QCD Evolution 20205 - Jefferson Lab

May 20, 2025

First Neural Network extraction of unpolarized TMDs

Chiara Bissolotti

Maria Goeppert Mayer Fellow
Argonne National Laboratory
PHY Division

Map Collaboration

with Valerio Bertone, Matteo Cerutti, Simone Rodini, Alessandro Bacchetta, Marco Radici, Lorenzo Rossi

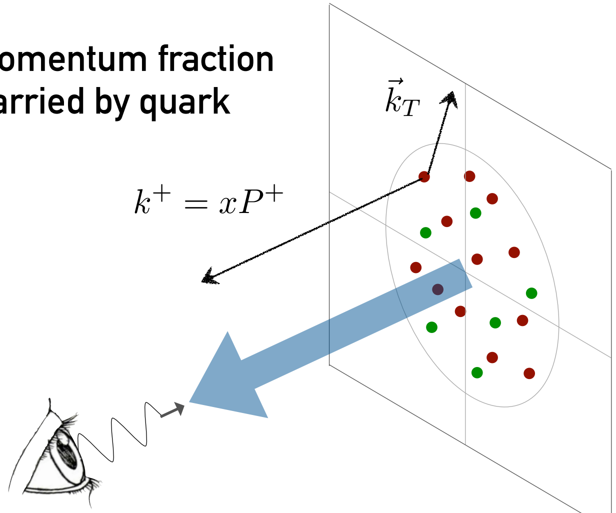
based on arXiv:2502.04166

Argonne NATIONAL LABORATORY

TMDs: 3D maps

in momentum space

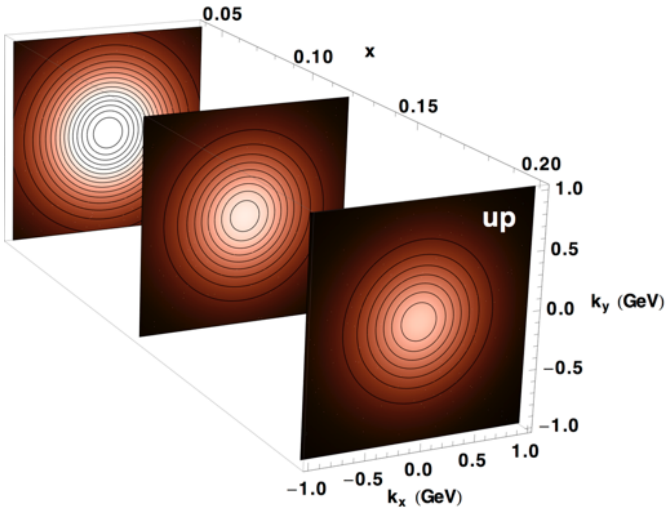
x: momentum fraction carried by quark



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

this work

Leading Quark TMDPDFs

Nucleon Spin   Quark Spin

Nucleon Polarization	Quark Polarization: Un-Polarized (U)	Quark Polarization: Longitudinally Polarized (L)	Quark Polarization: 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

from the TMD handbook
arXiv:2304.03302

TMDs — *unpolarized quark TMD PDF*

factorizes as hard and longitudinal non-perturbative

matching to the collinear region

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

$$
\begin{aligned}
F_{f/P}(x, \mathbf{b}_T; \mu, \zeta) \quad &= \quad \sum_j C_{f/j}(x, b_*; \mu_b, \zeta_F) \otimes f_{j/P}(x, \mu_b) \\
&\times \quad \exp\left\{ K(b_*; \mu_b) \ln \frac{\sqrt{\zeta_F}}{\mu_b} + \int_{\mu_b}^{\mu} \frac{d\mu'}{\mu'} \left[\gamma_F - \gamma_K \ln \frac{\sqrt{\zeta_F}}{\mu'} \right] \right\} \\
&\times \quad \exp\left\{ g_{j/P}(x, b_T) + g_K(b_T) \ln \frac{\sqrt{\zeta_F}}{\sqrt{\zeta_{F,0}}} \right\}
\end{aligned}
$$

CS and RGE evolution

non-perturbative transverse content

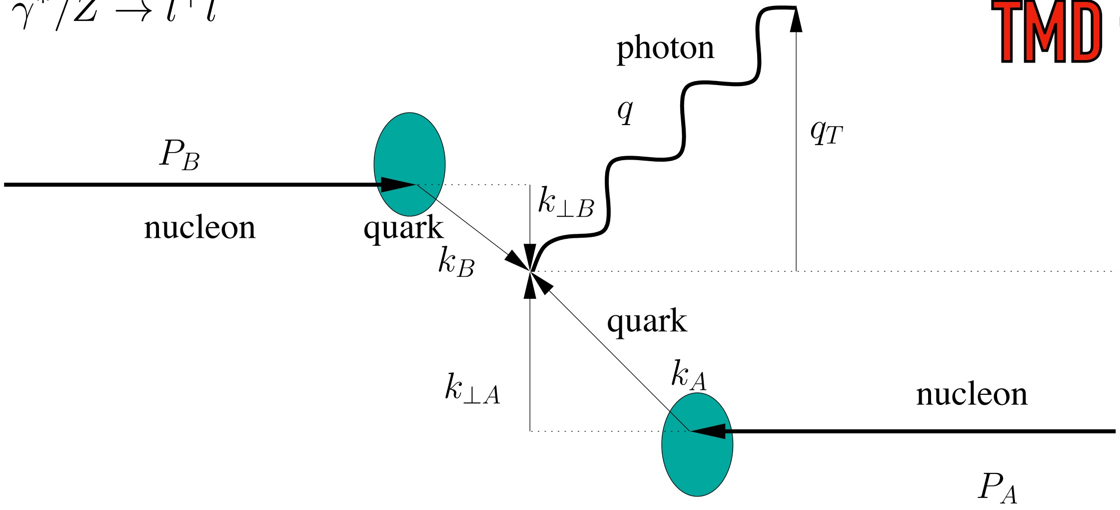
parametrized and fitted to data

Drell-Yan

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

for $q_T \ll Q$

TMD factorization



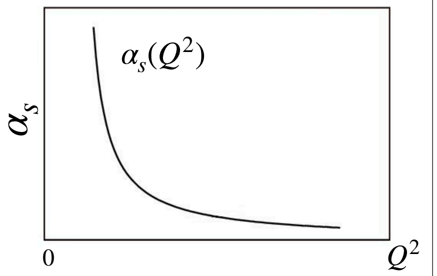
$$\left(\frac{d\sigma}{dq_T}\right)_{\text{res.}} \propto H(Q,\mu) \int \frac{d^2\mathbf{b}}{4\pi} e^{i\mathbf{b}\cdot\mathbf{q}_T} x_1 f_1^q(x_1,\mathbf{b};\mu,\zeta_1) x_2 f_1^{\bar{q}}(x_2,\mathbf{b};\mu,\zeta_2)$$

b* prescription

integration up to infinity

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

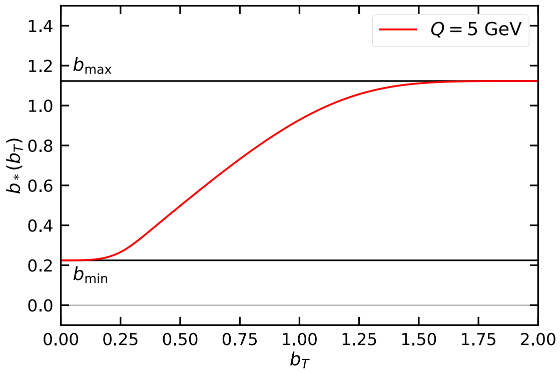
cross section



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_{NP}

perturbative

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

non perturbative

$$\boxed{f_{\mathrm{NP}}(x,b,\zeta)}$$

$\longrightarrow$ fit to data

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

Motivation

What would we like to accomplish?

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

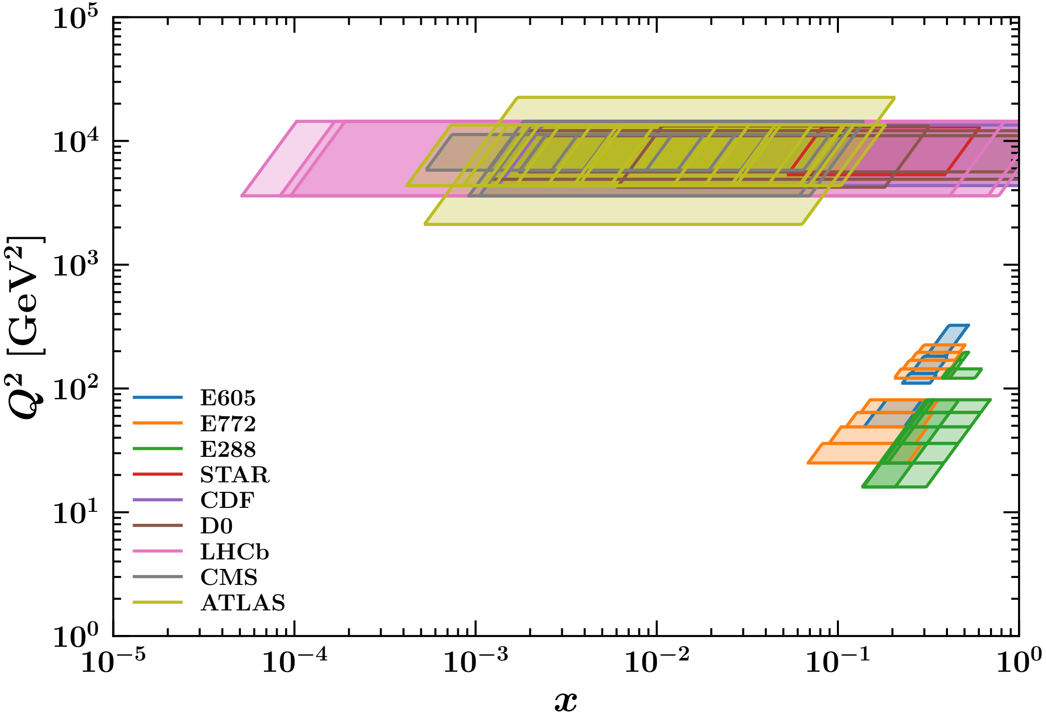
proof of concept

- do NN offer an advantage in fitting (real) data?
- we want to **test the methodology**

→ with **closure tests**

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)

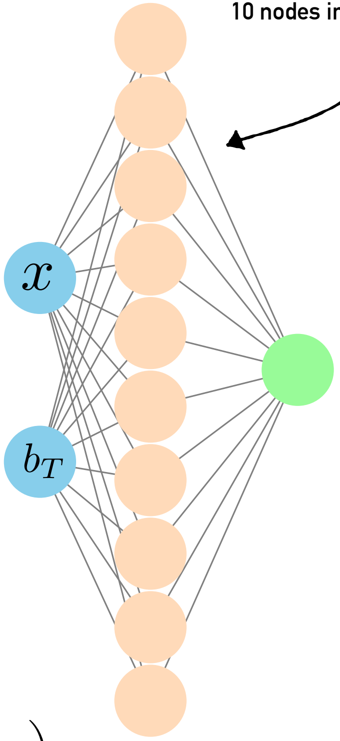
Parametrization

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

MAP Collaboration
arXiv:2502.04166

Neural Network

10 nodes in a single hidden layer



$\mathrm{NN}(x, b_T)$

only one output node because
there is no flavor dependence (yet)

41+1 parameters

parametrization

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

- physically required constraints

$f_{\mathrm{NP}} \to 1$ for $b_T \to 0$

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

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

Gaussians

weighted Gaussian

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

NP evolution

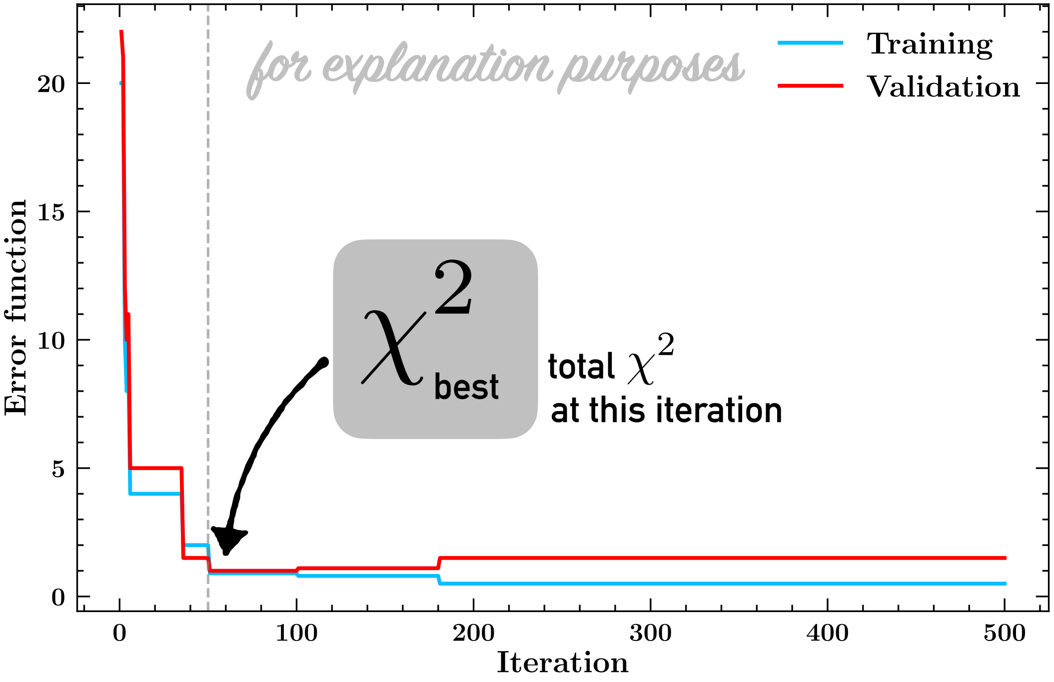
$$g_K(b_T^2) = -g_2^2 \frac{b_T^2}{4}$$

MAP22

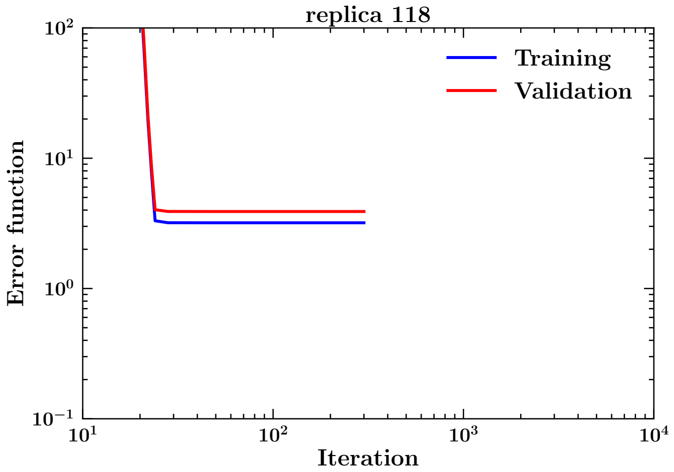
12 parameters

avoid overfitting

split data into training and validation set



χ^2_{best} total χ^2 at this iteration



random splitting

50% – 50%

dataset by dataset

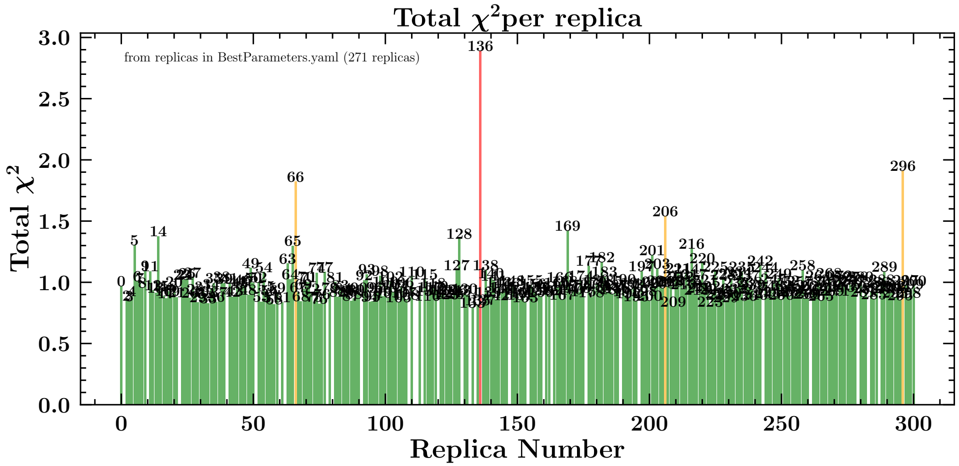
Monte Carlo replicas

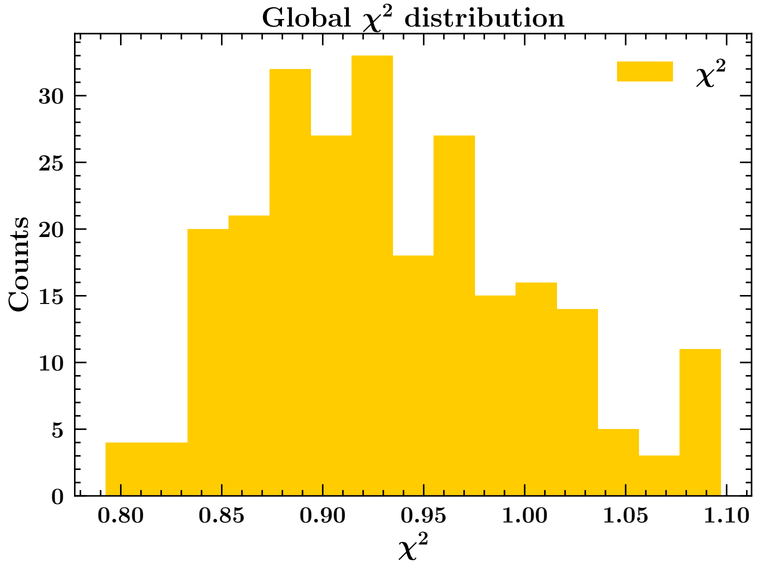
each replica is an independent fit

$\chi_0^2 = 0.97$

a fit runs in ~15 mins

nearly all replicas get to the end of the fit



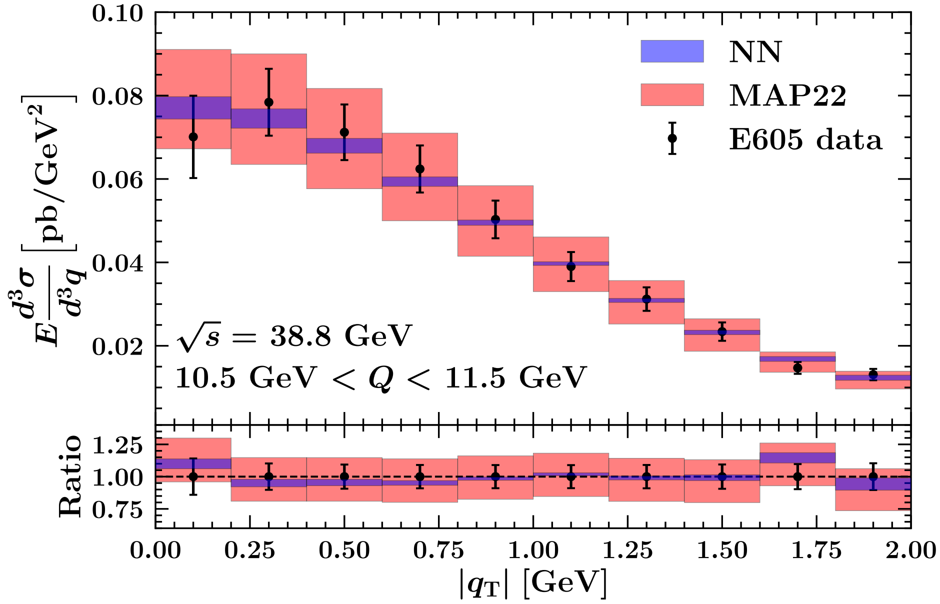


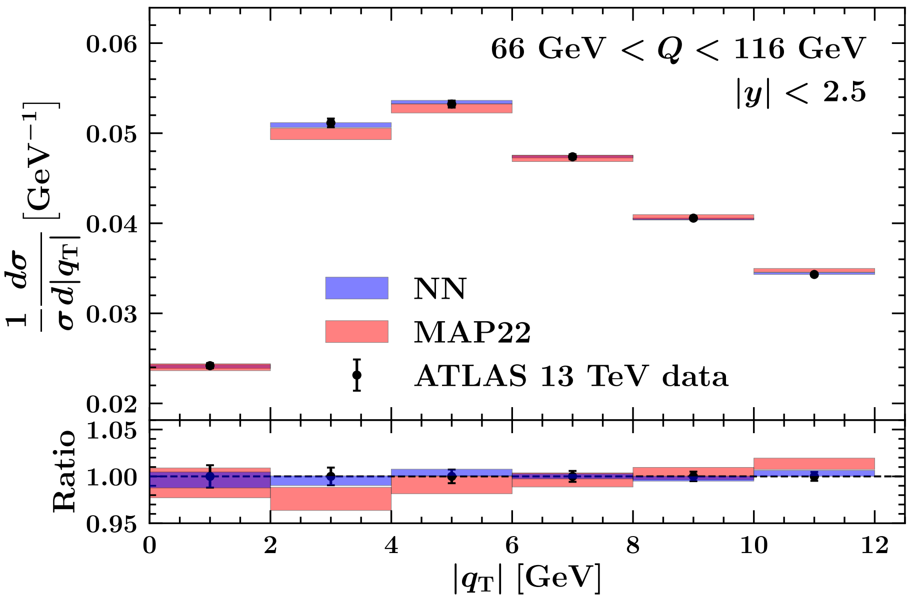
Comparison with data

logarithmic accuracy: N3LL

$\chi^2 = 0.97$

great description of data



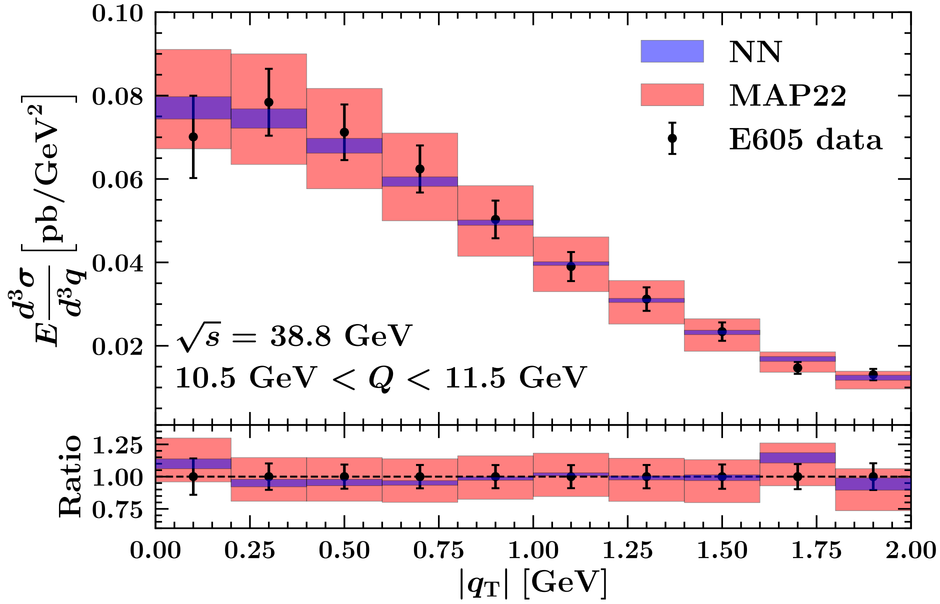


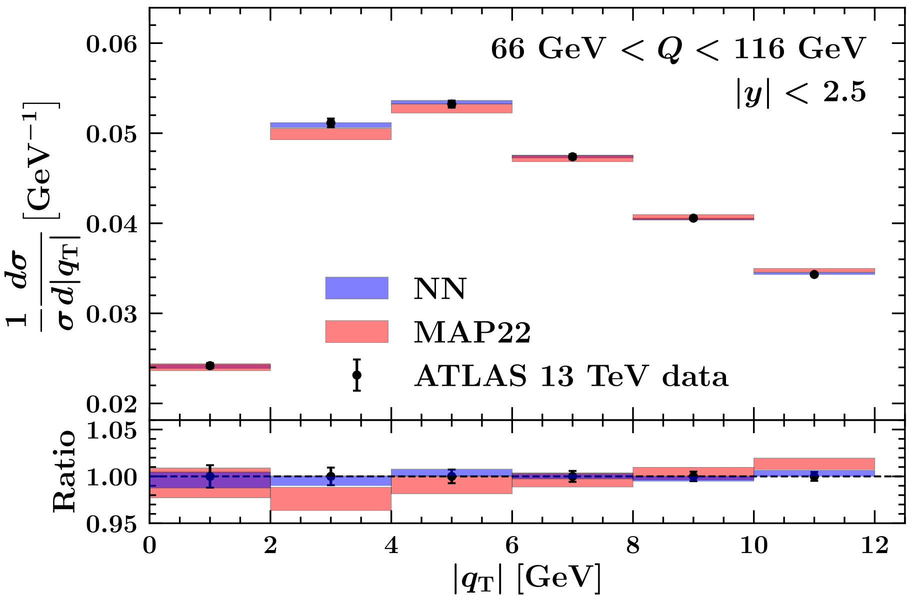
Comparison with data

logarithmic accuracy: N3LL

$\chi^2 = 0.97$

great description of data





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

U.S. DEPARTMENT OF ENERGY Argonne National Laboratory is a U.S. Department of Energy laboratory managed by UChicago Argonne, LLC.

Argonne NATIONAL LABORATORY

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

$$\chi^2 = \underbrace{\sum_{i=1}^{n} \left(\frac{m_i - \bar{t}_i}{s_i} \right)^2}_{\text{uncorrelated contribution}} + \underbrace{\sum_{\alpha=1}^{k} \lambda_\alpha^2}_{\text{penalty term}}$$

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

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

shifted prediction

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

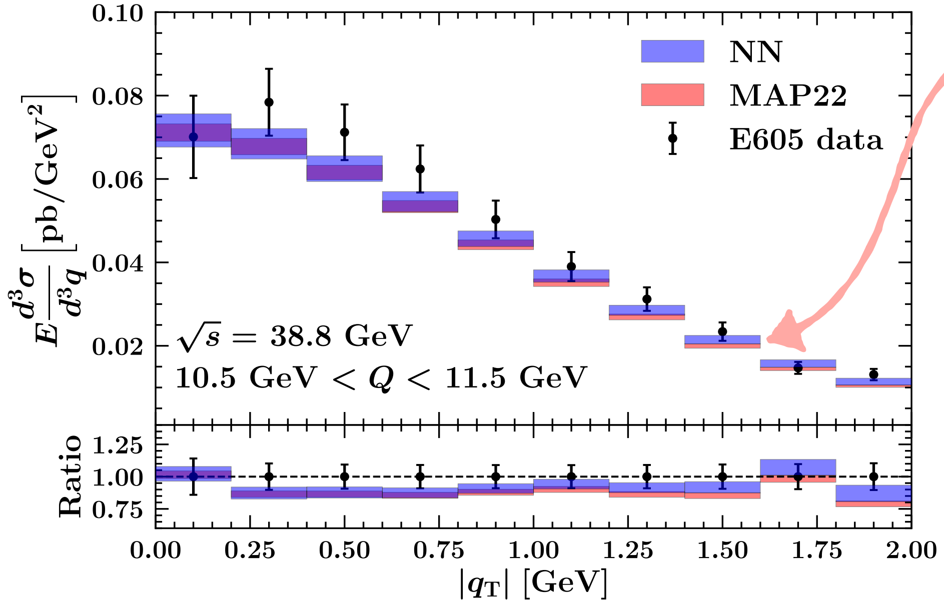
Comparison with data

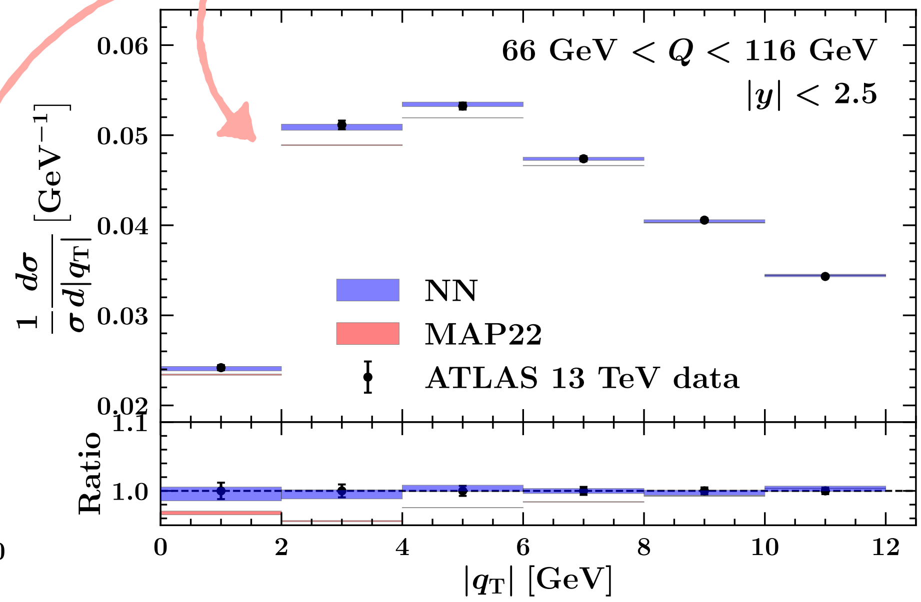
without including systematic shifts

significant reduction
of the bands for MAP22

$\chi^2 = 0.97$

$\chi^2 = 1.28$



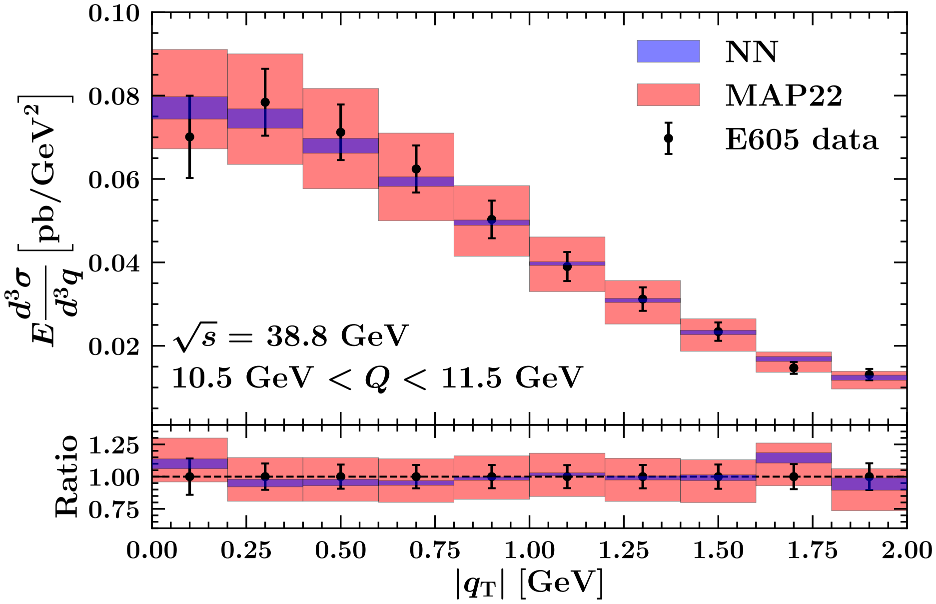


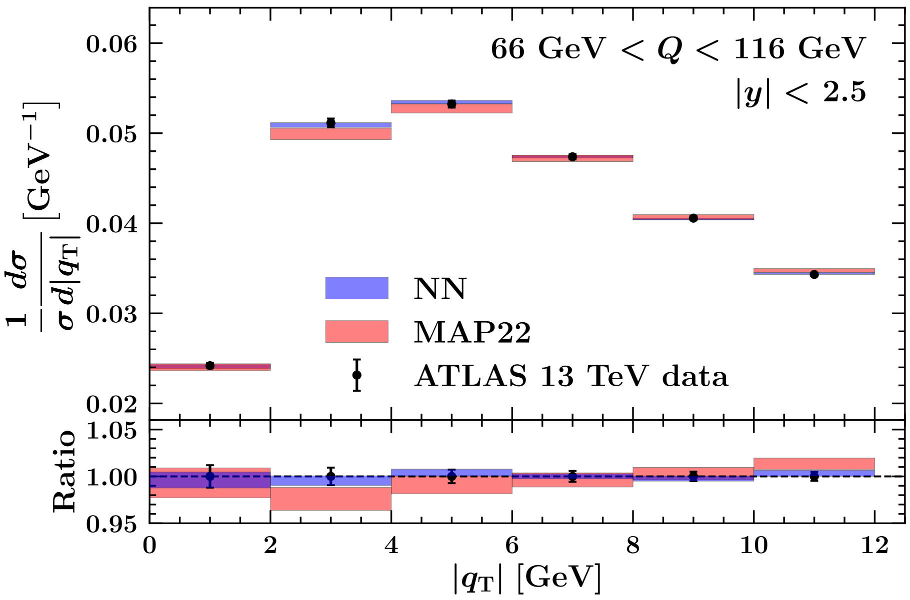
… just to compare again

N3LL

$\chi^2 = 0.97$

$\chi^2 = 1.28$





Description of data

numerical breakdown

$$\chi^2 = \underbrace{\sum_{i=1}^{n}\left(\frac{m_i - \bar{t}_i}{s_i}\right)^2}_{\text{uncorrelated contribution}} + \underbrace{\sum_{\alpha=1}^{k}\lambda_\alpha^2}_{\text{penalty term}}$$

Experiment	$N_{\rm dat}$	$\overline{\chi}^2$ $(\overline{\chi}^2_D + \overline{\chi}^2_\lambda)$ 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

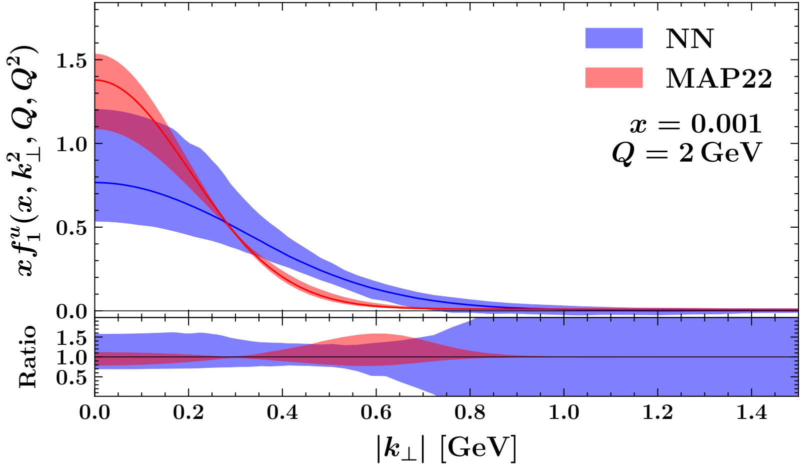
rule of thumb: it's best when it's small

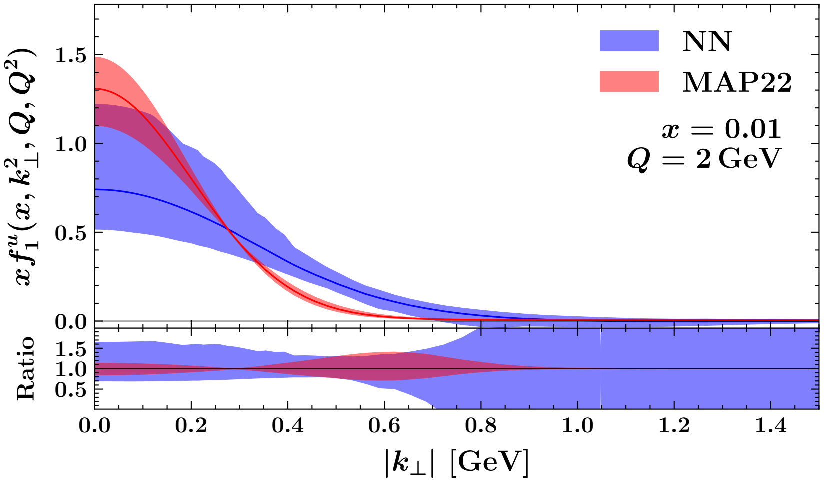
MAP22 relies more on the shifts with respect to the NN fit

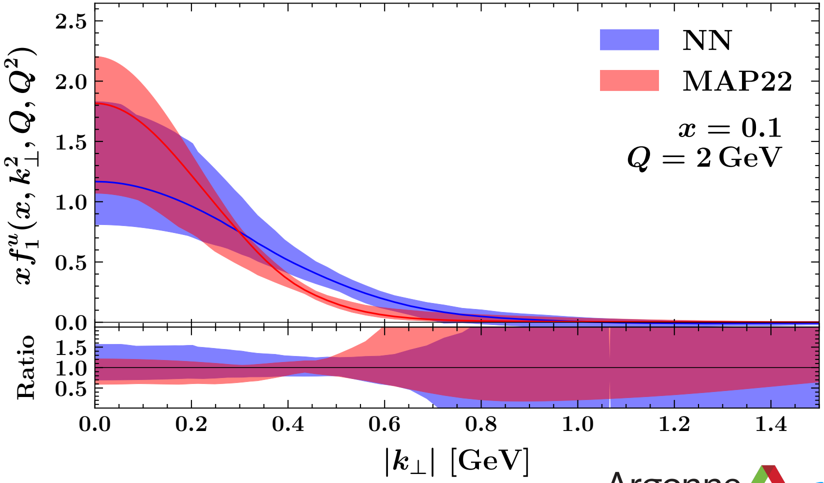
NN TMDs

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

MAP Collaboration
arXiv:2502.04166







NN fit has larger bands

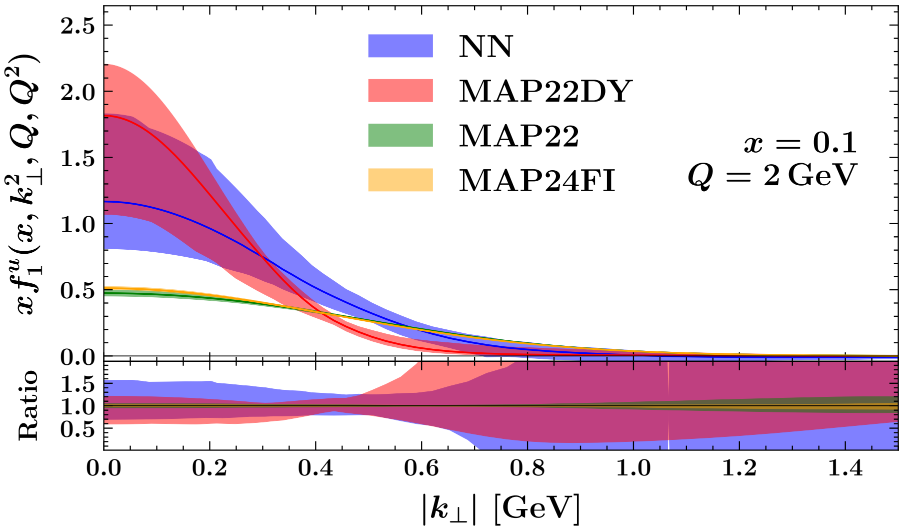
more reliable estimation uncertainties

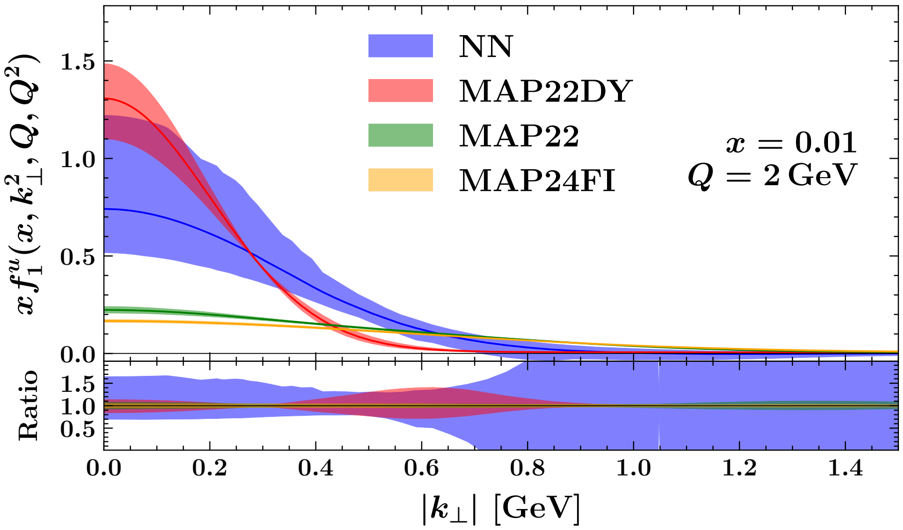
TMDs

comparison with other MAP fits

MAP22 and MAP24
full global fits, including SIDIS data

- narrower uncertainty bands
- wider TMDs

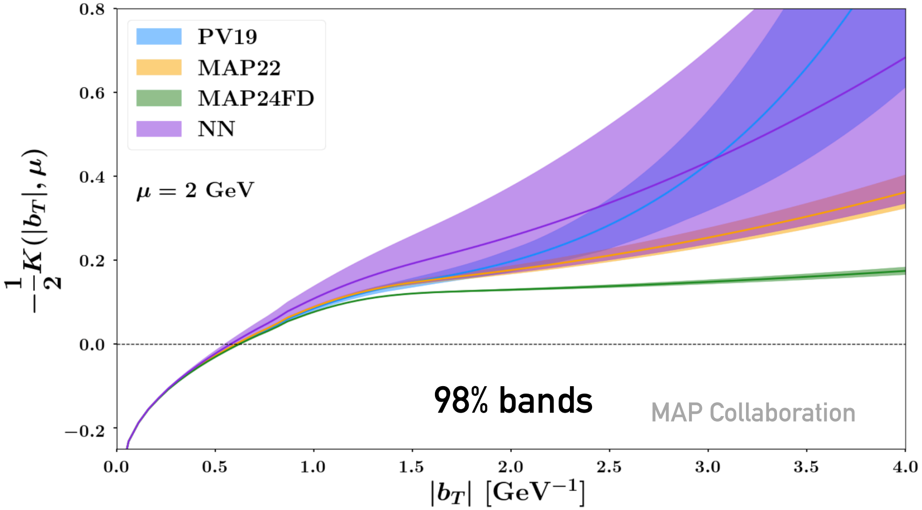


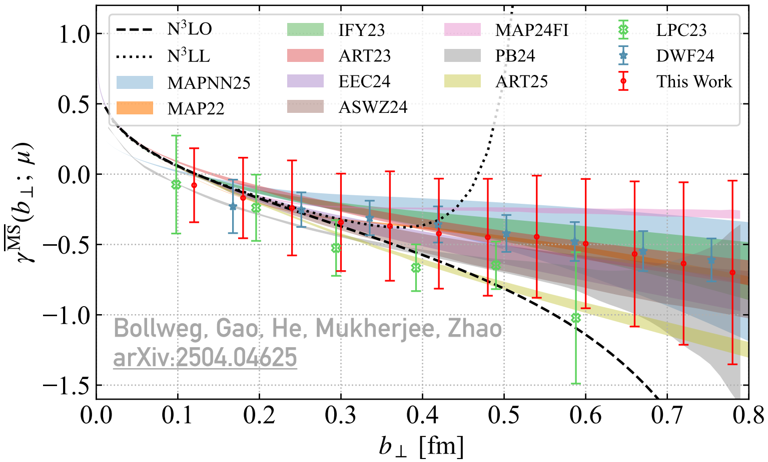


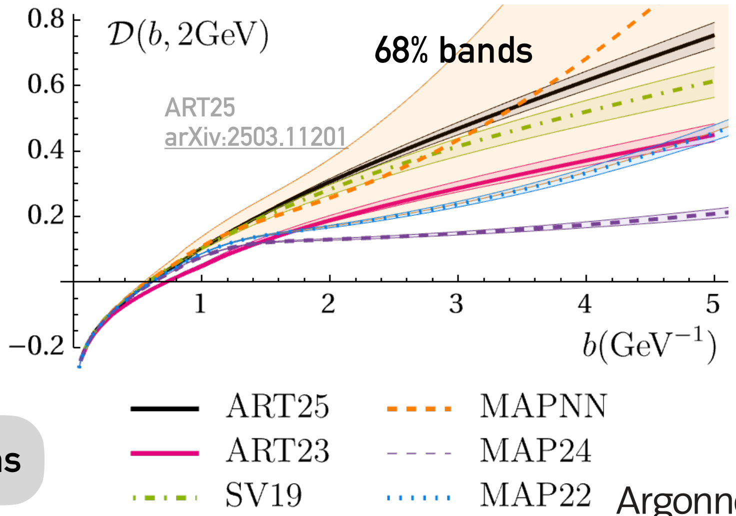
Collins-Soper kernel

“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



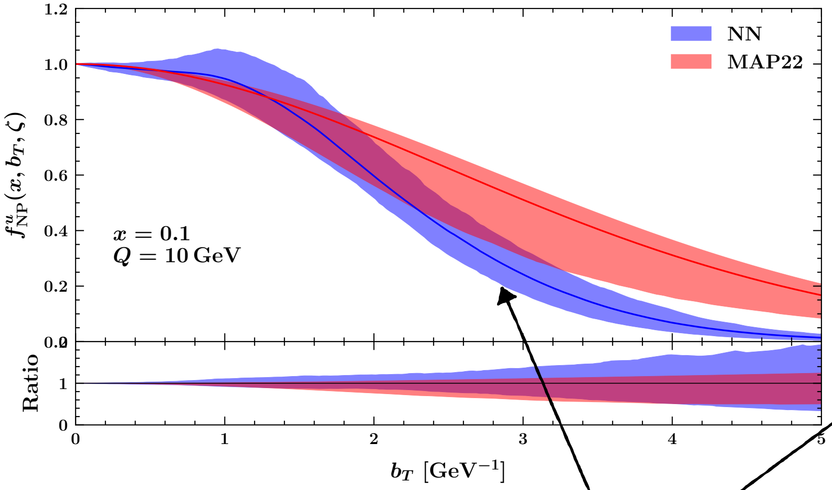


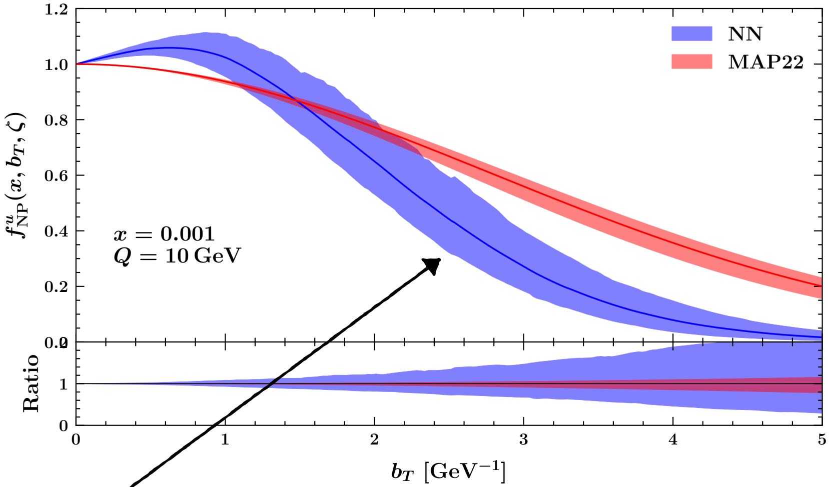


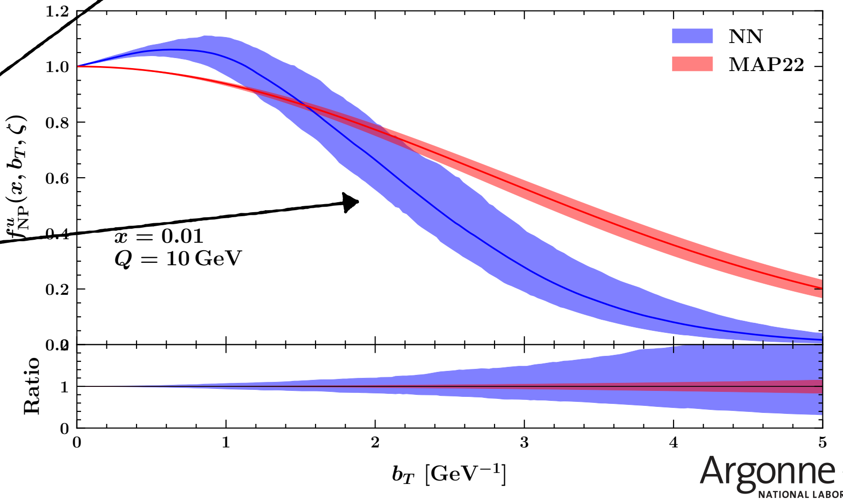
good agreement between the extractions

fNP

comparison with MAP22 only DY



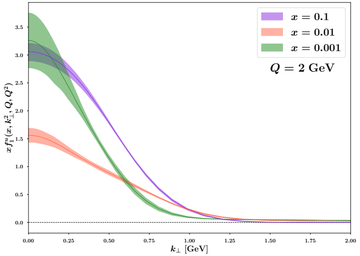




NN falls sharply with bT

Closure tests *only Drell-Yan data*

to validate the methodology



closure test – level 0

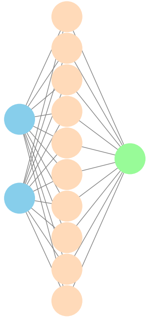
$$g_1 = 0.304$$
$$g_2 = 0.028$$

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

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!

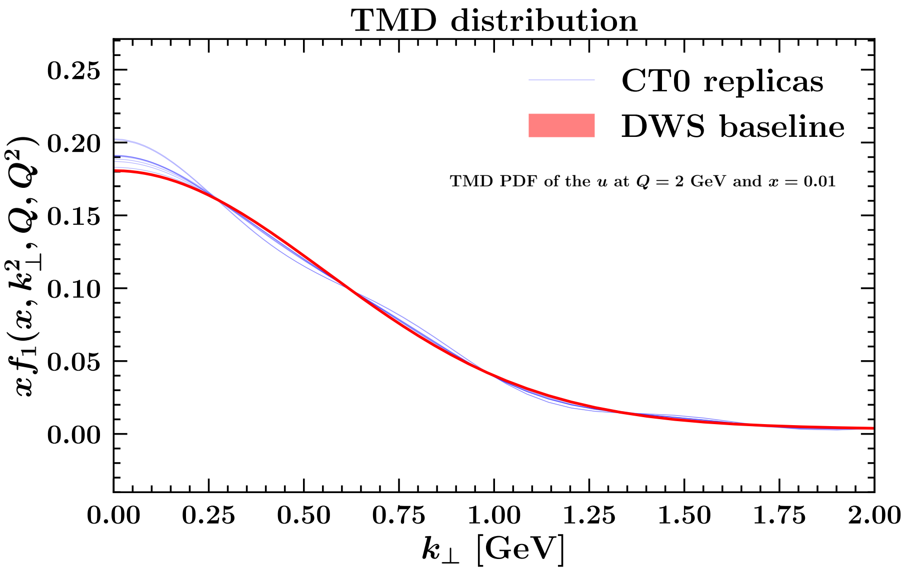
and for MAP22?

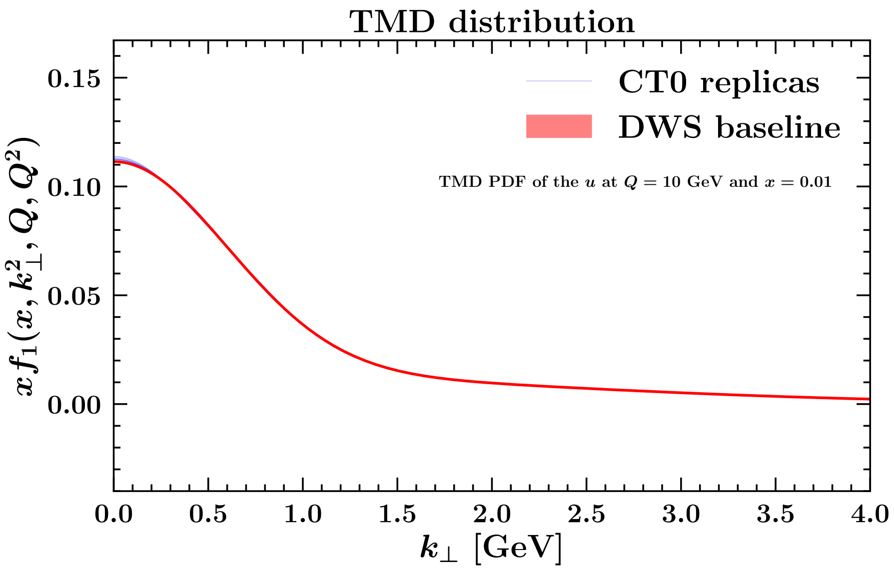
…work in progress

Closure tests

closure test – level 0

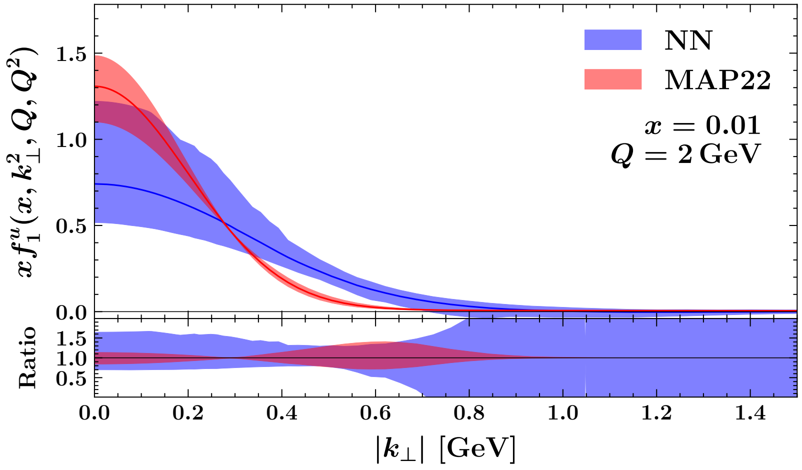
$$f_{\rm NP}^{\rm DWS}(b_T,\zeta) = \exp\left[-\frac{1}{2}\left(g_1 + g_2 \ln\left(\frac{\zeta}{2Q_0^2}\right)\right) b_T^2\right]$$





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

Conclusions



- **first Neural Networks extraction**

  of the unpolarized quark TMD PDFs in the proton

at N3LL ⊥ **Drell-Yan** data, 482 points

great description of data

$\chi^2 = 0.97$

more reliable uncertainties

OROS

the genesis of a new fitting framework

MAP Collaboration

ὄρος

— more flexible, modular and versatile

- suited for fitting all kinds distributions, PDFs, TMDs, GPDs
- very fast
- written in C++ but with a python wrapper

to do next:

- Study b* prescription - is there an impact on the extracted TMDs?
- Perform closure tests - validate the use of NN
- perform NN TMD fit for DY + SIDIS

Argonne
NATIONAL LABORATORY

TMD

every TMD has the same general structure

many subtleties involved in TMD analyses

b* prescription, ζ-prescription, logarithmic accuracy

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

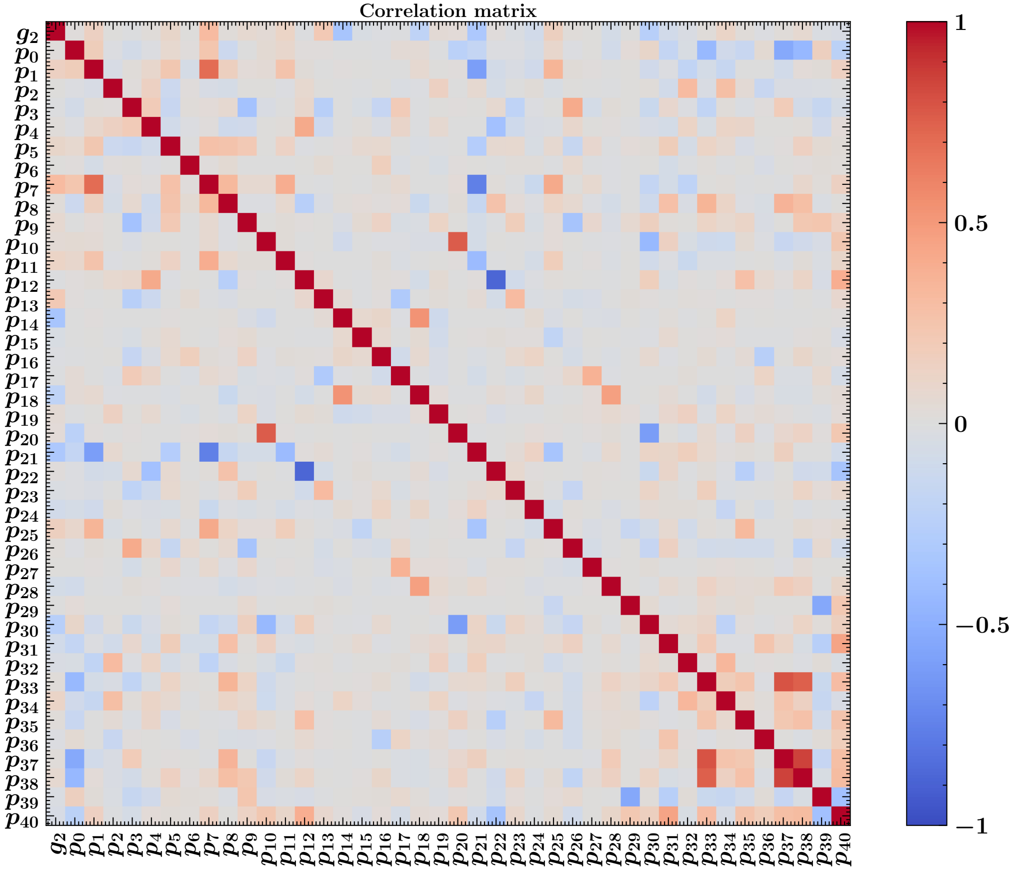
perturbative evolution

non perturbative transverse content

parametrized and fitted to data

correlation matrix

41 parameters



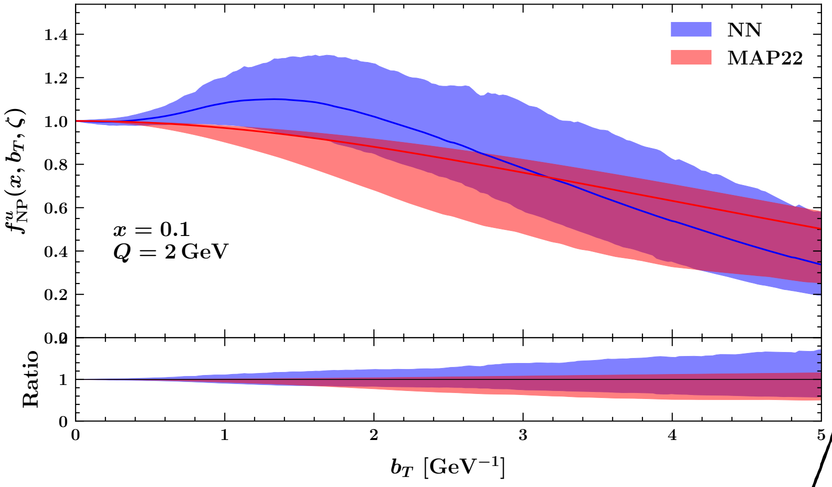
Closure tests

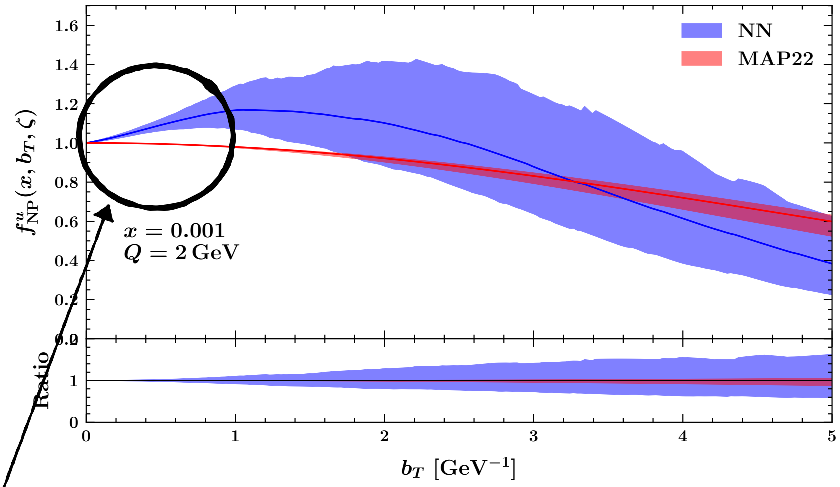
to validate the methodology

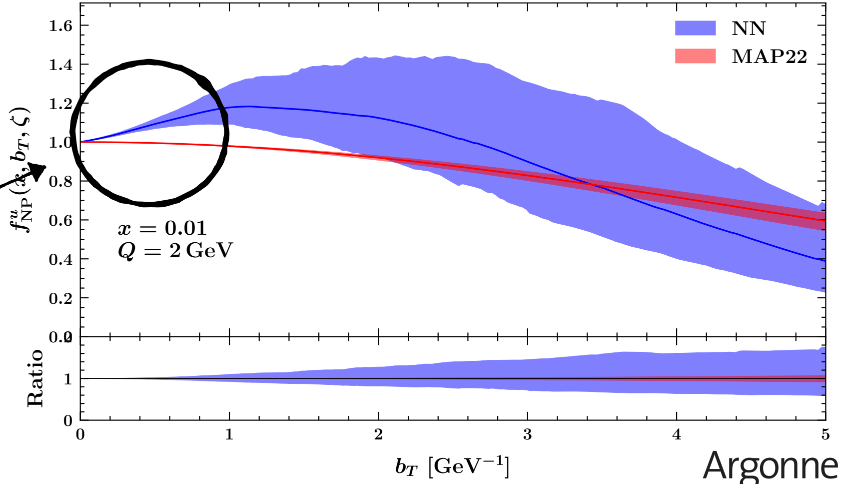
- **Level 0**
  - central pseudo-data is given by **central predictions** of the known model
  - **no Monte Carlo noise** is added on top of the central data — each replica is fitting the same set of data
- **Level 1**
  - central pseudo-data is shifted by some noise η drawn from the experimental covariance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

fNP

comparison with MAP22 only DY







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