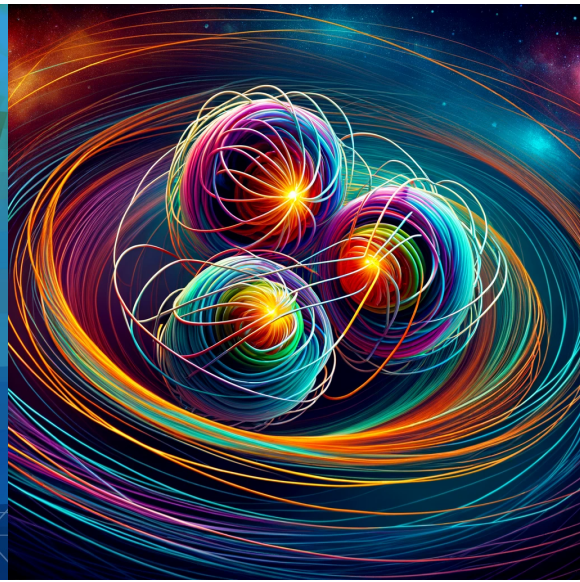
May 12, 2026

Neural Network extractions of unpolarized TMDs



Chiara Bissolotti
Maria Goeppert Mayer Fellow
Argonne National Laboratory
PHY Division

based on:

Phys. Rev. Lett. 135 (2025) 2, 021904
DOI: 10.1103/csc2-bj91

Map Collaboration

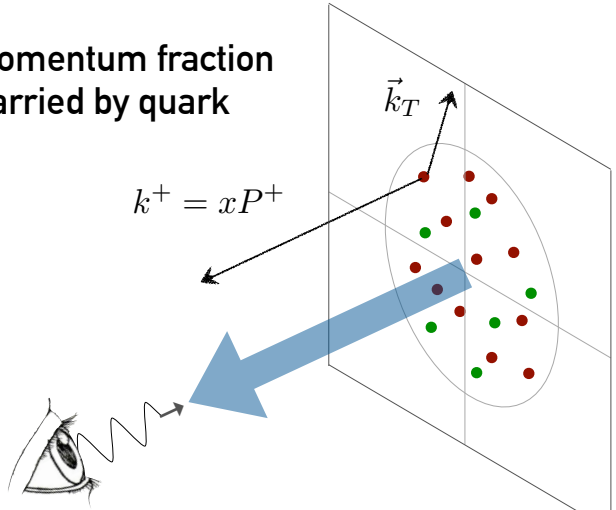
..and some work in progress with Valerio Bertone, Matteo Cerutti, Simone Rodini

Argonne NATIONAL LABORATORY

TMDs: 3D maps

in momentum space

x: momentum fraction carried by quark

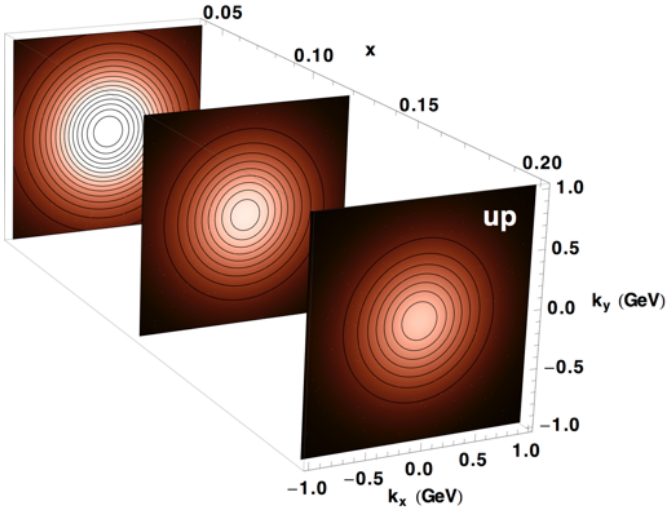




Transverse Momentum Distributions

TMDs

$$f^q(x, \mathbf{k_T})$$



Transverse Momentum Distributions

the distribution of quarks sharply depends on the orientation of their spins

Leading Quark TMDPDFs

Nucleon Spin Quark Spin

Nucleon Polarization	Quark Polarization: Un-Polarized (U)	Longitudinally Polarized (L)	Transversely Polarized (T)
U	f_1 = Unpolarized		$h_1^\perp$ = Boer-Mulders
L		g_1 = Helicity	$h_{1L}^\perp$ = Worm-gear
T	$f_{1T}^\perp$ = Sivers	$g_{1T}^\perp$ = Worm-gear	h_1 = Transversity; $h_{1T}^\perp$ = Pretzelosity

Leading Quark TMDFFs

Hadron Spin Quark Spin

	Quark Polarization: Un-Polarized (U)	Longitudinally Polarized (L)	Transversely Polarized (T)
Unpolarized (or Spin 0) Hadrons	D_1 = Unpolarized		$H_1^\perp$ = Collins
Polarized Hadrons L		G_1 = Helicity	$H_{1L}^\perp$ =
Polarized Hadrons T	$D_{1T}^\perp$ = Polarizing FF	$G_{1T}^\perp$ =	H_1 = Transversity; $H_{1T}^\perp$ =

from the TMD handbook, arXiv:2304.03302

this presentation

TMD

every TMD has the same general structure

matching to the collinear region

collinear PDFs

$$f_1^q(x, b; \mu, \zeta) = \sum_j \left(C_j \otimes f^j\right)(x, b_*; \mu_b) e^{R(b_*; \mu_b, \mu)} f_{\rm NP}(x, b)$$

double scale dependence

perturbative evolution

non perturbative transverse content

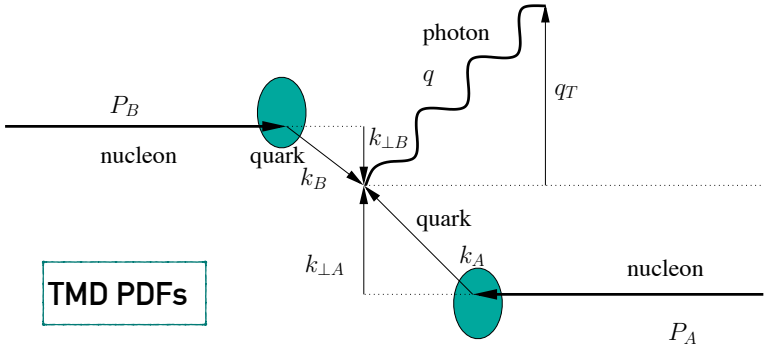
many subtleties involved in TMD analyses

b* prescription, ζ-prescription, logarithmic accuracy

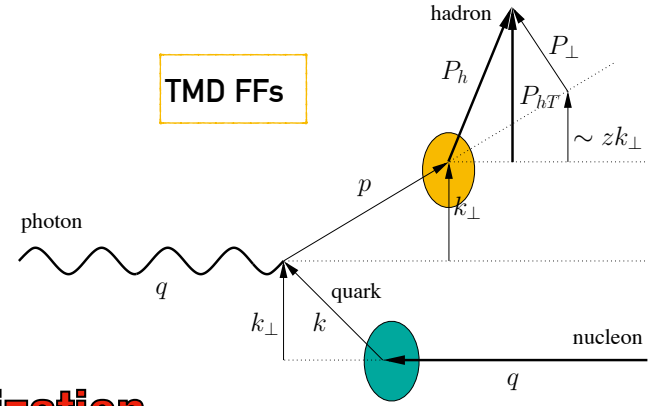
parametrized and fitted to data

Drell-Yan and SIDIS

$$N(P_A) + N(P_B) \to \gamma^*/Z \to l^+ l^-$$



$$\ell(l) + N(p) \to \ell(l') + h(P_h) + X$$



TMD factorization

$$q_T \ll Q$$

$$\left(\frac{d\sigma}{dq_T}\right) \propto \int \frac{d^2\mathbf{b}}{4\pi} e^{i\mathbf{b}\cdot\mathbf{q}_T} x_1 f_1^q(x_1, \mathbf{b}) x_2 f_1^{\bar{q}}(x_2, \mathbf{b})$$

$$M^2 \ll Q^2 \qquad P_{hT}^2 \ll Q^2$$

$$\left(\frac{d\sigma}{dq_T}\right) \propto \int \frac{d^2\mathbf{b}}{4\pi} e^{i\mathbf{b}\cdot\mathbf{q}_T} f_1^q(x, \mathbf{b}) D_1^{q\to h}(z, \mathbf{b})$$

Motivation

What would we like to accomplish?

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

main goal: **fit for the first time unpolarized TMDs with NN**

proof of concept ⇒ we fitted only Drell-Yan data

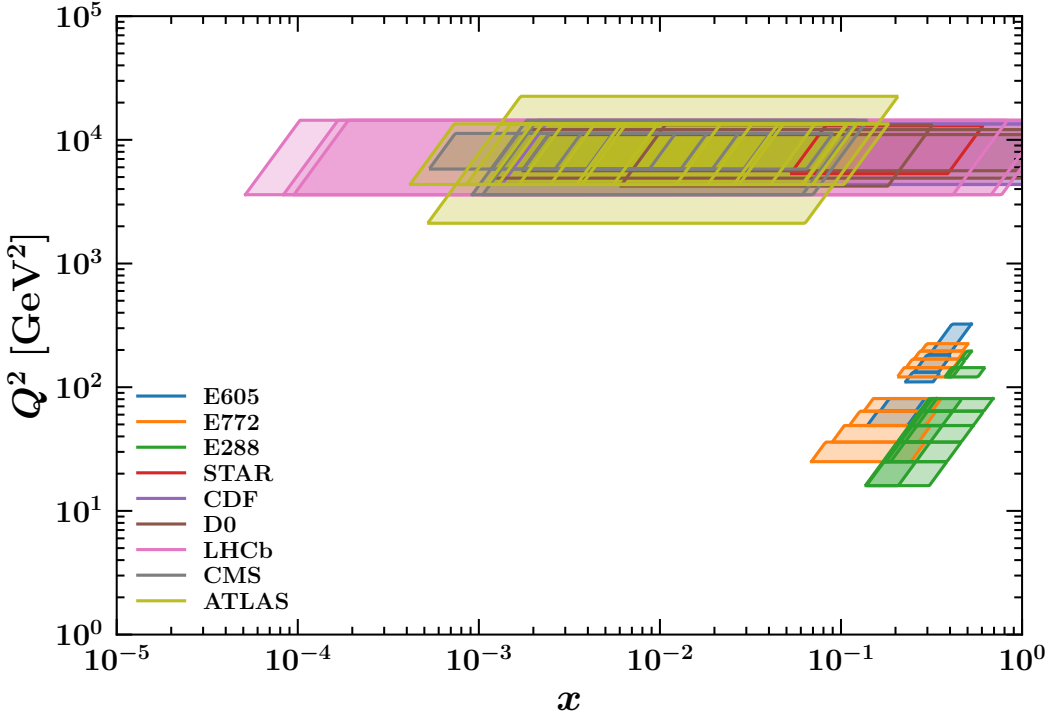
questions:

- NN have been used for years in the context of PDFs, to model the $x \in [0,1]$ dependence. Is this applicable also for TMDs, where $x \in [0,1]$ and $b_T \in (0,\infty)$?

- do NN offer an advantage in fitting (real) data?

Data kinematical coverage

high and low energy Drell-Yan data



TMD region

$$q_T/Q < 0.2$$

total number of data: 482

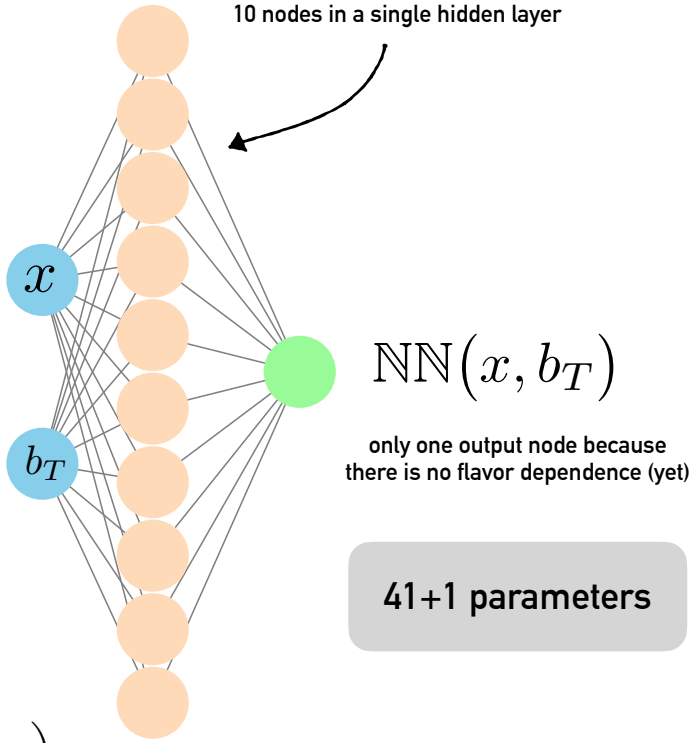
included all available DY data with the exception of PHENIX (2 data points)

Parametrization

of the non-perturbative part of $f_1(x, k_T)$

MAP Collaboration
arXiv:2502.04166

Neural Network



parametrization

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

41+1 parameters

- physically required constraints

$$f_{\rm NP} \to 1 \quad \text{for} \quad b_T \to 0$$

with activation function

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

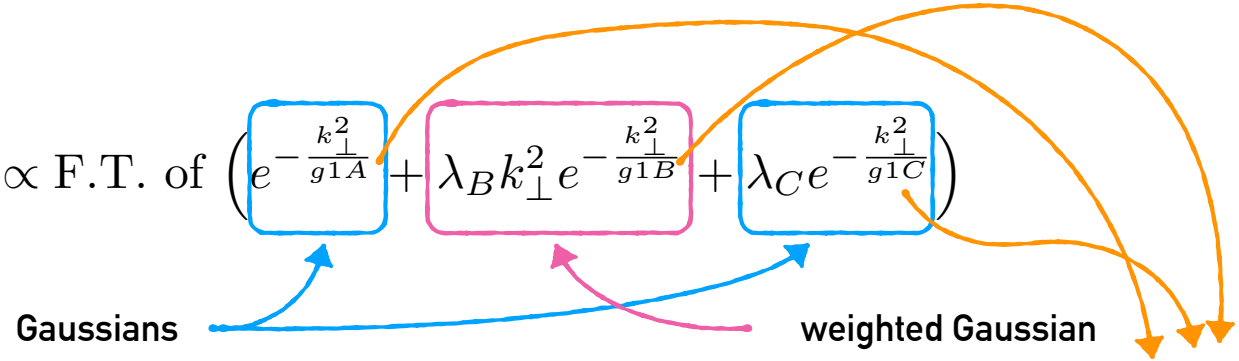
Parameterization

"analytic", "with a functional form"

of the non-perturbative part of TMDs

TMD PDF

$$f_{1\mathrm{NP}}(x, b_T^2) \propto \text{F.T. of } \left(e^{-\frac{k_\perp^2}{g_{1A}}} + \lambda_B k_\perp^2 e^{-\frac{k_\perp^2}{g_{1B}}} + \lambda_C e^{-\frac{k_\perp^2}{g_{1C}}} \right)$$



$$g_1(x) = N_1 \frac{(1-x)^\alpha \; x^\sigma}{(1-\hat{x})^\alpha \; \hat{x}^\sigma}$$

NP evolution

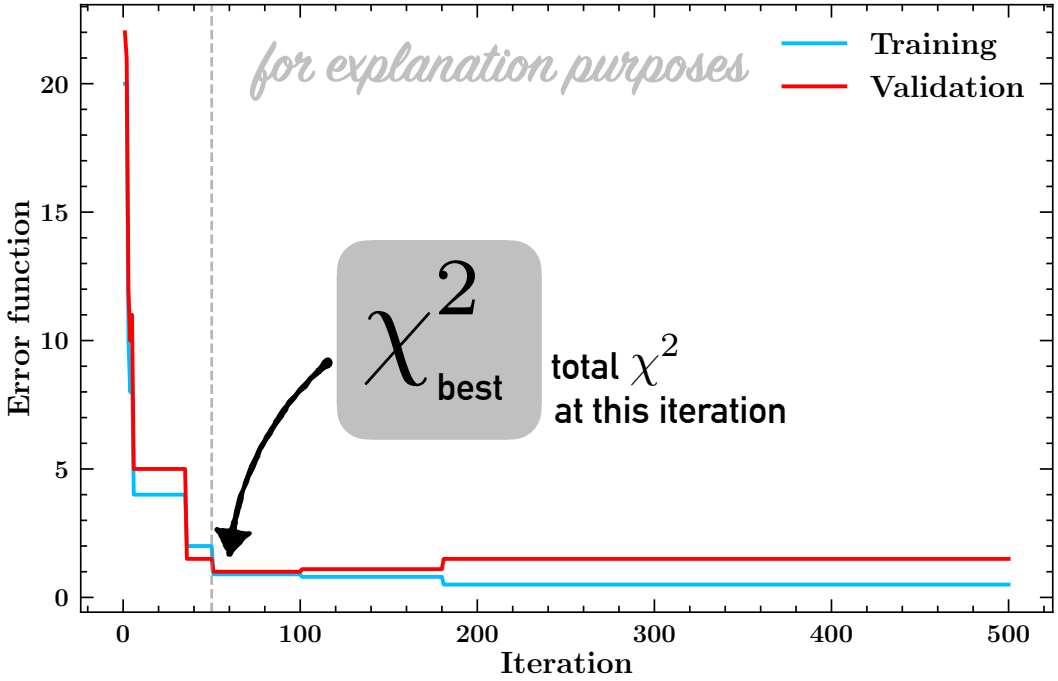
$$g_K(b_T^2) = -g_2^2 \frac{b_T^2}{4}$$

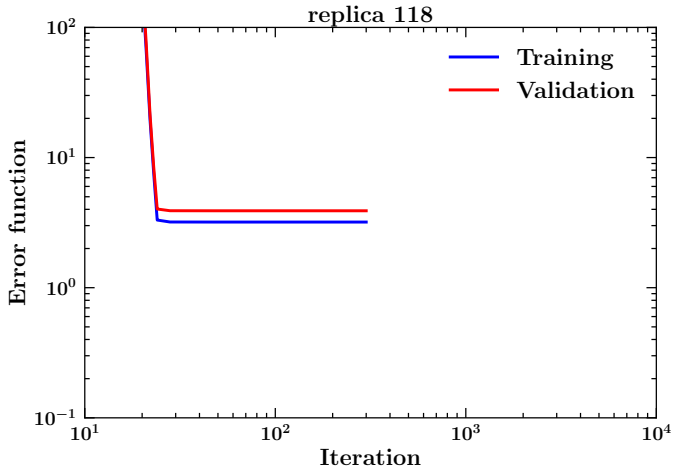
MAP22

12 parameters

avoid overfitting

split data into training and validation set





random splitting

50% – 50%

dataset by dataset

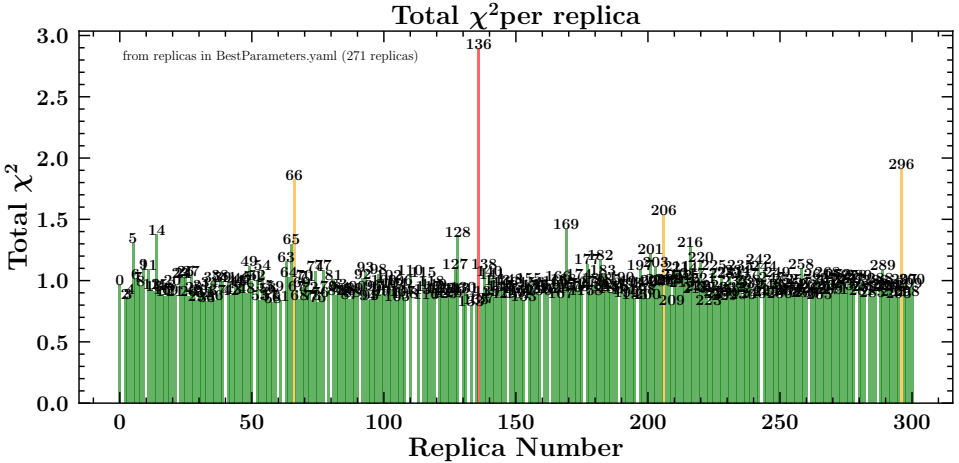
Monte Carlo replicas

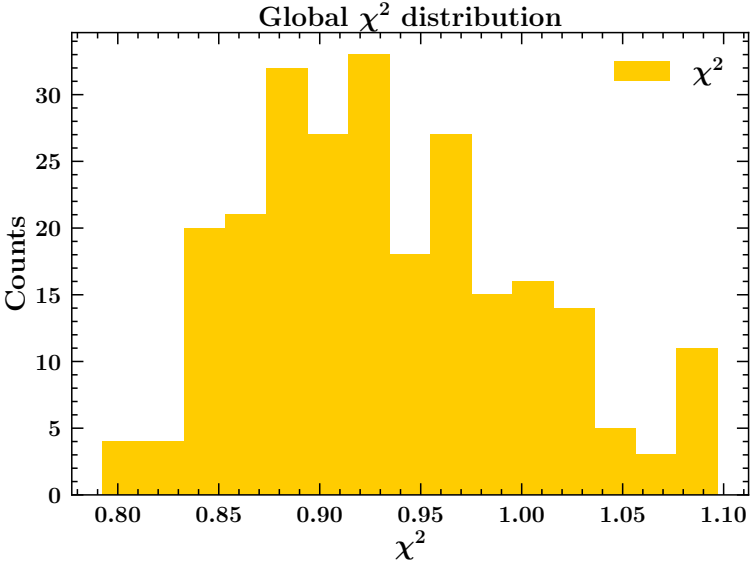
each replica is an independent fit

$$\chi^2_0 = 0.97$$

a fit runs in ~15 mins

nearly all replicas get to the end of the fit



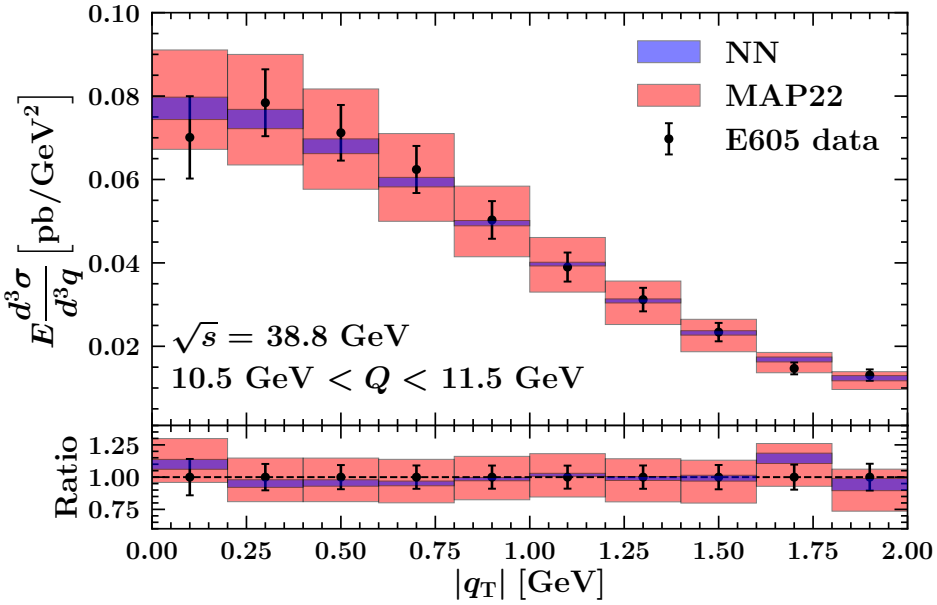


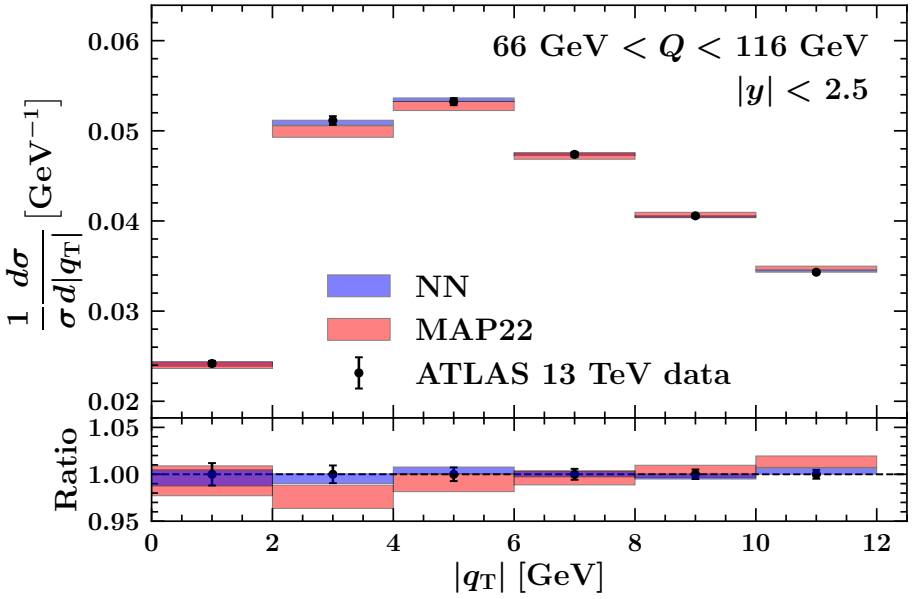
Comparison with data

logarithmic accuracy: N3LL

$\chi^2 = 0.97$ NN fit: great description of data

$\chi^2 = 1.28$ MAP22 (for comparison)





Description of data

numerical breakdown

$$\chi^2 = \sum_{i=1}^{n} \left(\frac{m_i - \bar{t}_i}{s_i} \right)^2 + \sum_{\alpha=1}^{k} \lambda_\alpha^2$$

uncorrelated contribution

penalty term



rule of thumb: it's best when it's small

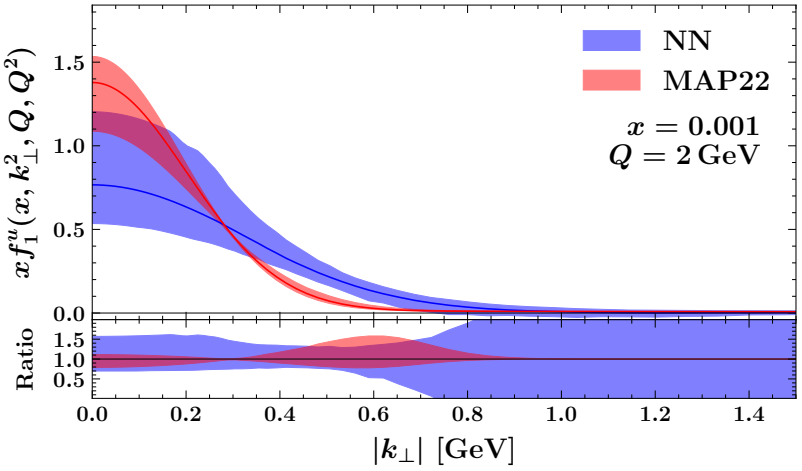
Experiment	N_{dat}	$\overline{\chi}^2$ $(\overline{\chi}^2_D + \overline{\chi}^2_\lambda)$	
		NN	MAP22
Fixed-target	233	1.08 (0.98 + 0.10)	0.91 (0.70 + 0.21)
RHIC	7	1.11 (1.03 + 0.07)	1.45 (1.37 + 0.08)
Tevatron	71	0.80 (0.73 + 0.06)	1.20 (1.17 + 0.04)
LHCb	21	0.98 (0.88 + 0.10)	1.25 (1.05 + 0.20)
CMS	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)

great improvement

NN TMDs

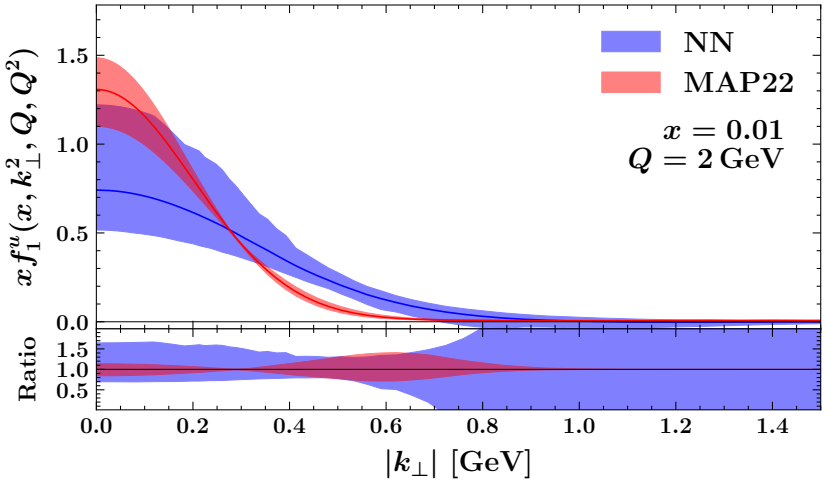
extraction of $f_1(x, k_T)$ from DY data

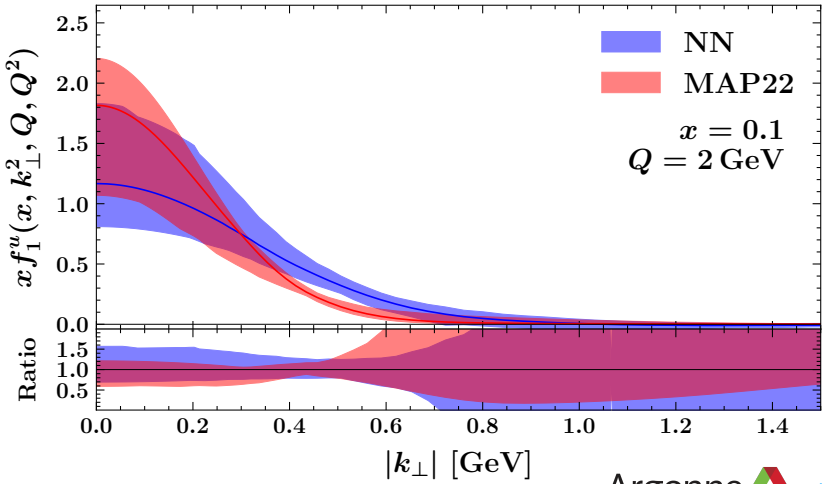
MAP Collaboration,
Phys. Rev. Lett. 135 (2025) 2, 021904



NN fit has larger bands

more reliable estimation uncertainties

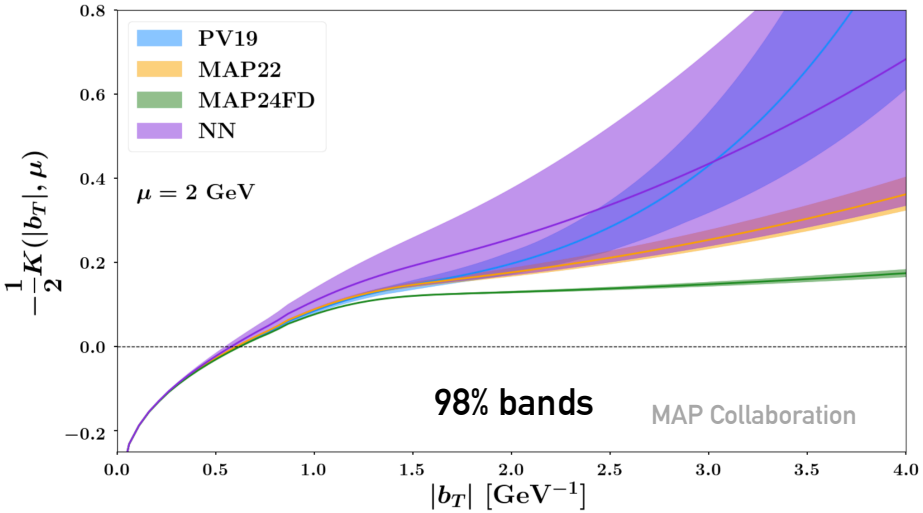


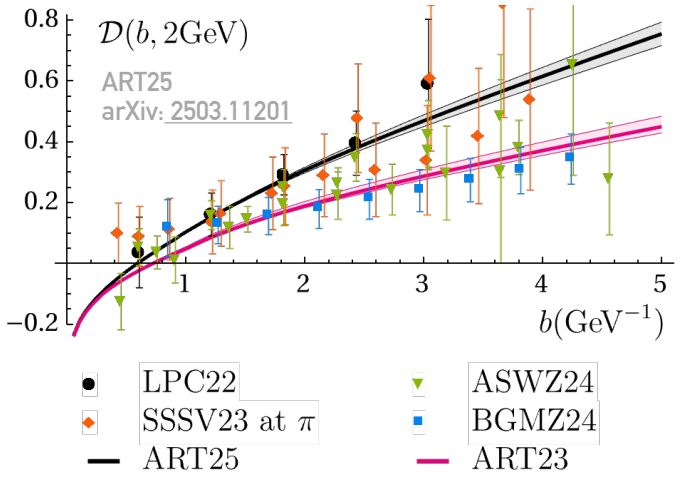


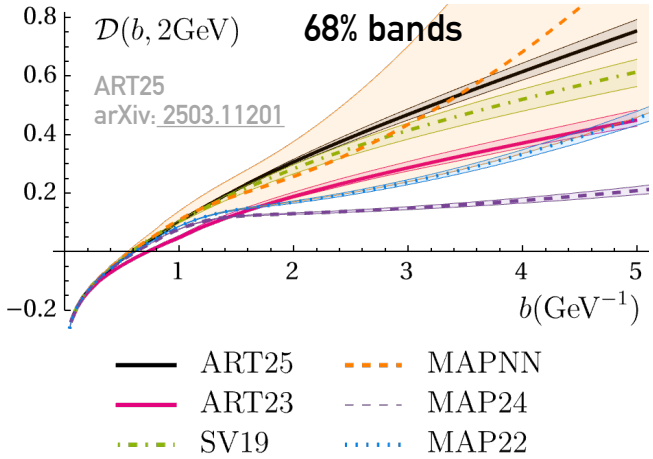
Collins-Soper kernel

"strong" universality

process independent,
insensitive to the types of external hadrons involved,
not dependent on polarization, on the flavors of the quarks, and on the scale Q







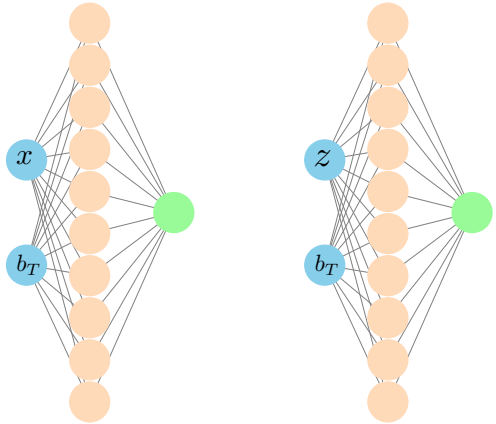
good agreement between the extractions

Extension of the NN fit to SIDIS data

global unpolarised NN TMD extraction

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

$$D_{\rm NP}(z, b_T; \zeta) = \frac{\mathbb{NN}(z, b_T, \{p_i\})}{\mathbb{NN}(z, 0, \{p_i\})} \exp\left[-g_2^2\, b_T^2\, \ln\left(\frac{\zeta}{Q_0^2}\right)\right]$$



41+ 41+1 parameters

Global NN unpolarized TMD fit

chisquare comparisons

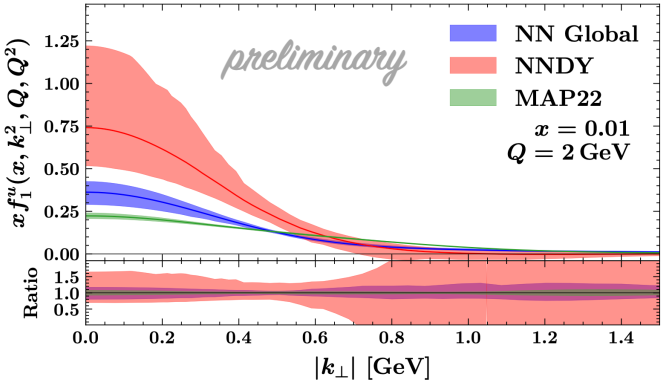


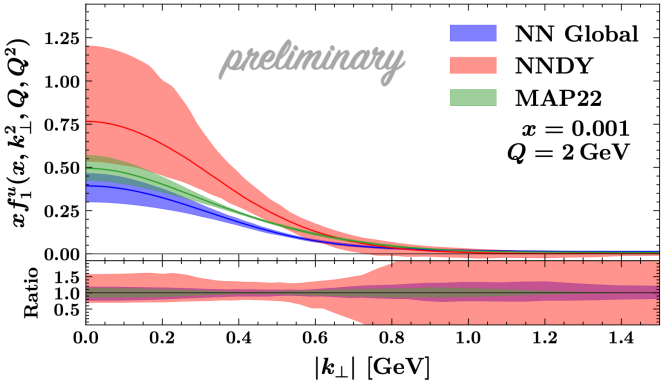
Experiment	$N_{\rm dat}$	$\overline{\chi}^2$ $(\overline{\chi}^2_D + \overline{\chi}^2_\lambda)$		
		NN Global	MAP22	NN DY-only
Fixed-target	233	0.85 (0.65 + 0.20)	1.24 (0.85 + 0.40)	1.08 (0.98 + 0.10)
RHIC	7	2.64 (2.62 + 0.04)	1.14 (1.05 + 0.10)	1.11 (1.03 + 0.07)
Tevatron	71	2.04 (1.86 + 0.18)	0.93 (0.87 + 0.06)	0.80 (0.73 + 0.06)
LHCb	21	1.32 (1.14 + 0.19)	1.45 (1.15 + 0.30)	0.98 (0.88 + 0.10)
CMS	78	0.63 (0.60 + 0.03)	0.55 (0.53 + 0.02)	0.40 (0.38 + 0.02)
ATLAS	72	2.87 (2.49 + 0.38)	5.05 (4.56 + 0.48)	1.38 (1.09 + 0.29)
DY Total	482	1.34 (1.14 + 0.19)	1.66 (1.37 + 0.29)	0.97 (0.86 + 0.11)
HERMES	344	1.33 (1.13 + 0.20)	0.71 (0.48 + 0.23)	–
COMPASS	1203	0.78 (0.59 + 0.19)	0.92 (0.62 + 0.30)	–
SIDIS Total	1547	0.90 (0.71 + 0.20)	0.87 (0.59 + 0.28)	–
Total	**2029**	**1.01 (0.81 + 0.19)**	**1.06 (0.77 + 0.29)**	—

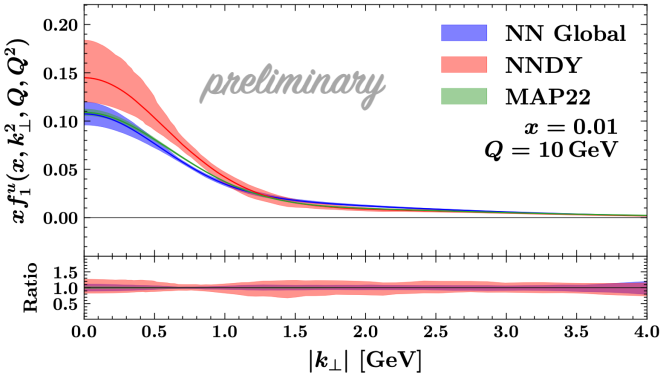
- NN global shows a general improvement in the fit quality for most of the datasets with respect to MAP22
- deterioration in the fit quality for DY data between NN global and NN-DY

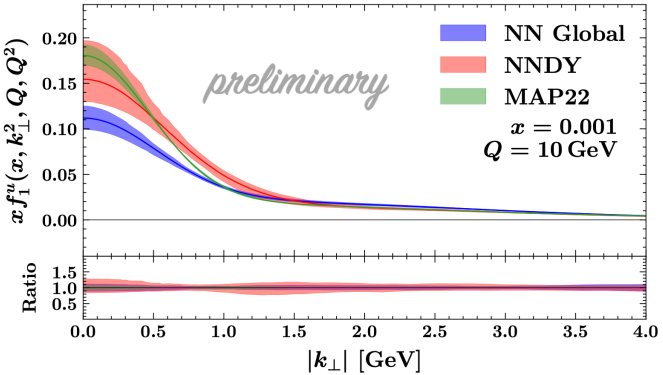
tensions between DY and SIDIS prevent optimal fit?

Global NN unpolarized TMD fit







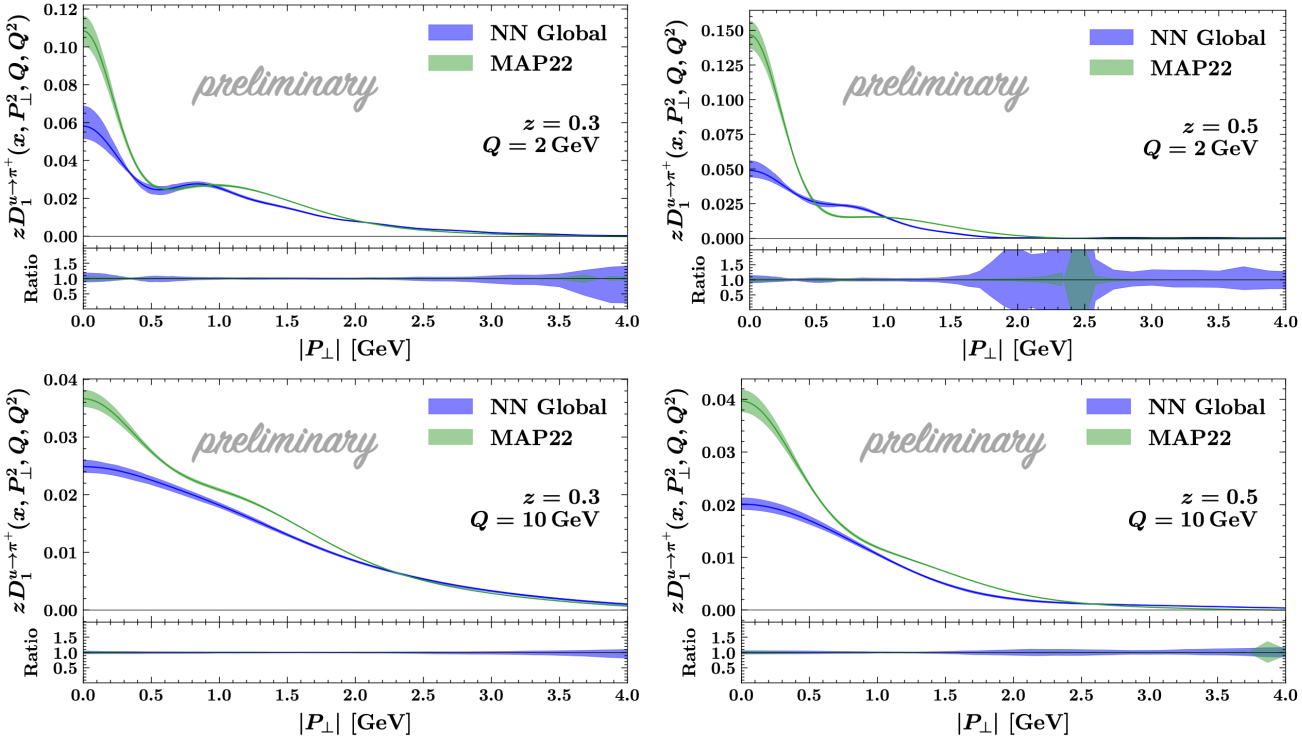




- NN global is broader than DY-only
- uncertainty bands are smaller than DY-only, but still larger than MAP22

Global NN unpolarized TMD fit

Fragmentation functions





- at Q = 2 GeV, the bump is still visible in NN global,

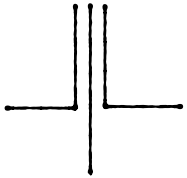
  suggest that this behavior may still be a genuine feature of the data at lower scales?

- at Q = 10 GeV NN global is smoother than MAP22, there is no bump

Conclusions

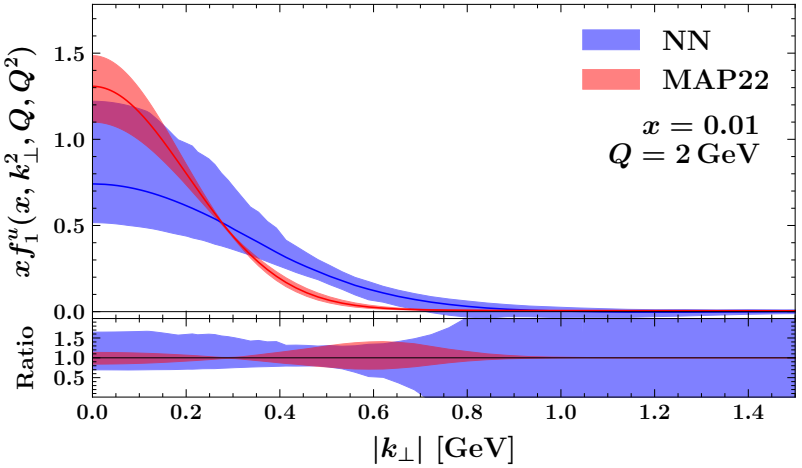
- **first Neural Networks extraction**

  of the unpolarized quark TMD PDFs in the proton

  at N3LL  Drell-Yan

  great description of data

  $\chi^2 = 0.97$



more reliable uncertainties

- combined **Drell-Yan + SIDIS** fit is still *preliminary*

  $\chi^2 = 1.01$



Argonne

NATIONAL LABORATORY

NangaParbat

main code used by the MAP Collaboration

— all fits so far have been done with



Nanga Parbat: a TMD fitting framework

Nanga Parbat is a fitting framework aimed at the determination of the non-perturbative component of TMD distributions.

Download

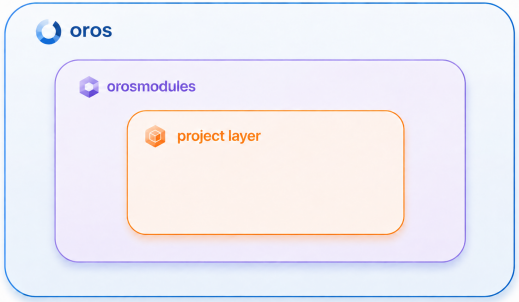
You can obtain NangaParbat directly from the github repository: https://github.com/MapCollaboration/NangaParbat

- suited for fitting unpolarized TMDs
- written in C++

OROS

the genesis of a fitting a model to data

MAP Collaboration





— more flexible, modular and versatile

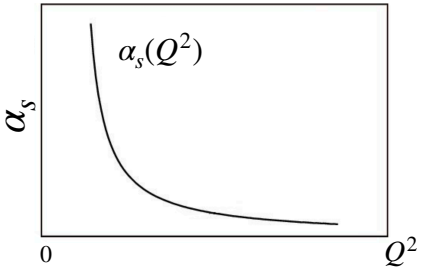
- suited for fitting not only TMDs, but also PDFs and eventually GPDs
- very fast
- written in C++ but with a python wrapper

b* prescription

integration up to infinity

$$\frac{d\sigma}{dq_T} \propto \int \frac{d^2\mathbf{b}}{4\pi} e^{i\mathbf{b}\cdot\mathbf{q}_T} f_1^q(x_1, \mathbf{b}) f_1^{\bar{q}}(x_2, \mathbf{b})$$

cross section



when b_T becomes large

$$\alpha_s(\mu_b) = \alpha\left(\frac{2e^{-\gamma_E}}{b}\right) \gg 1$$

invalidates perturbative calculations

$\Rightarrow b_{\max}$

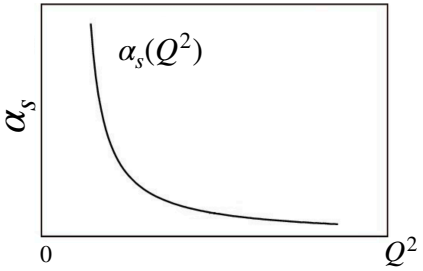
necessary to introduce a prescription

b* prescription

integration up to infinity

$$\frac{d\sigma}{dq_T} \propto \int \frac{d^2\mathbf{b}}{4\pi} e^{i\mathbf{b}\cdot\mathbf{q}_T} f_1^q(x_1,\mathbf{b}) f_1^{\bar{q}}(x_2,\mathbf{b})$$

cross section

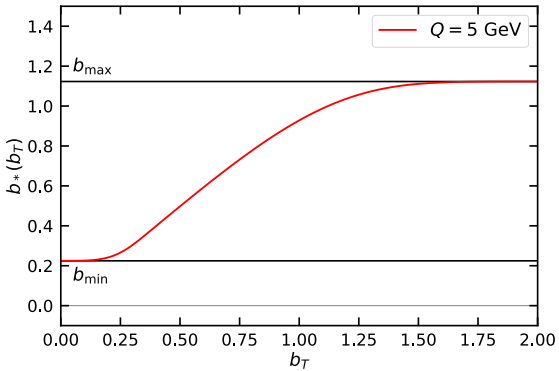


when b_T becomes large

$$\alpha_s(\mu_b) = \alpha\left(\frac{2e^{-\gamma_E}}{b}\right) \gg 1$$

invalidates perturbative calculations

$$\Rightarrow b_{\max}$$



$$b_{\max} = 2e^{-\gamma_E}$$

$$b_{\min} = 2e^{-\gamma_E}/Q$$

necessary to introduce a prescription

$$b_*(b) = b_{\max}\left(\frac{1-\exp\left(-\frac{b^4}{b_{\max}^4}\right)}{1-\exp\left(-\frac{b^4}{b_{\min}^4}\right)}\right)^{\frac{1}{4}}$$

b* prescription and definition of $f_{\rm NP}$

perturbative

$$f(x,b;\mu,\zeta) = \left[\frac{f(x,b;\mu,\zeta)}{f(x,b_*(b);\mu,\zeta)}\right] f(x,b_*(b);\mu,\zeta)$$

non perturbative

$f_{\rm NP}(x,b,\zeta)$ → fit to data

Non perturbative function depends on the choice of b*-prescription

TMDs — *unpolarized quark TMD PDF*

factorizes as **hard** and **longitudinal non-perturbative**

matching to the collinear region

$b_T \ll 1/\Lambda_{\rm QCD}$

$$F_{f/P}(x, \mathbf{b}_T; \mu, \zeta) = \sum_j C_{f/j}(x, b_*; \mu_b, \zeta_F) \otimes f_{j/P}(x, \mu_b)$$

$$\times \exp\left\{ K(b_*; \mu_b) \ln \frac{\sqrt{\zeta_F}}{\mu_b} + \int_{\mu_b}^{\mu} \frac{d\mu'}{\mu'} \left[\gamma_F - \gamma_K \ln \frac{\sqrt{\zeta_F}}{\mu'} \right] \right\}$$

CS and RGE evolution

$$\times \exp\left\{ g_{j/P}(x, b_T) + g_K(b_T) \ln \frac{\sqrt{\zeta_F}}{\sqrt{\zeta_{F,0}}} \right\}$$

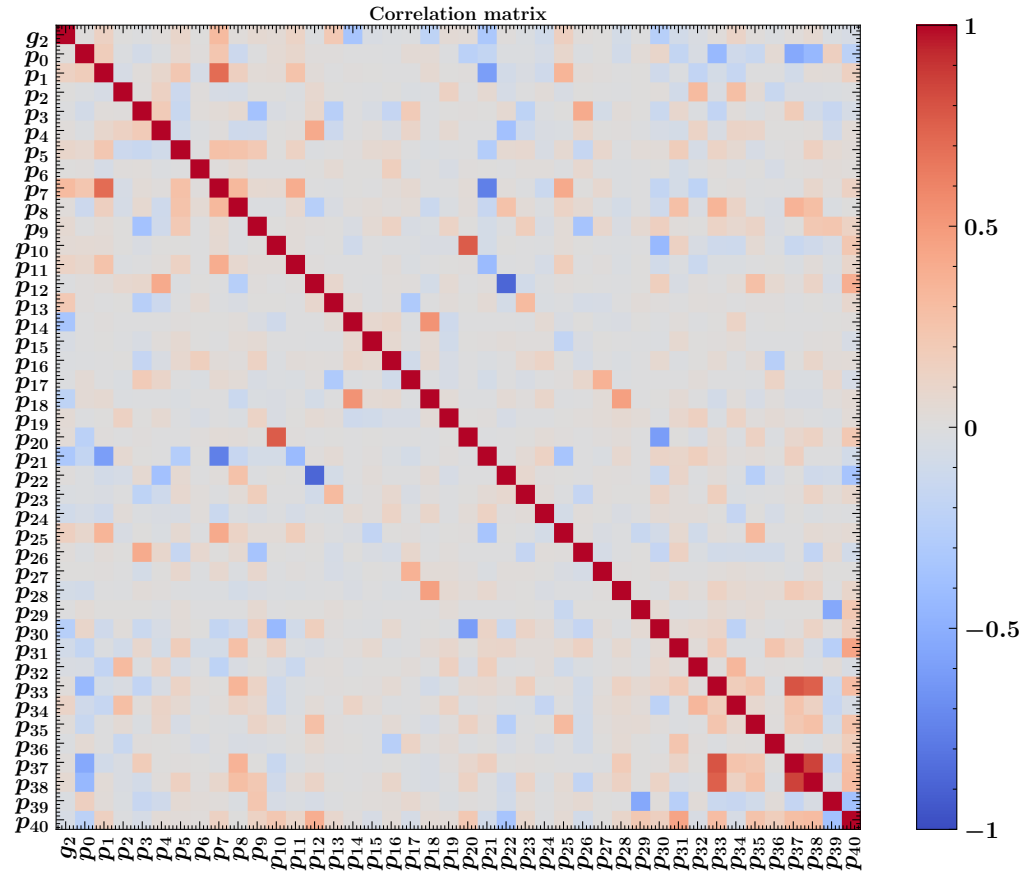
non-perturbative transverse content

parametrized and fitted to data

correlation matrix

MAP Collaboration
arXiv:2502.04166

41 parameters

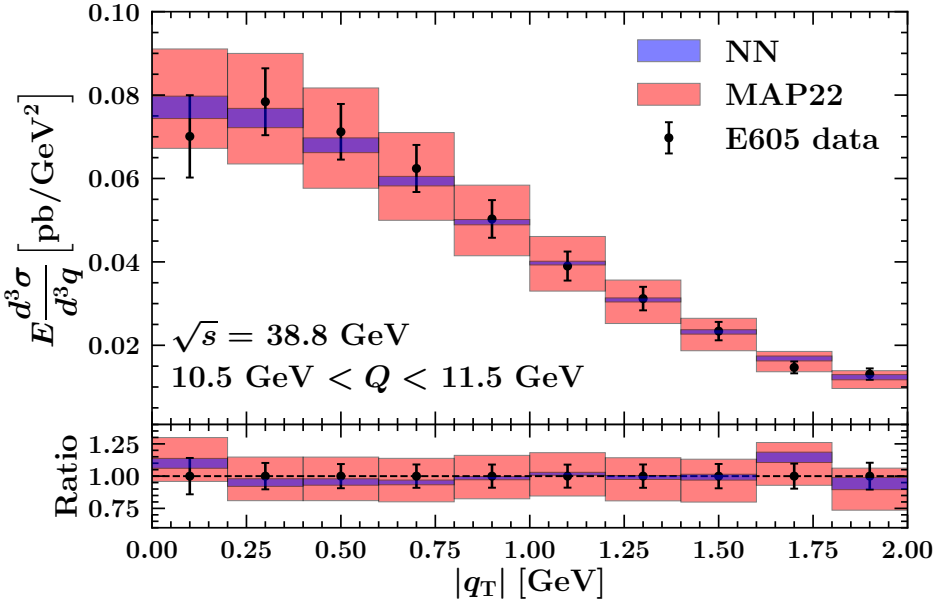


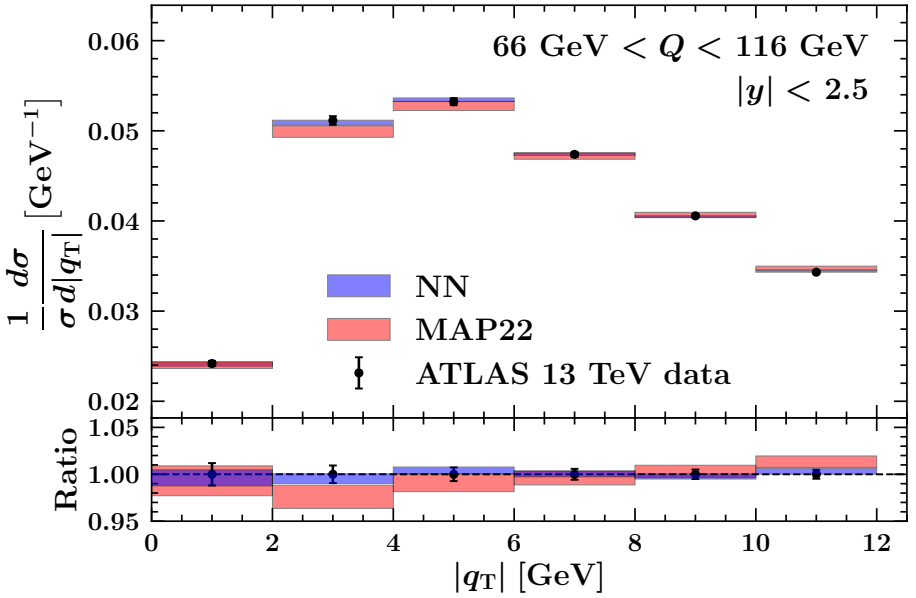
Comparison with data

logarithmic accuracy: N3LL

$\chi^2 = 0.97$

great description of data





why are NN bands so much smaller than MAP22's?

Shifted predictions

& decomposition of the χ^2

experimental data

$$m_i \pm \sigma_{i,\mathrm{unc}} \pm \sigma^{(1)}_{i,\mathrm{corr}} \pm \cdots \pm \sigma^{(k)}_{i,\mathrm{corr}}$$

uncorrelated correlated

$$\chi^2 = \sum_{i=1}^{n} \left(\frac{m_i - \bar{t}_i}{s_i} \right)^2 + \sum_{\alpha=1}^{k} \lambda_\alpha^2$$

uncorrelated contribution penalty term

nuisance parameters $\dfrac{\partial \chi^2}{\partial \lambda_\alpha} = 0$

$$\bar{t}_i = t_i + d_i$$

shifted prediction

shift $d_i = \sum_{\alpha=1}^{k} \lambda_\alpha \sigma^{(\alpha)}_{i,\mathrm{corr}}$

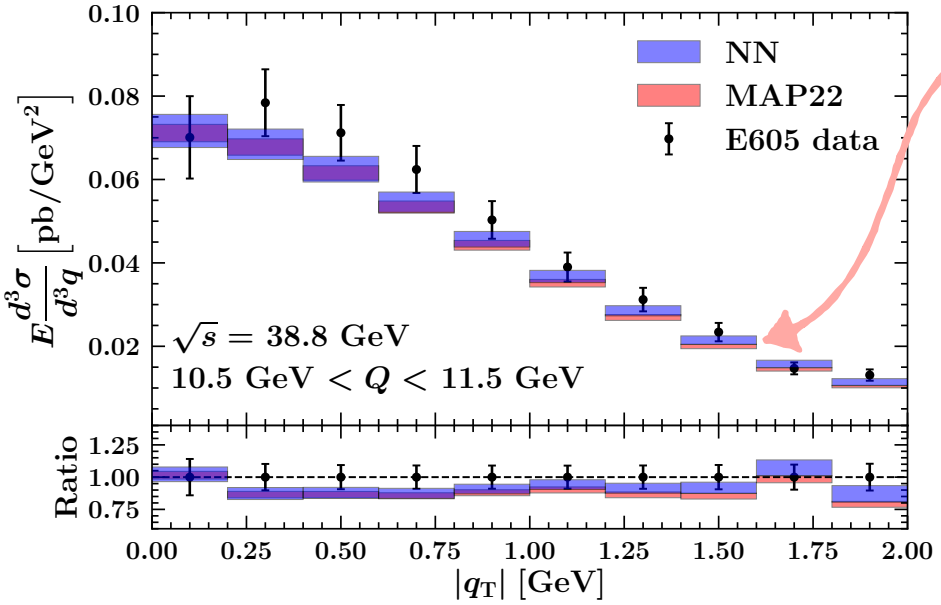
Comparison with data

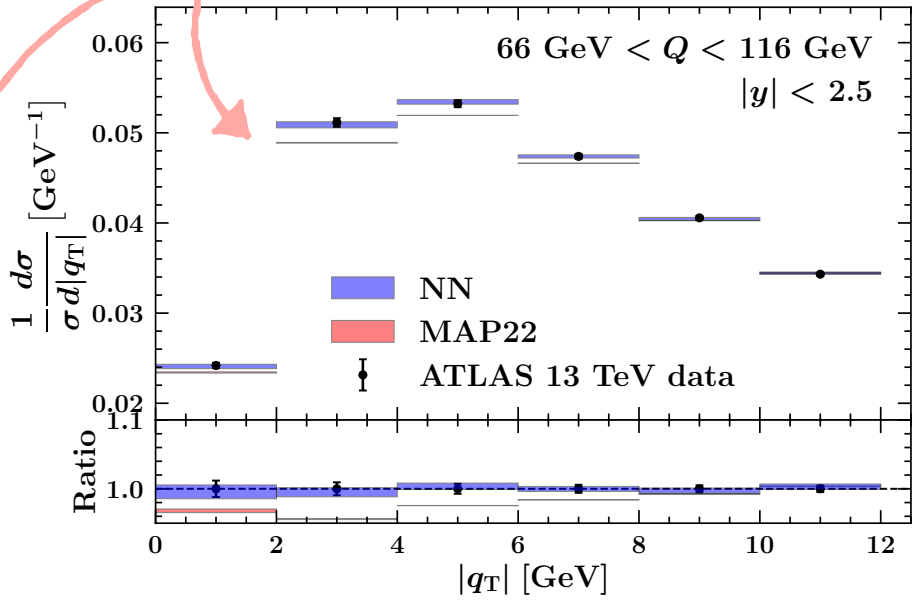
without including systematic shifts

$\chi^2 = 0.97$

$\chi^2 = 1.28$

significant reduction
of the bands for MAP22



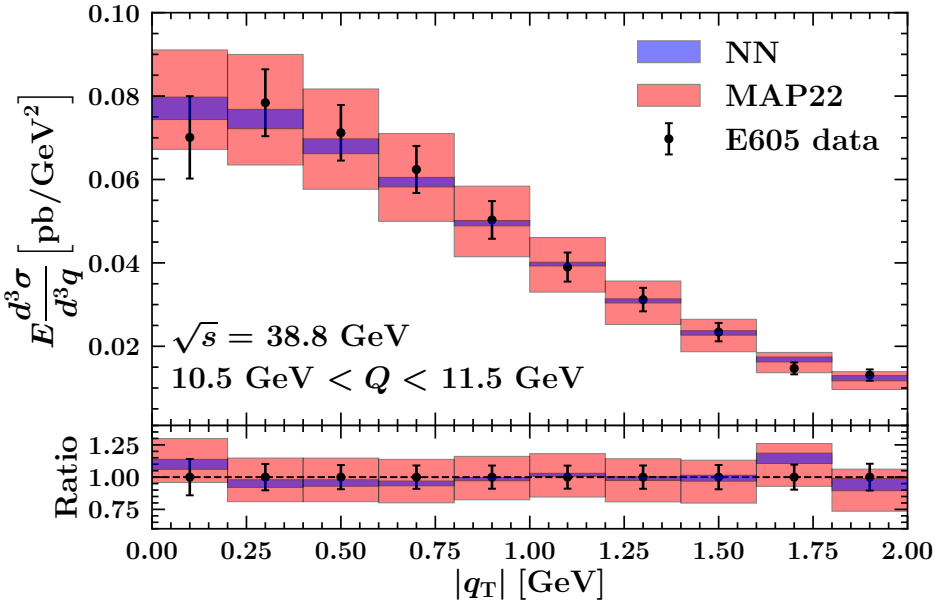


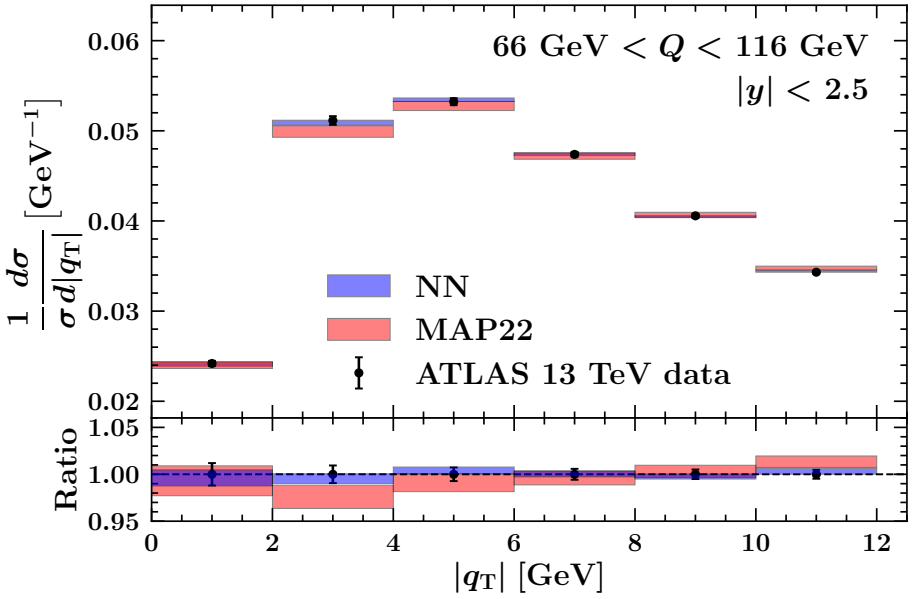
. . . just to compare again

N3LL

$\chi^2 = 0.97$

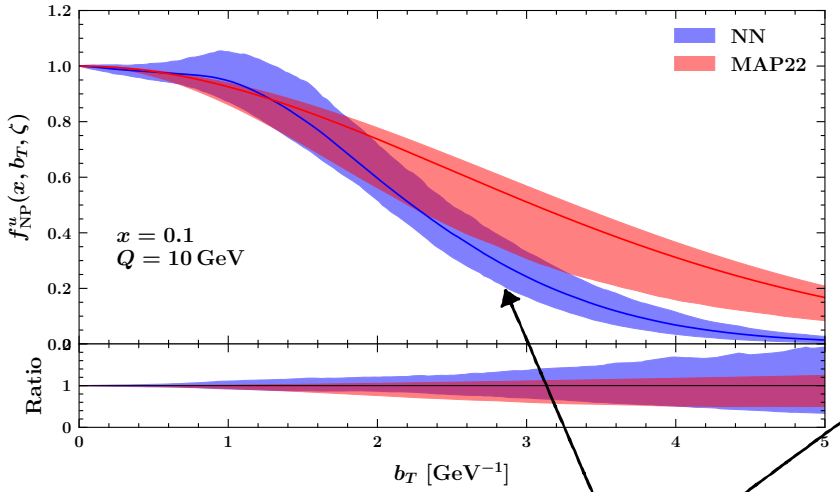
$\chi^2 = 1.28$

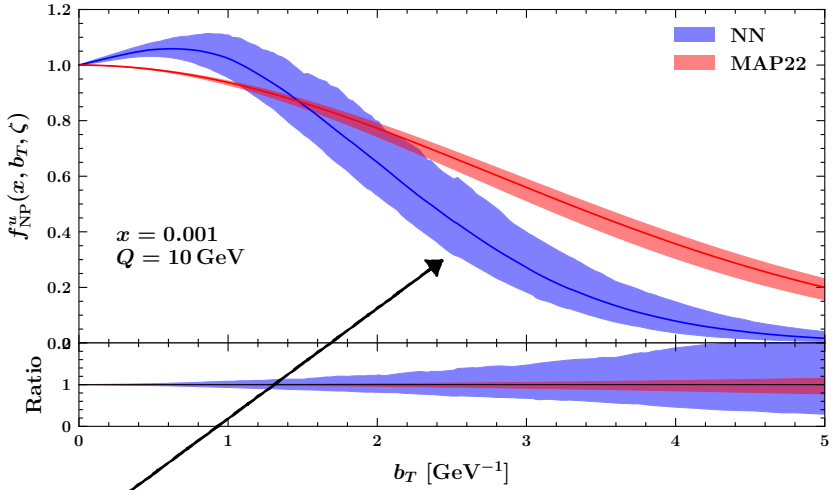


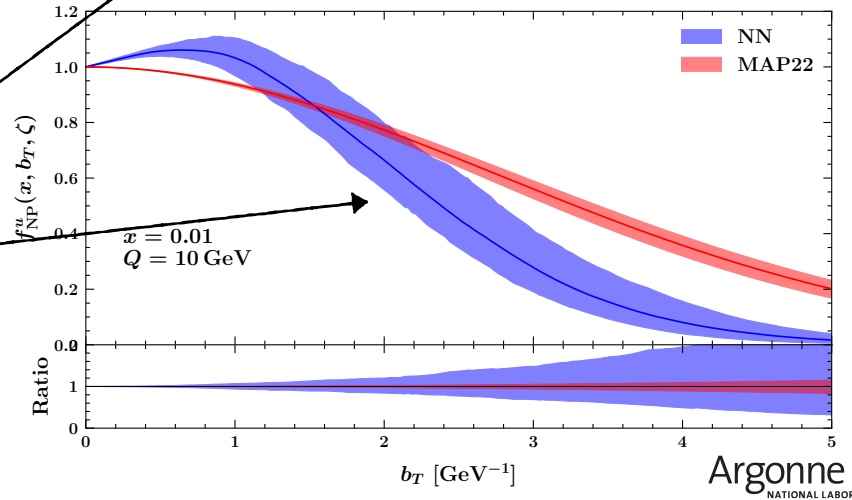


fNP

comparison with MAP22 only DY



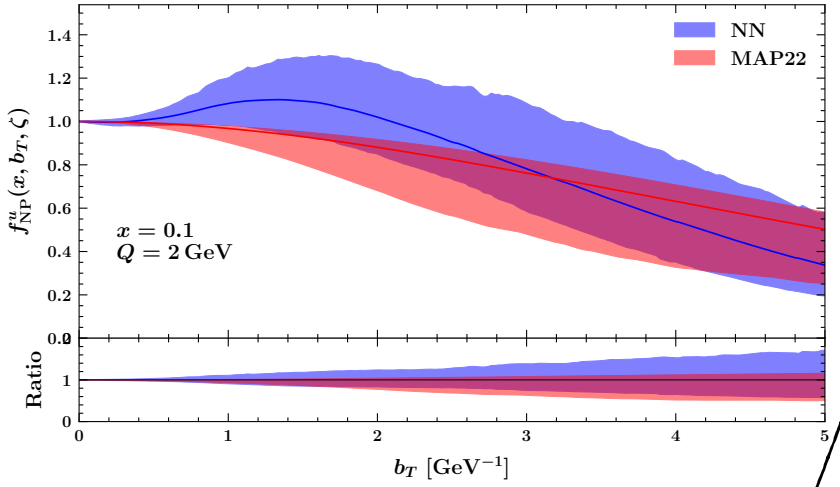


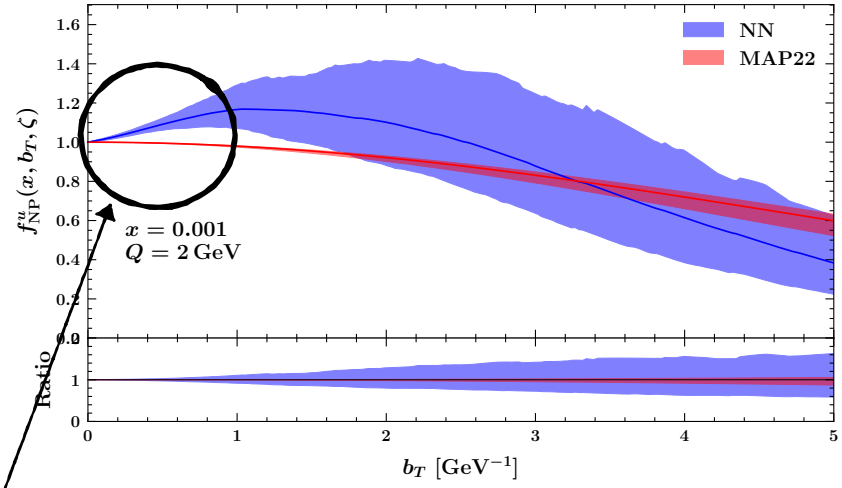


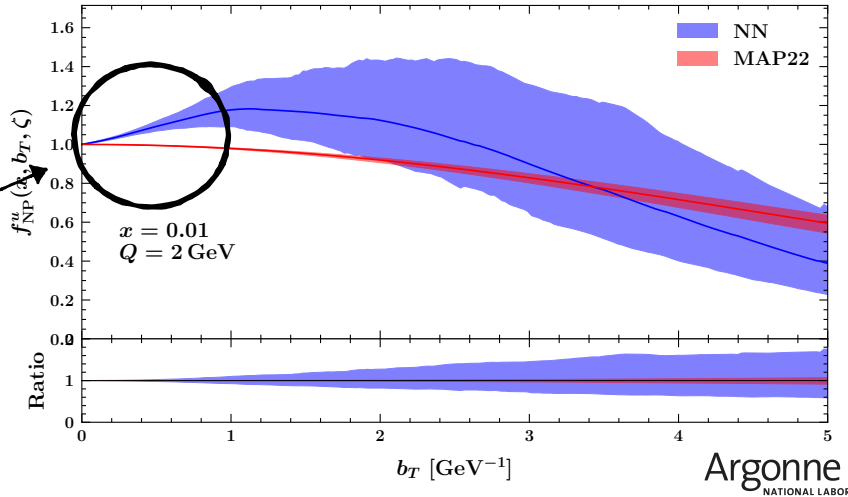
NN falls sharply with bT

fNP

comparison with MAP22 only DY







NN has a bump at very small bT
it reaches 1 from above

Closure tests

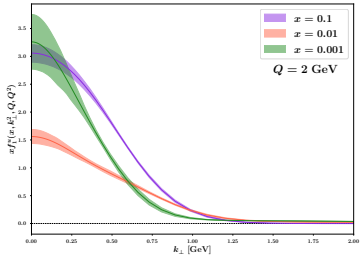
to validate the methodology

- **Level 0**
  - central pseudo-data is given by **central predictions** of the known model
  - **no Monte Carlo noise** is added on top of the central data — each replica is fitting the same set of data
- **Level 1**
  - central pseudo-data is shifted by some noise η drawn from the experimental covariance matrix
  - no MC noise is added — each replica fits a subset of the same shifted data
- **Level 2**
  - central pseudo-data is shifted by level 1 noise η
  - MC noise is added on top of the level 1 shift

Closure tests

only Drell-Yan data

to validate the methodology



closure test – level 0

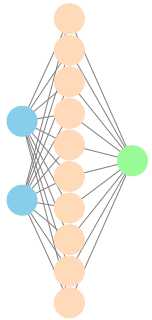
$$g_1 = 0.304$$
$$g_2 = 0.028$$

$$f_{\rm NP}^{\rm DWS}(b_T,\zeta) = \exp\left[-\frac{1}{2}\left(g_1 + g_2 \ln\left(\frac{\zeta}{2Q_0^2}\right)\right) b_T^2\right]$$

Generate **pseudodata** based on a model that we know

- central value: computed with **DWS** parameters
- uncertainties: real DY data

fit with **NN**



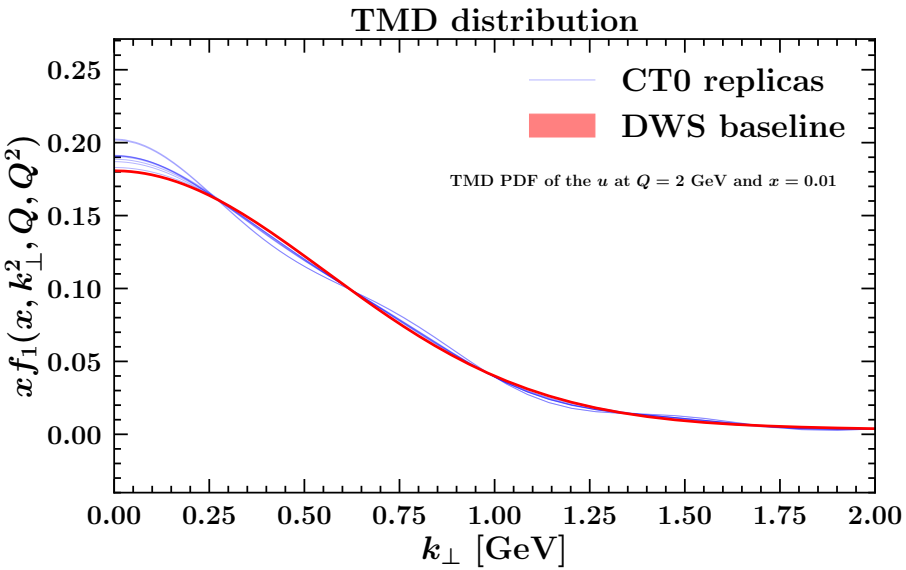
can we recover the DWS result?

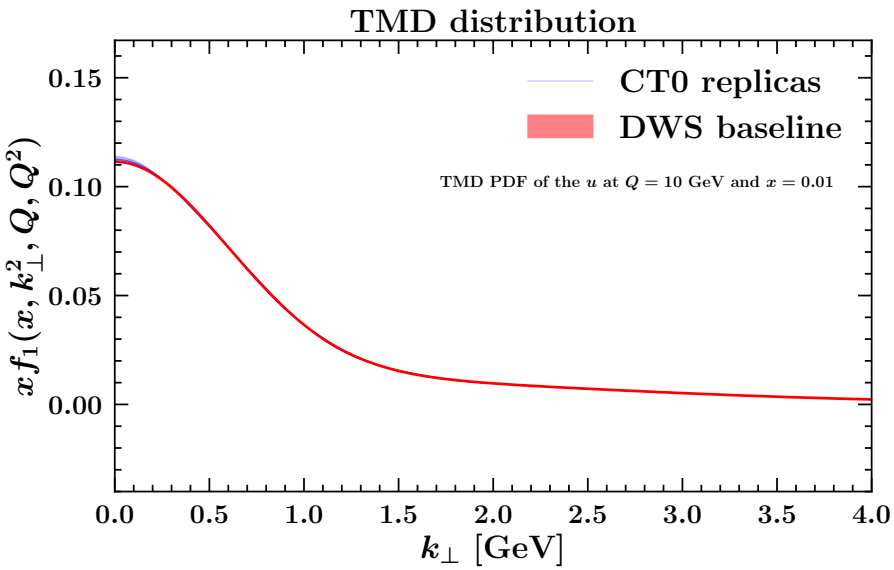
yes!

Closure tests

closure test – level 0

$$f_{\rm NP}^{\rm DWS}(b_T,\zeta) = \exp\left[-\frac{1}{2}\left(g_1 + g_2 \ln\left(\frac{\zeta}{2Q_0^2}\right)\right) b_T^2\right]$$





$$\chi^2 \sim 10^{-6}$$