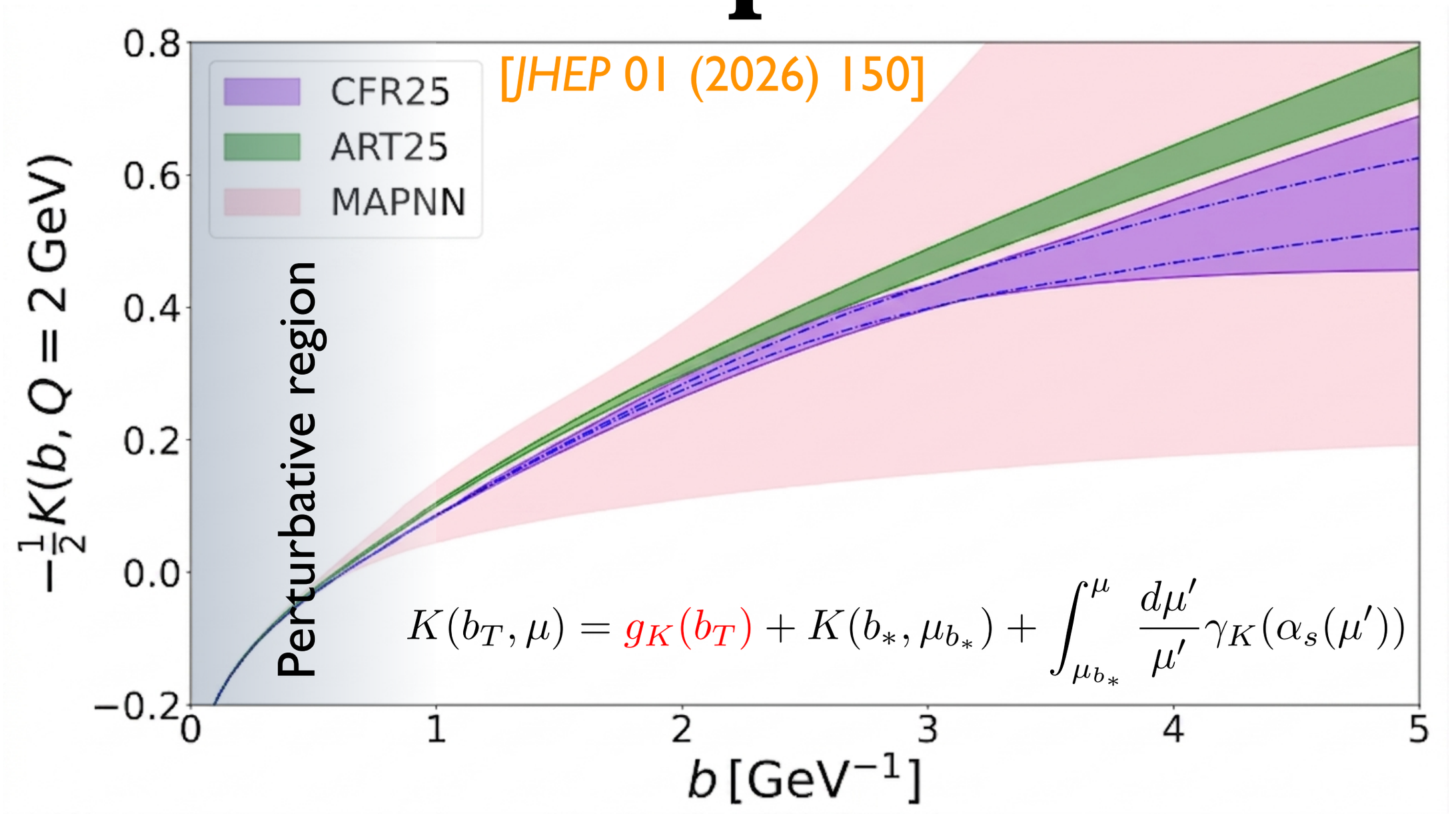
Neural-Network TMDs

Moving to a fully-fledged analysis

- The work in [*PRL*135(2025)2,021904] was a **proof of concept**.
- We are now working on a major analysis based on NNs that will include:
  - **SIDIS data** ⇒ simultaneous extraction of TMD PDFs and FFs.
  - **flavour dependence** ⇒ many more parameters to fit.
  - **Closure tests** ⇒ validation of the robustness of the methodology.



The Collins-Soper kernel



- In the NN TMD we also extracted the Collins-Soper kernel (CSK):
  - in spite of a simple parametrisation ($g_K(b_T) = -g_2^2 b_T^2$), the CSK is affected by large uncertainties ⇒ **small functional bias**.
  - Current (DY) data have **limited constraint power** on the CSK.

The Collins-Soper kernel

The impact of lattice data

- In view of the scarce sensitivity of current experimental data to the CSK, **lattice simulations** become an important resource.

- In [*Phys. Rev. Lett.* 136 (2026) 17, 171902], we included lattice data for the CSK in the NN analysis:
  - lattice data from [*Phys. Rev. Lett.* 132 (2024) 23, 231901].

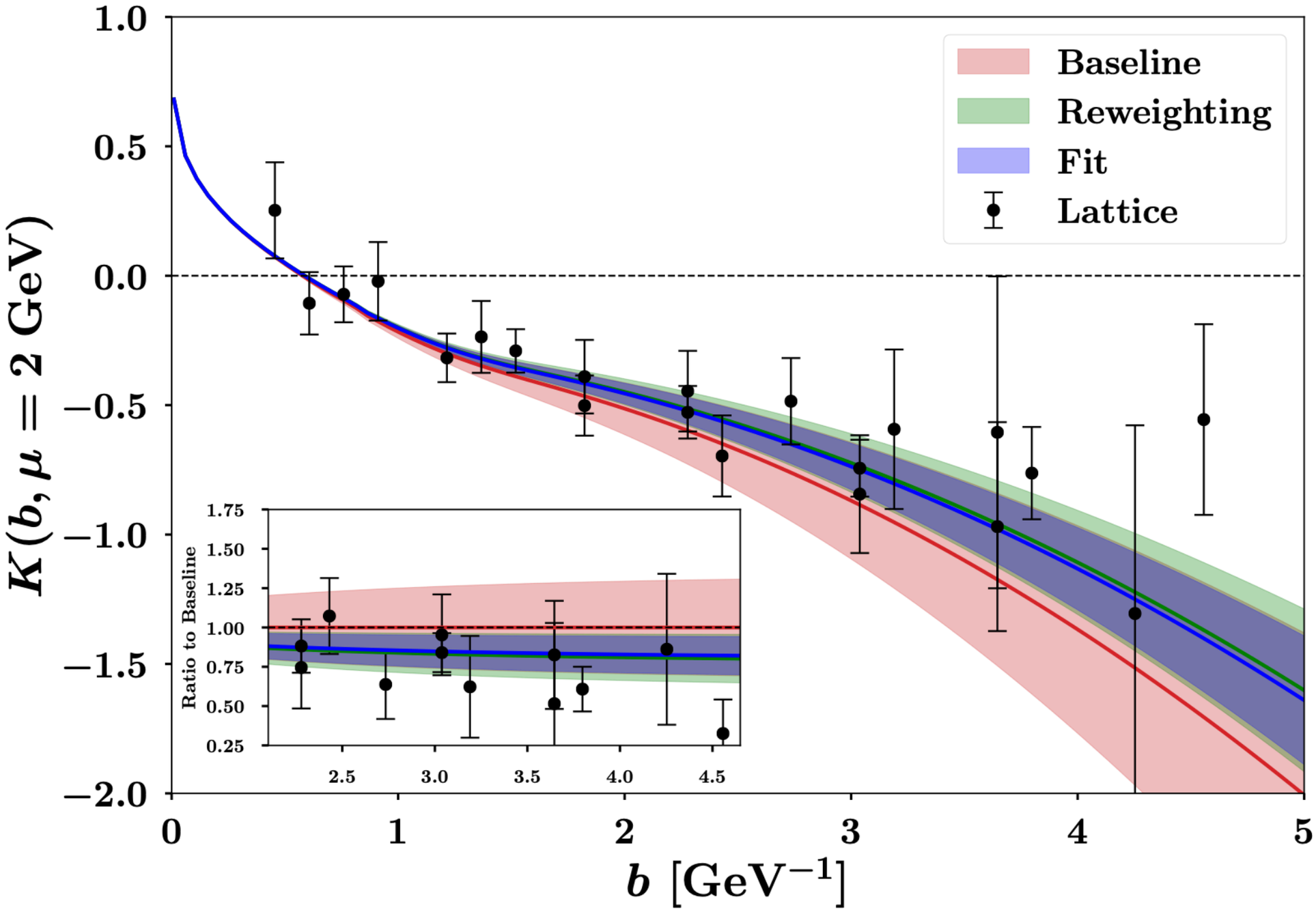
- We followed two different strategies:

  1. Bayesian **reweighting**,

  2. **Refitting**.

- Mutual consistency indicates **robustness**.

The Collins-Soper kernel



- Both strategies consistently produce an upward shift of the central value (~10%) and a **significant reduction of the uncertainty** (~50%).
- Proof of the potential of lattice data on the CSK.

Conclusions

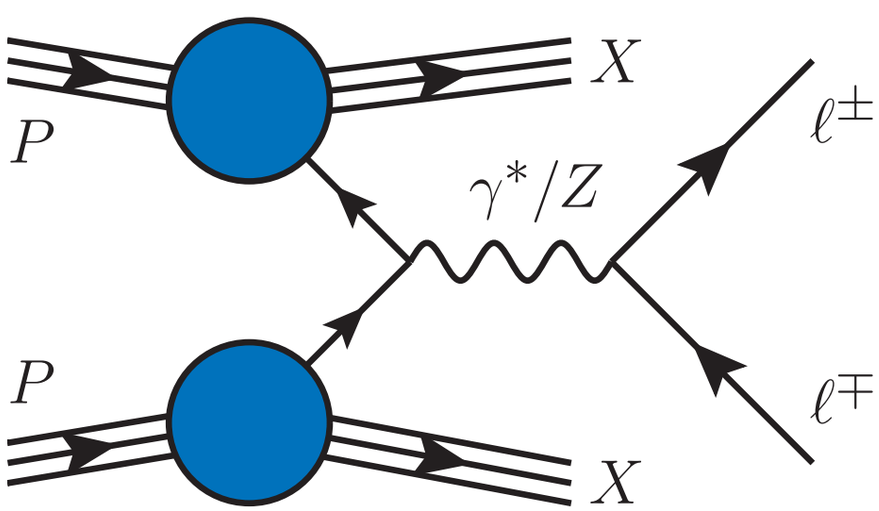
- TMD extractions have become highly demanding:
  - high **perturbative accuracy**,
  - very **precise data** available.
- This led the community to adopt highly **sophisticated fitting methodologies**:
  - complicated parametrisations.
- **Neural networks** are currently being employed to minimise the functional biases:
  - they proved to have a **strong potential**.
  - Only a *proof of concept* exists (DY only, flavour blind) but we are currently working on a **TMD extraction** based on NNs

Backup

Factorising processes (quarks)

- Processes for which leading-power TMD factorisation has been **proven**:

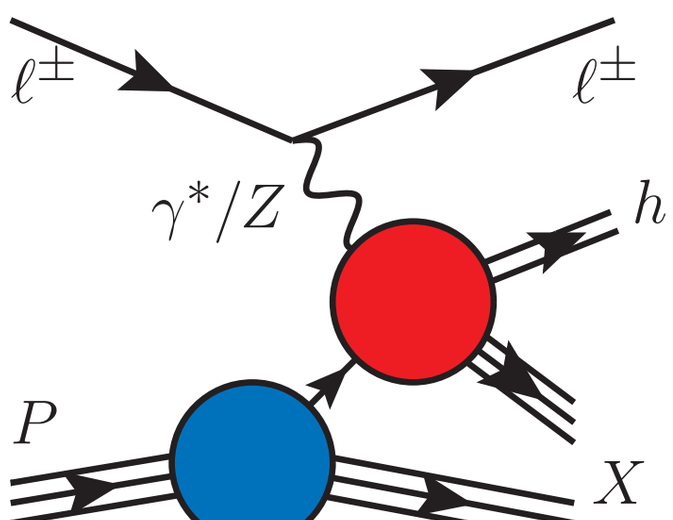
Drell-Yan



$$P P \longrightarrow \ell^\pm \ell^\mp X$$

- **Two** TMD **PDFs**:
- Lots of data:
  - low-energy: FNAL,
  - mid-energy: RHIC,
  - high-energy: Tevatron, LHC.

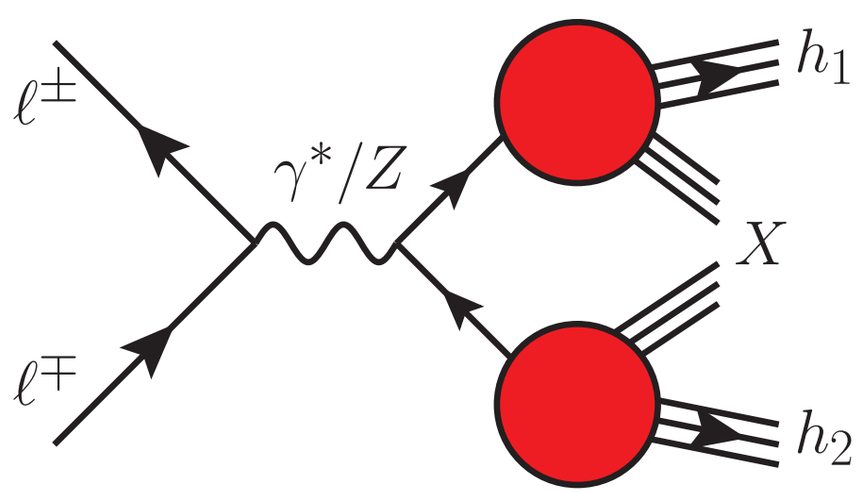
Semi-inclusive DIS



$$P \ell^\pm \longrightarrow \ell^\pm h X$$

- One TMD **PDF** one **FF**:
- many precise data points:
  - HERMES at DESY,
  - COMPASS at CERN.

e^+e^- annihilation



$$\ell^\pm \ell^\mp \to h_1 h_2 X$$

- **Two** TMD **FFs**:
- di-hadron prod. from:
  - BELLE at KEK,
  - BABAR at SLAC.

- Examples of other precesses:
  - thrust and p_{hT} distributions in single-hadron production in e^+e^-,
  - hadron-in-jet production,
  - …

Unpolarised TMD extractions

A selection of fits

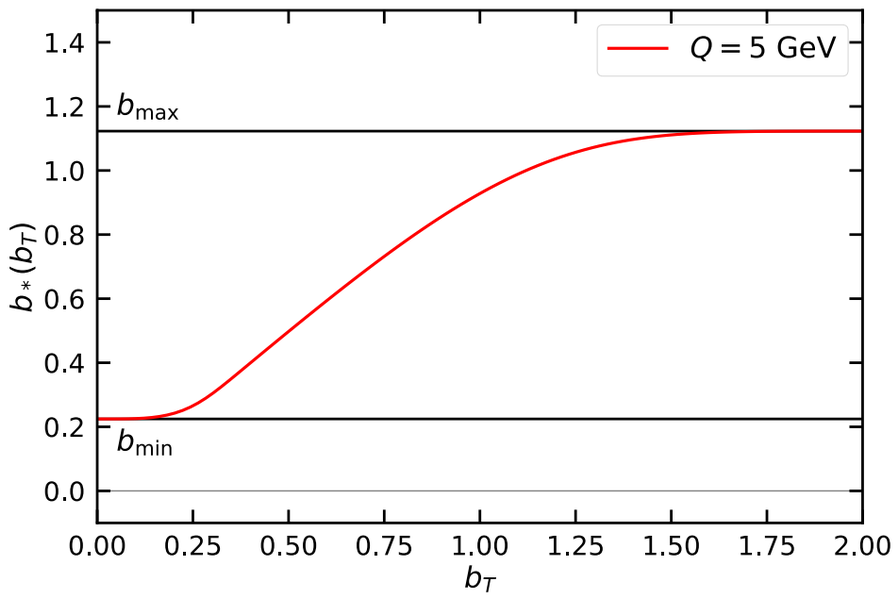
	Accuracy	SIDIS	Drell-Yan	N. of points	Flavour dep.
DWS 1984, CERN-TH.3987/84	NLL	✘	✔	a few	✘
BLNY 2003, hep-ph/0212159	NLL'-NNLL	✘	✔	116	✘
Pavia 2017, 1703.10157	NLL	✔	✔	8059	✘
SV 2017, 1706.01473	N^3LL	✘	✔ (LHC)	309	✘
SV 2019,1912.06532	$N^3LL^{(-)}$	✔	✔ (LHC)	1039	✘
Pavia 2019, 1912.07550	N^3LL	✘	✔ (LHC)	353	✘
MAPTMD22, 2206.07598	$N^3LL^{(-)}$	✔	✔ (LHC)	2031	✘
ART23, 2305.07473	$N^4LL^{(-)}$	✘	✔ (LHC)	627	✔
MAPTMD24, 2405.13833	N^3LL	✔	✔ (LHC)	2031	✔
ART25, 2503.11201	$N^4LL^{(-)}$	✔	✔ (LHC)	1209	✔
Neural Network TMDs (MAP) 2502.04166	N^3LL	✘	✔ (LHC)	482	✘

MAPTMD 2024

Main settings

- b_* prescription:

$$b_*(b_T) = b_{\max}\left(\frac{1-e^{-b_T^4/b_{\max}^4}}{1-e^{-b_T^4/b_{\min}^4}}\right)^{1/4} \quad \text{with} \quad \begin{cases} b_{\max} = 2e^{-\gamma_E} \\ b_{\min} = b_{\max}/Q \end{cases}$$



- Non-perturbative function $f_{\rm NP}$:
  - evolution (CS kernel): $g_K(\boldsymbol{b}_T^2) = -g_2^2\,\frac{\boldsymbol{b}_T^2}{2}$
  - 5 PDFs ($u, \bar{u}, d, \bar{d}, sea$):

$$f_{1\,NP}(x,\boldsymbol{b}_T^2;\zeta,Q_0) = \frac{g_1(x)\,e^{-g_1(x)\frac{\boldsymbol{b}_T^2}{4}} + \lambda^2\,g_{1B}^2(x)\left[1-g_{1B}(x)\frac{\boldsymbol{b}_T^2}{4}\right]e^{-g_{1B}(x)\frac{\boldsymbol{b}_T^2}{4}} + \lambda_2^2\,g_{1C}(x)\,e^{-g_{1C}(x)\frac{\boldsymbol{b}_T^2}{4}}}{g_1(x)+\lambda^2\,g_{1B}^2(x)+\lambda_2^2\,g_{1C}(x)}\left[\frac{\zeta}{Q_0^2}\right]^{g_K(\boldsymbol{b}_T^2)/2}$$

  - 5 FFs (π and K):

$$D_{1\,NP}(z,\boldsymbol{b}_T^2;\zeta,Q_0) = \frac{g_3(z)\,e^{-g_3(z)\frac{\boldsymbol{b}_T^2}{4z^2}} + \frac{\lambda_F}{z^2}\,g_{3B}^2(z)\left[1-g_{3B}(z)\frac{\boldsymbol{b}_T^2}{4z^2}\right]e^{-g_{3B}(z)\frac{\boldsymbol{b}_T^2}{4z^2}}}{g_3(z)+\frac{\lambda_F}{z^2}\,g_{3B}^2(z)}\left[\frac{\zeta}{Q_0^2}\right]^{g_K(\boldsymbol{b}_T^2)/2}$$

$$g_{\{1,1B,1C\}}(x) = N_{\{1,1B,1C\}}\frac{x^{\sigma_{\{1,2,3\}}}(1-x)^{\alpha^2_{\{1,2,3\}}}}{\hat{x}^{\sigma_{\{1,2,3\}}}(1-\hat{x})^{\alpha^2_{\{1,2,3\}}}} \qquad g_{\{3,3B\}}(z) = N_{\{3,3B\}}\frac{(z^{\beta_{\{1,2\}}}+\delta^2_{\{1,2\}})(1-z)^{\gamma^2_{\{1,2\}}}}{(\hat{z}^{\beta_{\{1,2\}}}+\delta^2_{\{1,2\}})(1-\hat{z})^{\gamma^2_{\{1,2\}}}}$$

- **96 free parameters** to fit to data.
- Perturbative accuracies: **N³LL**.
- **Monte Carlo** method for the experimental error propagation.

MAPTMD 2024

Dataset

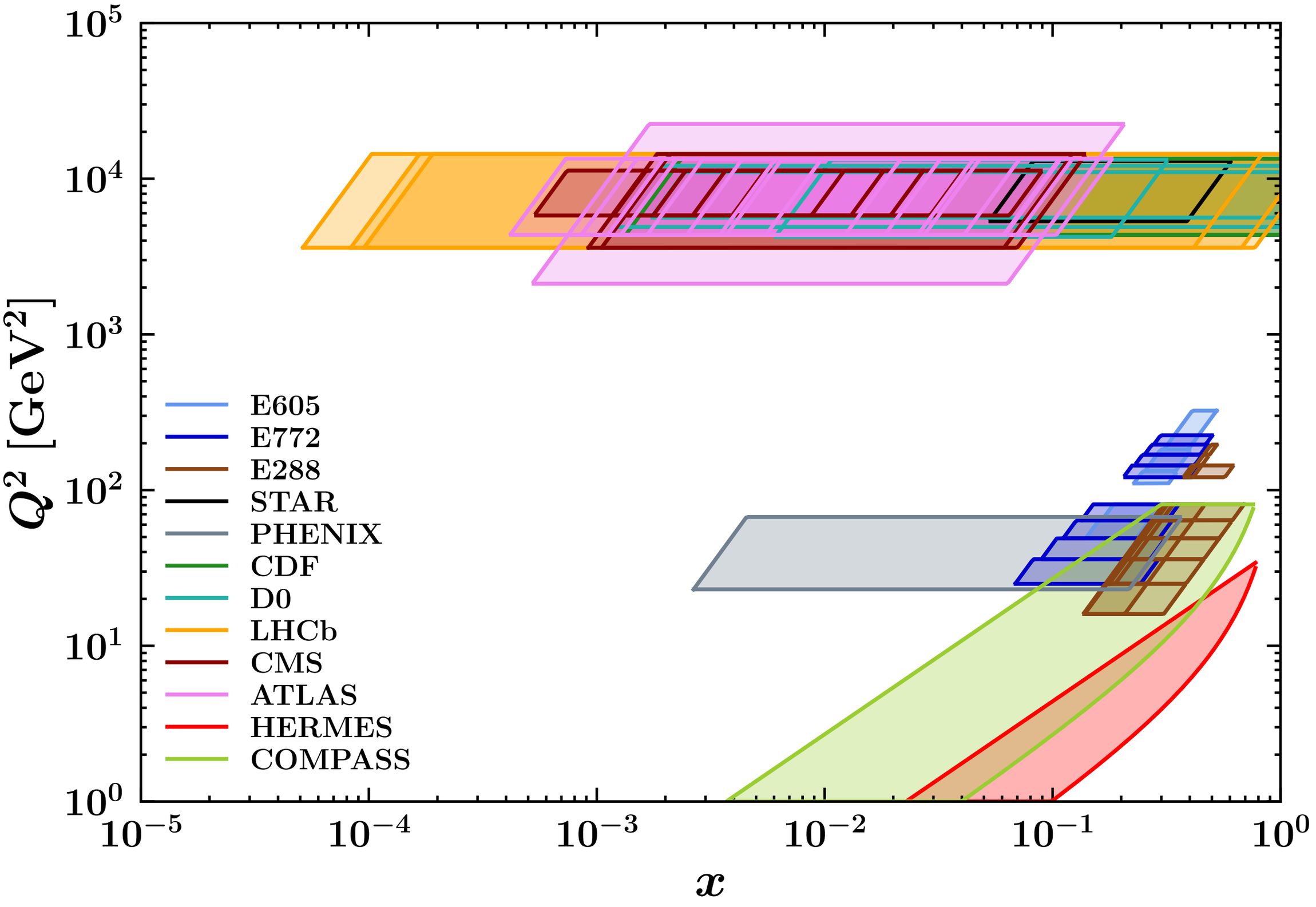
- DY data:
  - fixed-target low-energy DY,
  - RHIC data,
  - LHC and Tevatron data,
  - selection cut q_T / $Q < 0.2$,
  - 484 data points.
- SIDIS data:
  - HERMES and COMPASS,
  - $P_{hT}|_{\max} = \min[\min[0.2Q, 0.5zQ] + 0.3 \text{ GeV}, zQ]$
  - $Q > 1.4$ GeV, $0.2 < z < 0.7$,
  - 1547 points.

Experiment	$N_{\rm dat}$	Observable	$\sqrt{s}$ [GeV]	Q [GeV]	y or x_F	Lepton cuts	Ref.
E605	50	$Ed^3\sigma/d^3\boldsymbol{q}$	38.8	7 - 18	$x_F = 0.1$	-	[55]
E772	53	$Ed^3\sigma/d^3\boldsymbol{q}$	38.8	5 - 15	$0.1 < x_F < 0.3$	-	[51]
E288 200 GeV	30	$Ed^3\sigma/d^3\boldsymbol{q}$	19.4	4 - 9	$y = 0.40$	-	[56]
E288 300 GeV	39	$Ed^3\sigma/d^3\boldsymbol{q}$	23.8	4 - 12	$y = 0.21$	-	[56]
E288 400 GeV	61	$Ed^3\sigma/d^3\boldsymbol{q}$	27.4	5 - 14	$y = 0.03$	-	[56]
STAR 510	7	$d\sigma/d\lvert\boldsymbol{q}_T\rvert$	510	73 - 114	$\lvert y\rvert < 1$	$p_{T\ell} > 25$ GeV $\lvert\eta_\ell\rvert < 1$	-
PHENIX200	2	$d\sigma/d\lvert\boldsymbol{q}_T\rvert$	200	4.8 - 8.2	$1.2 < y < 2.2$	-	[52]
CDF Run I	25	$d\sigma/d\lvert\boldsymbol{q}_T\rvert$	1800	66 - 116	Inclusive	-	[57]
CDF Run II	26	$d\sigma/d\lvert\boldsymbol{q}_T\rvert$	1960	66 - 116	Inclusive	-	[58]
D0 Run I	12	$d\sigma/d\lvert\boldsymbol{q}_T\rvert$	1800	75 - 105	Inclusive	-	[59]
D0 Run II	5	$(1/\sigma)d\sigma/d\lvert\boldsymbol{q}_T\rvert$	1960	70 - 110	Inclusive	-	[60]
D0 Run II (μ)	3	$(1/\sigma)d\sigma/d\lvert\boldsymbol{q}_T\rvert$	1960	65 - 115	$\lvert y\rvert < 1.7$	$p_{T\ell} > 15$ GeV $\lvert\eta_\ell\rvert < 1.7$	[61]
LHCb 7 TeV	7	$d\sigma/d\lvert\boldsymbol{q}_T\rvert$	7000	60 - 120	$2 < y < 4.5$	$p_{T\ell} > 20$ GeV $2 < \eta_\ell < 4.5$	[62]
LHCb 8 TeV	7	$d\sigma/d\lvert\boldsymbol{q}_T\rvert$	8000	60 - 120	$2 < y < 4.5$	$p_{T\ell} > 20$ GeV $2 < \eta_\ell < 4.5$	[63]
LHCb 13 TeV	7	$d\sigma/d\lvert\boldsymbol{q}_T\rvert$	13000	60 - 120	$2 < y < 4.5$	$p_{T\ell} > 20$ GeV $2 < \eta_\ell < 4.5$	[64]
CMS 7 TeV	4	$(1/\sigma)d\sigma/d\lvert\boldsymbol{q}_T\rvert$	7000	60 - 120	$\lvert y\rvert < 2.1$	$p_{T\ell} > 20$ GeV $\lvert\eta_\ell\rvert < 2.1$	[65]
CMS 8 TeV	4	$(1/\sigma)d\sigma/d\lvert\boldsymbol{q}_T\rvert$	8000	60 - 120	$\lvert y\rvert < 2.1$	$p_{T\ell} > 15$ GeV $\lvert\eta_\ell\rvert < 2.1$	[66]
CMS 13 TeV	70	$d\sigma/d\lvert\boldsymbol{q}_T\rvert$	13000	76 - 106	$\lvert y\rvert < 0.4$ $0.4 < \lvert y\rvert < 0.8$ $0.8 < \lvert y\rvert < 1.2$ $1.2 < \lvert y\rvert < 1.6$ $1.6 < \lvert y\rvert < 2.4$	$p_{T\ell} > 25$ GeV $\lvert\eta_\ell\rvert < 2.4$	[53]
ATLAS 7 TeV	6 6 6	$(1/\sigma)d\sigma/d\lvert\boldsymbol{q}_T\rvert$	7000	66 - 116	$\lvert y\rvert < 1$ $1 < \lvert y\rvert < 2$ $2 < \lvert y\rvert < 2.4$	$p_{T\ell} > 20$ GeV $\lvert\eta_\ell\rvert < 2.4$	[67]
ATLAS 8 TeV on-peak	6 6 6 6 6 6	$(1/\sigma)d\sigma/d\lvert\boldsymbol{q}_T\rvert$	8000	66 - 116	$\lvert y\rvert < 0.4$ $0.4 < \lvert y\rvert < 0.8$ $0.8 < \lvert y\rvert < 1.2$ $1.2 < \lvert y\rvert < 1.6$ $1.6 < \lvert y\rvert < 2$ $2 < \lvert y\rvert < 2.4$	$p_{T\ell} > 20$ GeV $\lvert\eta_\ell\rvert < 2.4$	[68]
ATLAS 8 TeV off-peak	4 8	$(1/\sigma)d\sigma/d\lvert\boldsymbol{q}_T\rvert$	8000	46 - 66 116 - 150	$\lvert y\rvert < 2.4$	$p_{T\ell} > 20$ GeV $\lvert\eta_\ell\rvert < 2.4$	[68]
ATLAS 13 TeV	6	$(1/\sigma)d\sigma/d\lvert\boldsymbol{q}_T\rvert$	13000	66 - 113	$\lvert y\rvert < 2.5$	$p_{T\ell} > 27$ GeV $\lvert\eta_\ell\rvert < 2.5$	[54]
Total	484						

Experiment	$N_{\rm dat}$	Observable	Channels	Q [GeV]	x	z	Phase space cuts	Ref.
HERMES	344	$M(x, z, \lvert\boldsymbol{P}_{hT}\rvert, Q)$	$p \to \pi^+$ $p \to \pi^-$ $p \to K^+$ $p \to K^-$ $d \to \pi^+$ $d \to \pi^-$ $d \to K^+$ $d \to K^-$	1 - $\sqrt{15}$	$0.023 < x < 0.6$ (6 bins)	$0.1 < z < 1.1$ (8 bins)	$W^2 > 10$ GeV^2 $0.1 < y < 0.85$	[46]
COMPASS	1203	$M(x, z, \boldsymbol{P}_{hT}^2, Q)$	$d \to h^+$ $d \to h^-$	1 - 9 (5 bins)	$0.003 < x < 0.4$ (8 bins)	$0.2 < z < 0.8$ (4 bins)	$W^2 > 25$ GeV^2 $0.1 < y < 0.9$	[72]
Total	1547							

MAPTMD 2024

Kinematic coverage



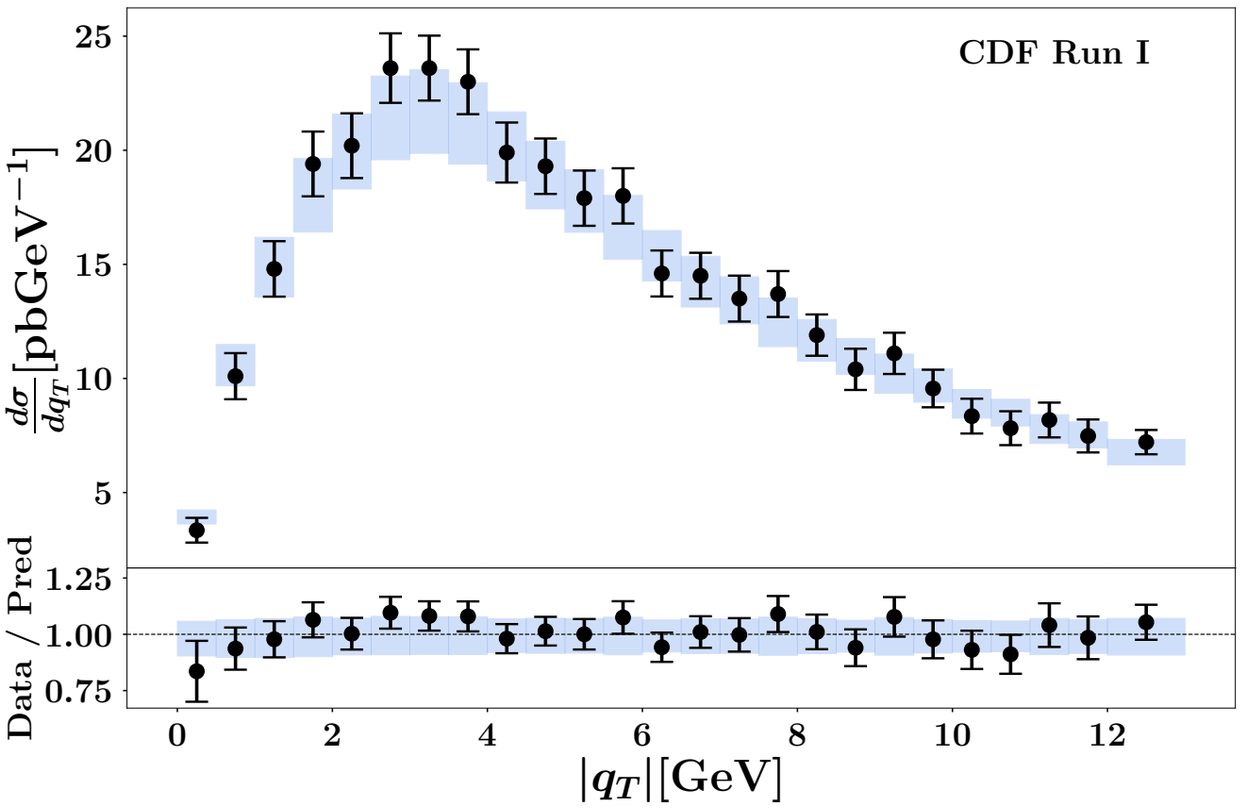
MAPTMD 2024

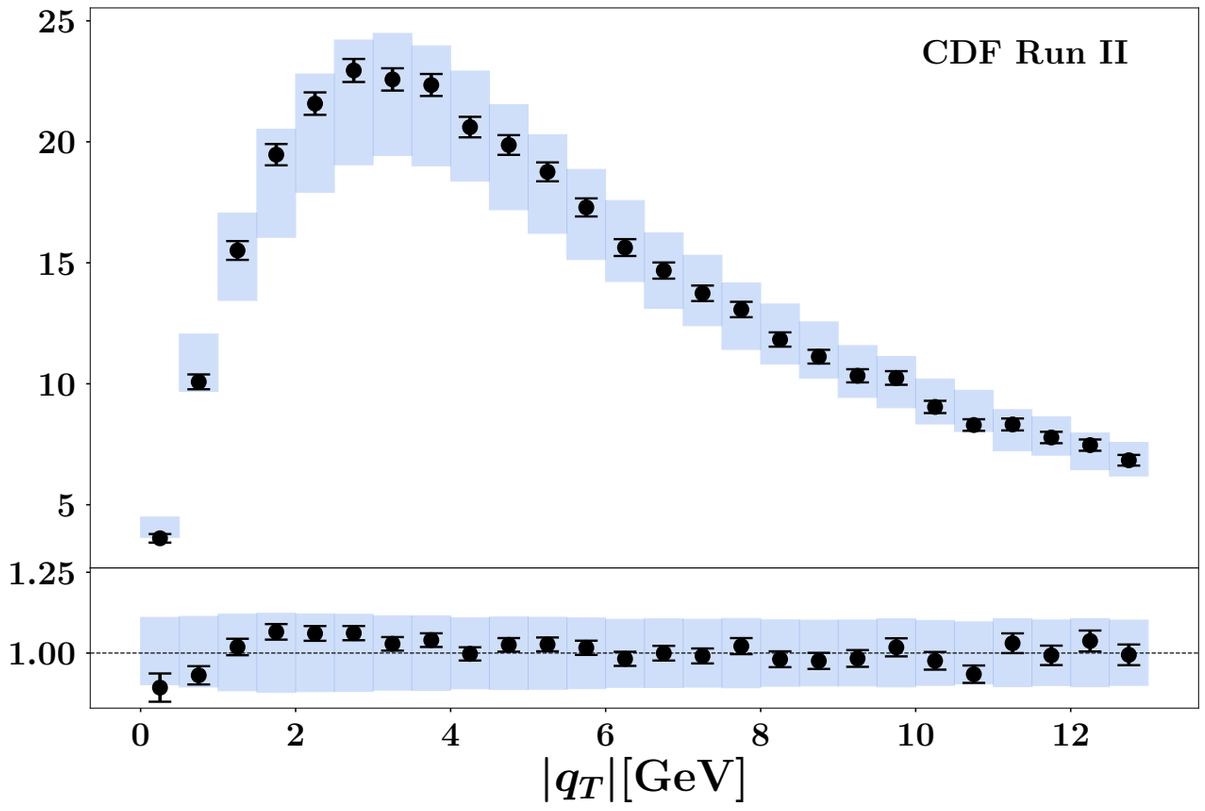
Fit quality

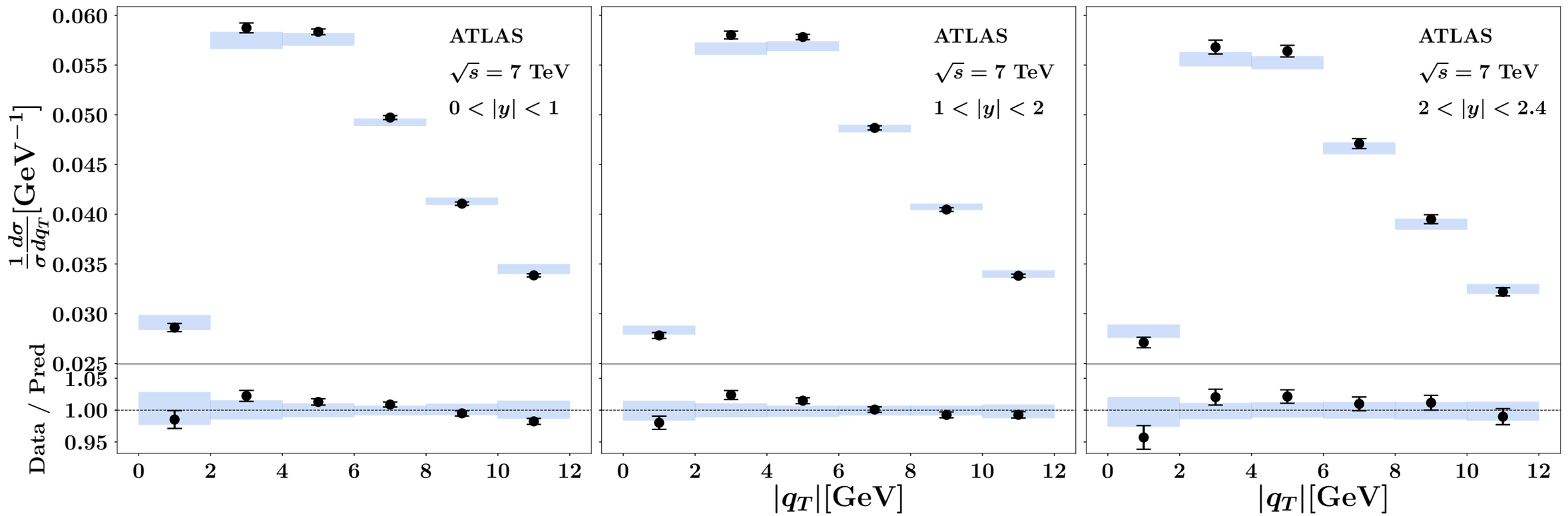
	N^{3}LL			
Data set	$N_{\rm dat}$	χ^2_D	χ^2_λ	χ^2_0
Tevatron total	71	1.10	0.07	1.17
LHCb total	21	3.56	0.96	4.52
ATLAS total	72	3.54	0.82	4.36
CMS total	78	0.38	0.05	0.43
PHENIX 200	2	2.76	1.04	3.80
STAR 510	7	1.12	0.26	1.38
DY collider total	251	1.37	0.28	1.65
E288 200 GeV	30	0.13	0.40	0.53
E288 300 GeV	39	0.16	0.26	0.42
E288 400 GeV	61	0.11	0.08	0.19
E772	53	0.88	0.20	1.08
E605	50	0.70	0.22	0.92
DY fixed-target total	233	0.63	0.31	0.94
DY total	484	1.02	0.29	1.31
HERMES total	344	0.81	0.24	1.05
COMPASS total	1203	0.67	0.27	0.94
SIDIS total	1547	0.70	0.26	0.96
Total	**2031**	**0.81**	**0.27**	**1.08**

MAPTMD 2024

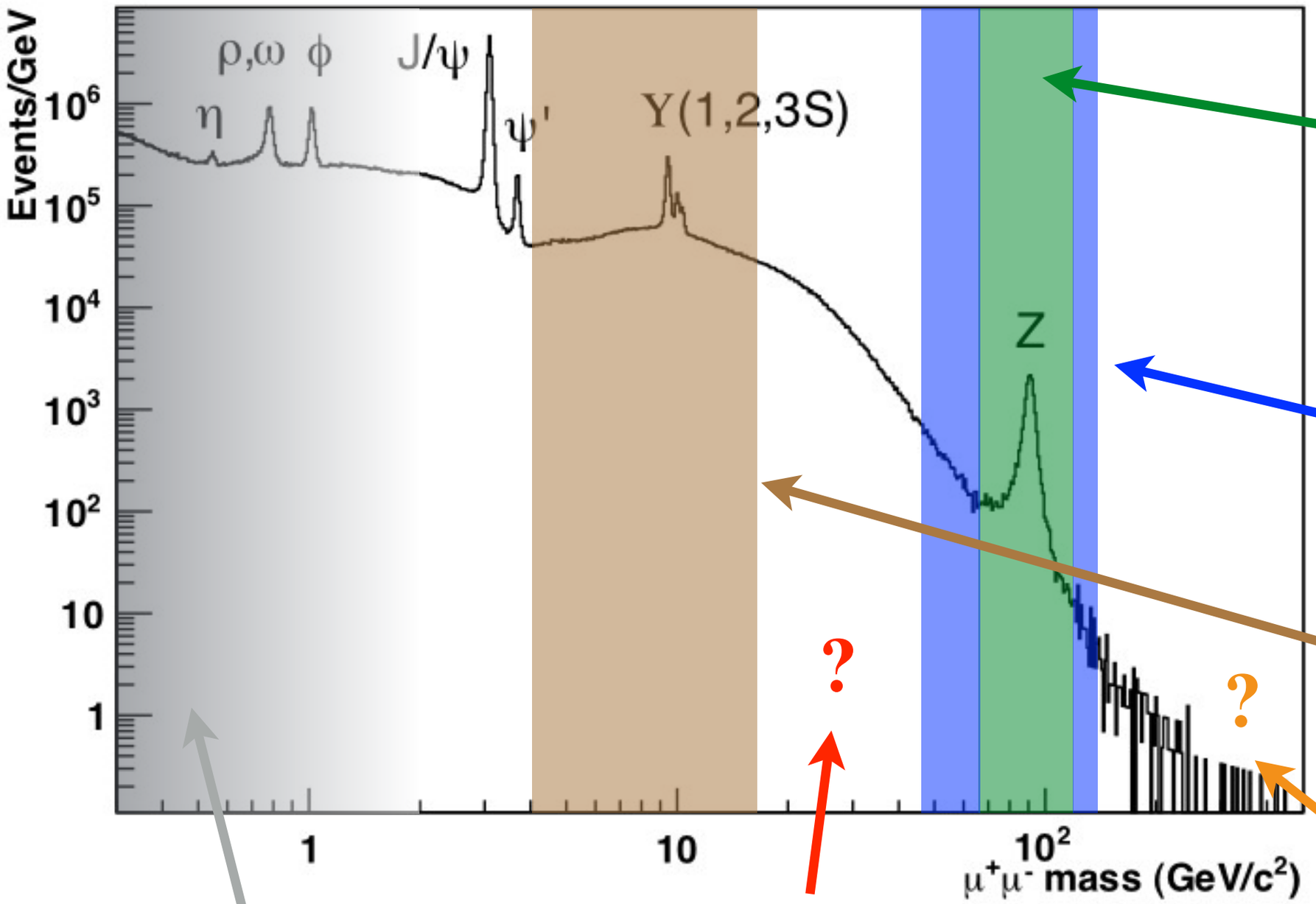
Fit quality: DY







Q^2 data coverage



Precise data from the LHC and Tevatron

Less precise data from the LHC that extends to low q_T

Some old fixed-target data

Perturbation theory breakdown and resonances

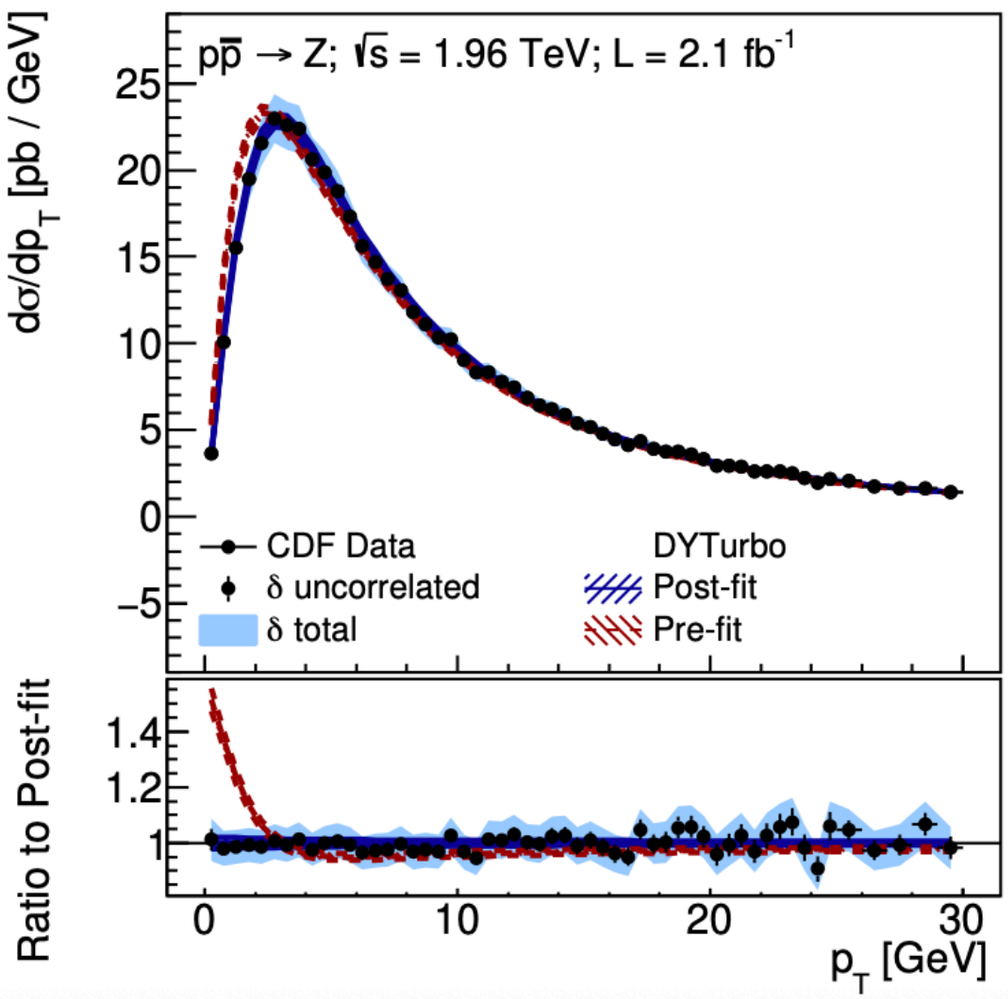
No low-q_T data in this region

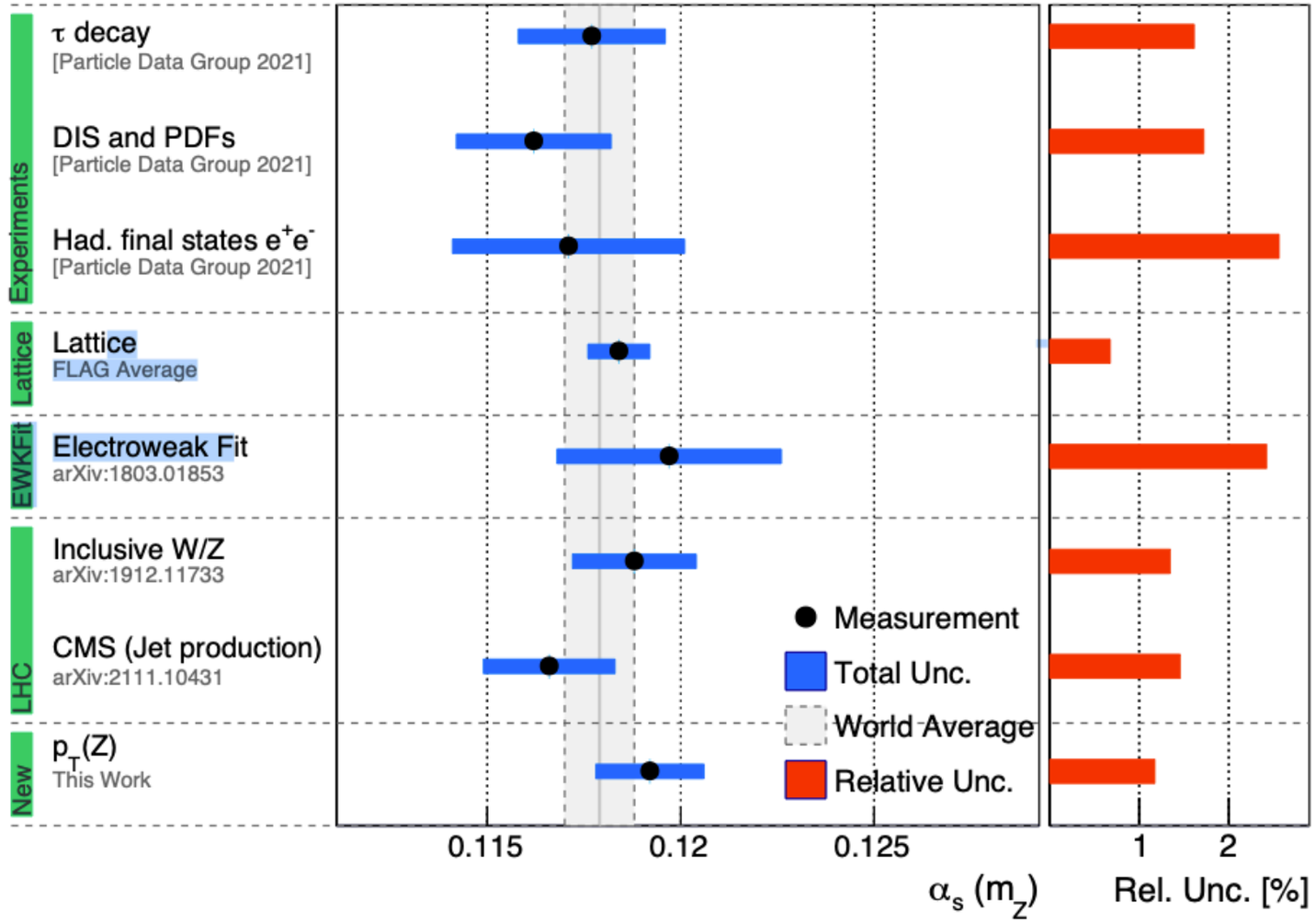
$q_T \simeq 0$ hardly attainable here with good precision

TMDs at the LHC

The strong coupling

- A relatively recent paper [2203.05394], has recently demonstrated the *strong* sensitivity of the low-q_T Drell-Yan spectrum to α_s ($\alpha_s(M_Z) = 0.1191^{+0.0013}_{-0.0016}$).

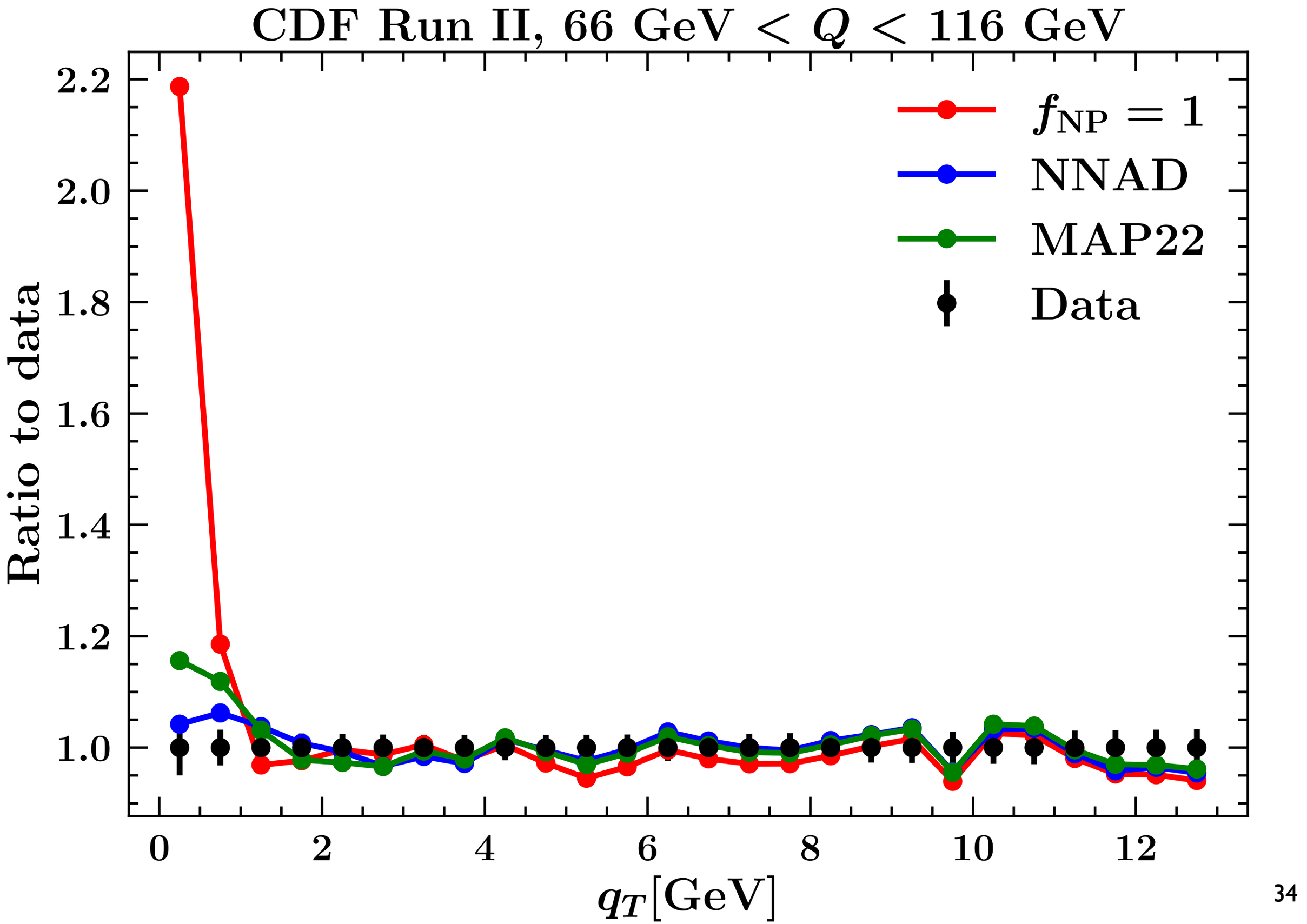




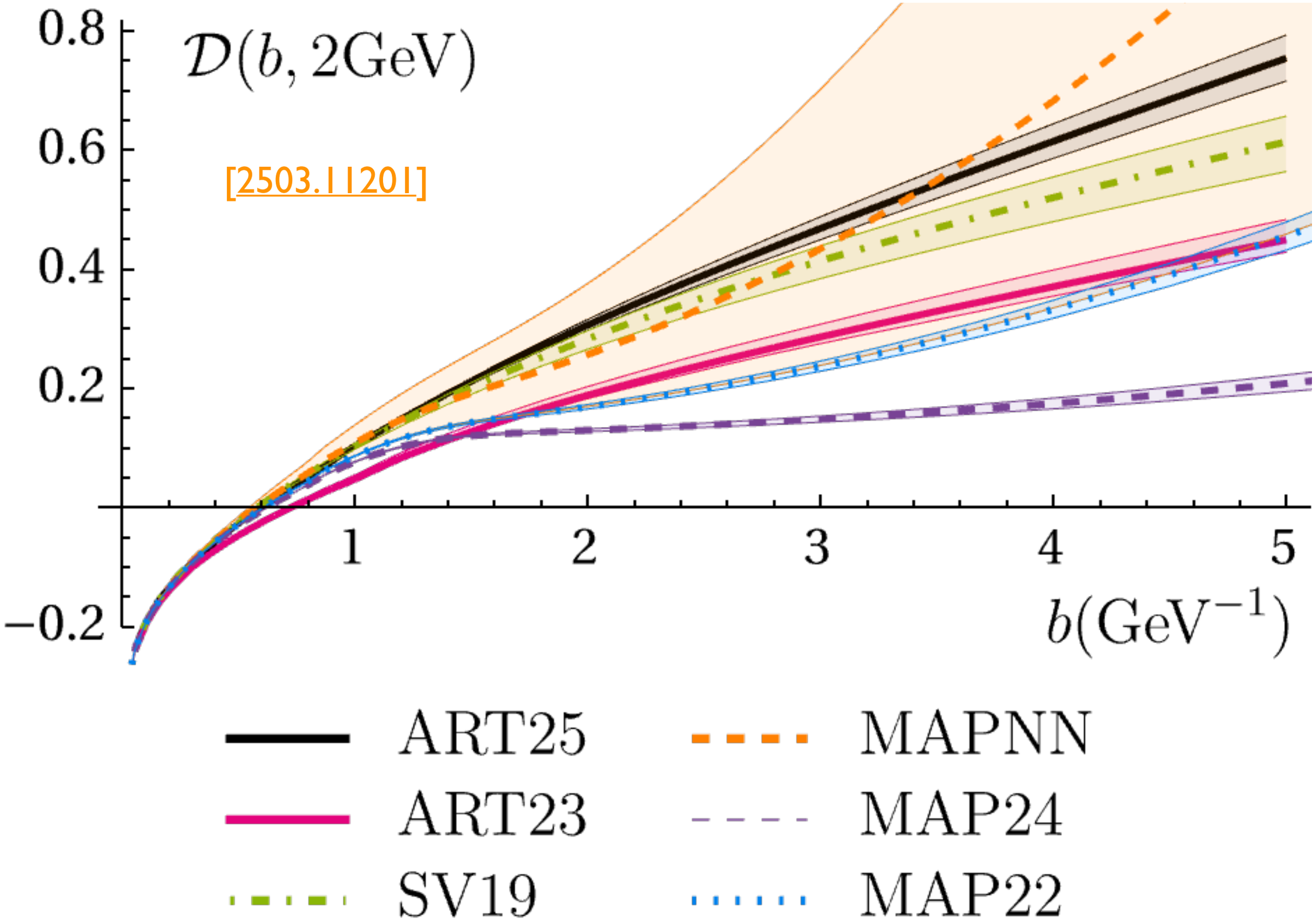
- Importantly, this work has required a determination of the non-perturbative component of TMDs.

- Work based of Tevatron (CDF) data ⇒ huge **potential of the LHC.**

On the importance of TMDs



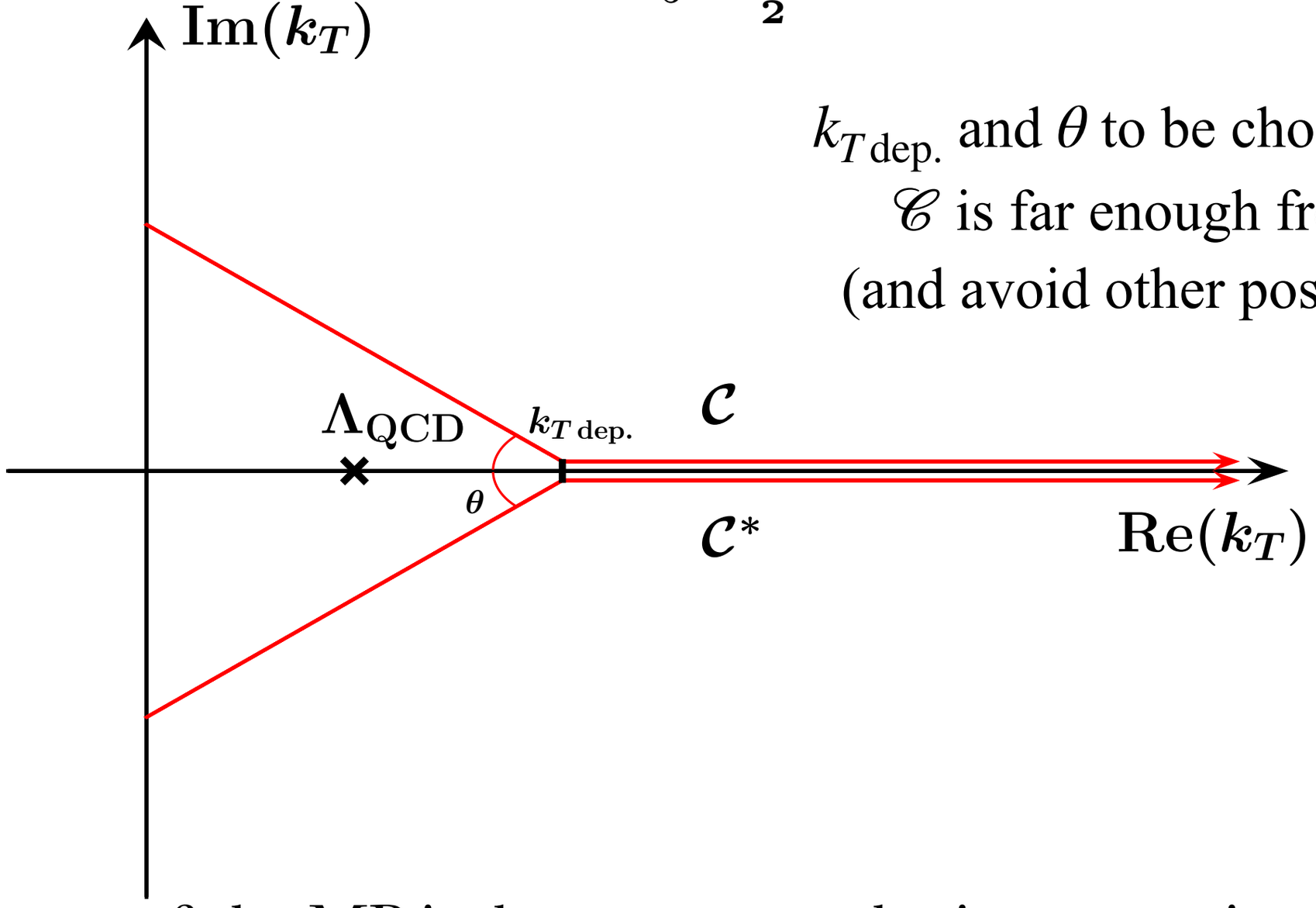
Collis-Soper kernel



Origin of $\mathcal{O}(\Lambda_{\rm QCD}/q_T)$ corrections

2) The **Minimal Prescription**:

$$\int_0^{Q} d\boldsymbol{k}_{\mathbf{T}}\, \alpha_s^p\left(\boldsymbol{k}_{\mathbf{T}}\right)\cdots \rightarrow \int_{\frac{C+C^*}{2}} d\boldsymbol{k}_{\mathbf{T}}\, \alpha_s^p\left(\boldsymbol{k}_{\mathbf{T}}\right)\ldots$$



$k_{T\,\text{dep.}}$ and θ to be chosen such that $\mathscr{C}$ is far enough from $\Lambda_{\rm QCD}$ (and avoid other possible poles)

The advantage of the MP is that non-perturbative corrections scale as $\exp[-\beta q_T/\Lambda_{\rm QCD}]$, relegating them to smaller values of q_T.