TMD factorisation and phenomenology…

with a sprinkle of GPDs

Valerio Bertone

IRFU, CEA, Université Paris-Saclay



October 20, 2026, Lecture, REF 2026

A "constructive" introduction

- Let us start with a generic **bi-local** (quark) operator:

$$\mathcal{O} = \overline{\psi}(b)\,\Gamma\psi(0)$$

- Γ is generic Dirac structure, *i.e.* a linear combination of $\{\mathbb{I}, \gamma^5, \gamma^\mu, \gamma^5\gamma^\mu, \sigma^{\mu\nu}\}$.

$$\Gamma = A\mathbb{I} + B\gamma^5 + C_\mu\gamma^\mu + D_\mu\gamma^5\gamma^\mu + E_{\mu\nu}\sigma^{\mu\nu}$$

- In order to give $\mathcal{O}$ a physical meaning, we need to make it gauge invariant.
  - Introduce the parallel-transport operator W (often called **Wilson line** in this context):

$$W(y,x) = \mathcal{P}\exp\left[-igt^a\int_y^x dz^\mu A_\mu^a(z)\right]$$

- The gauge invariant version of $\mathcal{O}$ is then:

$$\mathcal{O} = \overline{\psi}(b)\,\Gamma\,W(b,0)\psi(0)$$

- Now consider the case in which $\mathcal{O}$ is **highly boosted** along $-z$ (as if it was involved in a high-energy collision): this frame is called infinite-momentum frame (IMF).
- Working in the IMF has **two important consequences**:
  - $b_z \simeq -cb_t$, therefore in light-cone coordinates $b \simeq (0, b^-, \mathbf{b}_T)$.
  - In addition, the coefficients $\{A, B, C_\mu, D_\mu, E_{\mu\nu}\}$ get enhanced, unchanged, or suppressed:
    - C_+, D_+, E_{+i} enhanced (twist 2), $A, B, C_i, D_i, E_{ij}, E_{+-}$ unchanged (twist 3), C_-, D_-, E_{-i} suppressed (twist 4).

A "constructive" introduction

- As an example let us consider the "unpolarised" operator:

$$\mathcal{O} = \overline{\psi}(b)\,\gamma^+\,W(b,0)\psi(0)\big|_{b^+=0}$$

- To connect this operator to an observable we need to take a matrix element.
- We bracket it between two, generally different hadronic states:

$$\mathcal{M} = \langle H'(p',\lambda')|\overline{\psi}(b)\,\gamma^+\,W(b,0)\psi(0)\,|H(p,\lambda)\rangle\big|_{b^+=0}$$

- Finally, it is usually more phenomenologically relevant to study the **momentum** behaviour of any such matrix element. We thus take its Fourier transform:

$$\Phi = \int db^- d^2\mathbf{b}_T\, e^{ib^-k^+ - i\mathbf{b}_T\cdot\mathbf{k}_T}\ \langle H'(p',\lambda')|\overline{\psi}(b)\,\gamma^+\,W(b,0)\psi(0)\,|H(p,\lambda)\rangle\big|_{b^+=0}$$

- This is a sketched definition of **generalised transverse-momentum dependent** (GTMD) correlation (the Wilson line needs to be chosen appropriately).
- GTMDs can be regarded as "**mother distributions**" (*cit.* Meißner, Metz, Schlegel [*JHEP* 08 (2009) 056])
- They encode "the most general **one-body** information of partons, corresponding to the full one-quark density matrix in momentum space" (*cit.* Lorcé, Parquini, Vanderhaeghen [*JHEP* 05 (2011) 041]).

Further readings: Ji [*Phys.Rev.Lett.* 91 (2003) 062001], Belitsky, Ji, Yuan [*Phys.Rev.D* 69 (2004) 074014], Belitsky, Radyushkin [*Phys.Rept.* 418 (2005) 1-387]

A "constructive" introduction

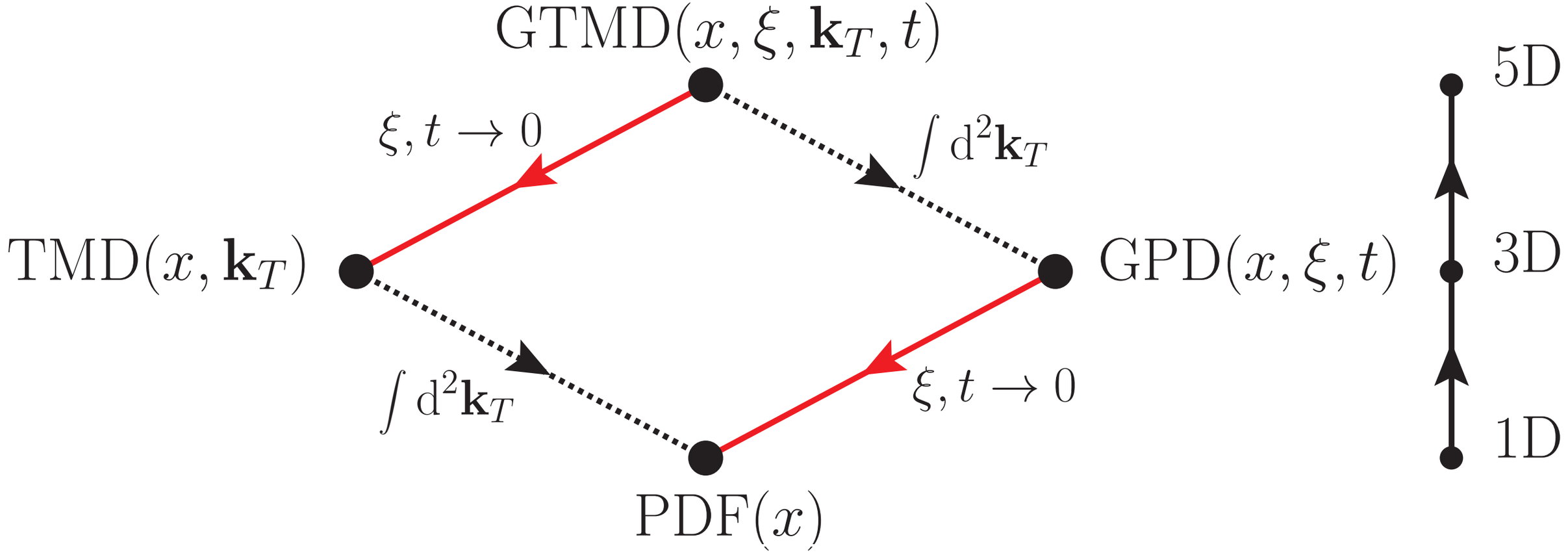
- Many relevant hadronic distributions in high-energy physics can be made descend from GTMDs. Introducing the definitions:

$$P = \frac{p + p'}{2} \qquad \Delta \equiv p - p'$$

- A common set of kinematic variables used to parameterise GTMDs is:

$$k^+ \equiv x P^+ \qquad \Delta^+ \equiv \xi \frac{P^+}{2} \qquad t = \Delta^2$$

- A (partial) genealogy of GTMDs then looks like this:



- All of these distributions are related to physical observables through **factorisation**.

A "constructive" introduction

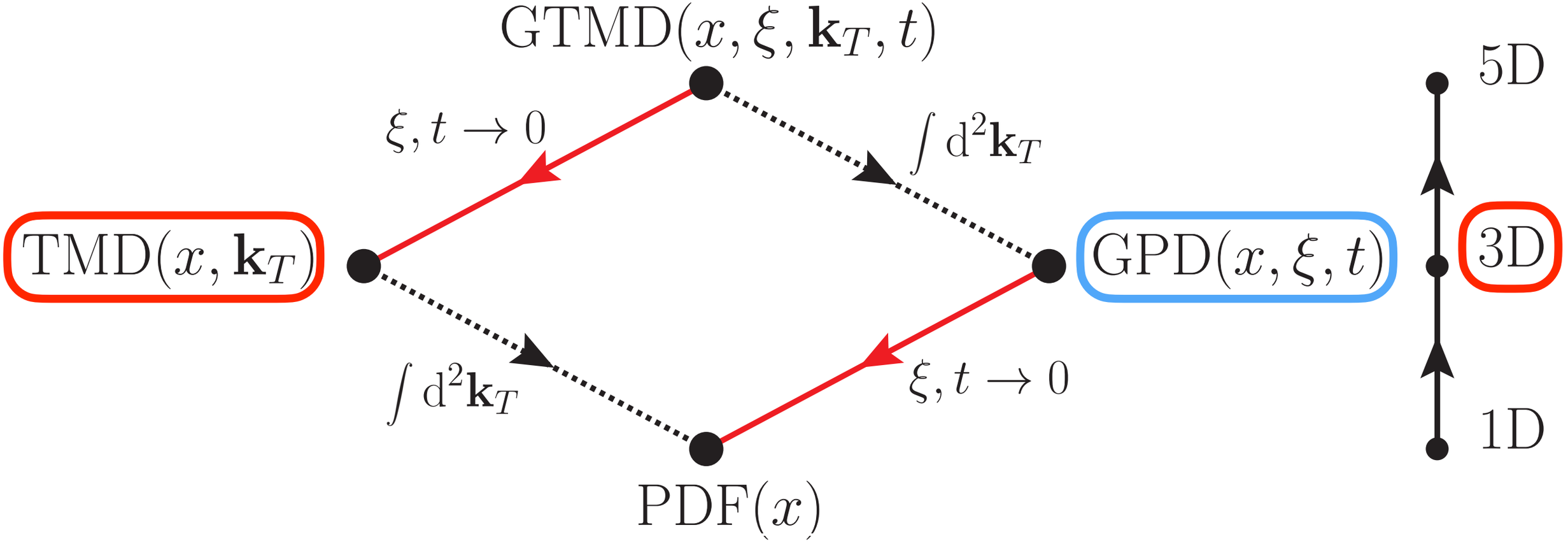
- Many relevant hadronic distributions in high-energy physics can be made descend from GTMDs. Introducing the definitions:

$$P = \frac{p + p'}{2} \qquad \Delta \equiv p - p'$$

- A common set of kinematic variables used to parameterise GTMDs is:

$$k^+ \equiv x P^+ \qquad \Delta^+ \equiv \xi \frac{P^+}{2} \qquad t = \Delta^2$$

- A (partial) genealogy of GTMDs then looks like this:

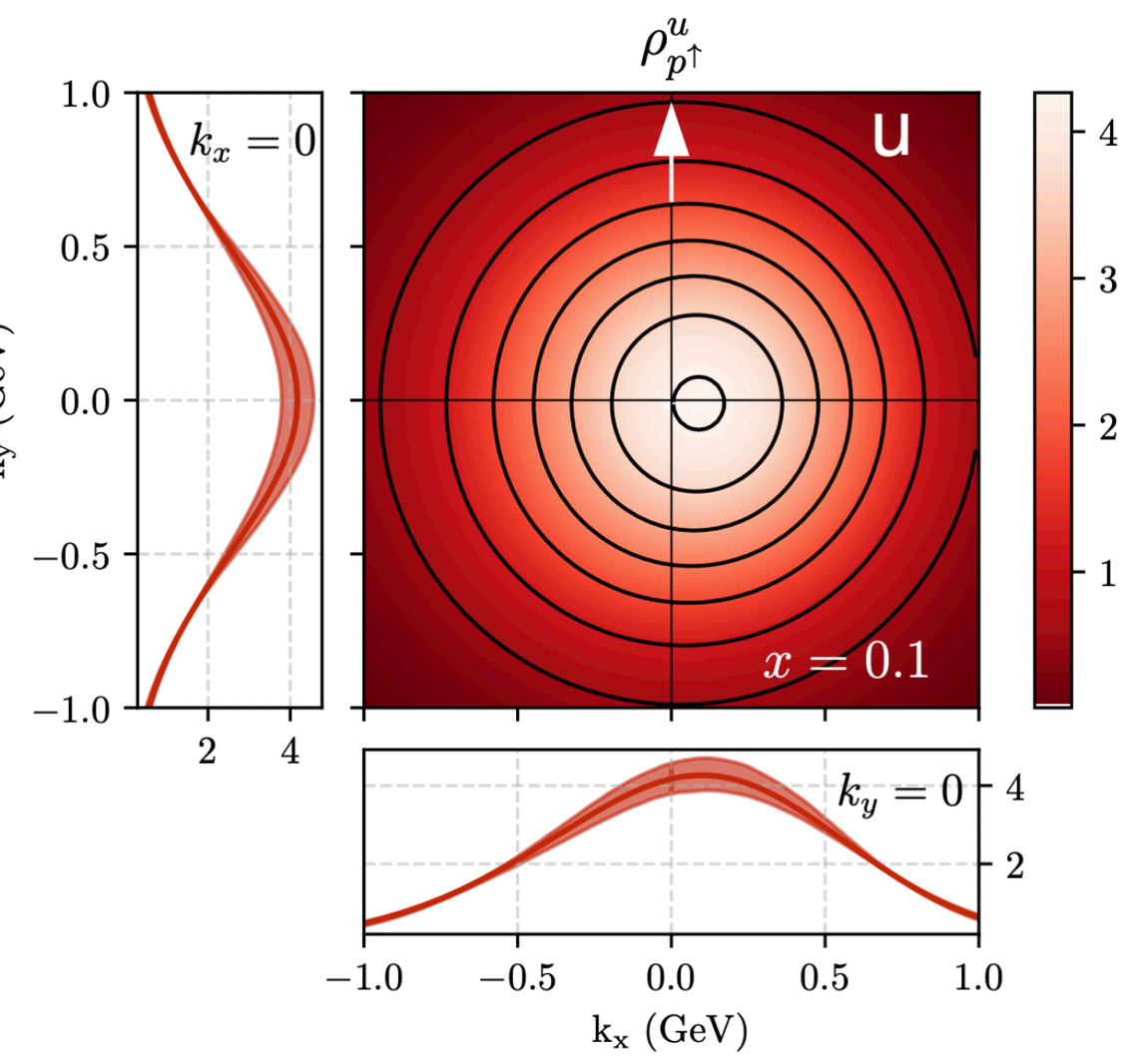


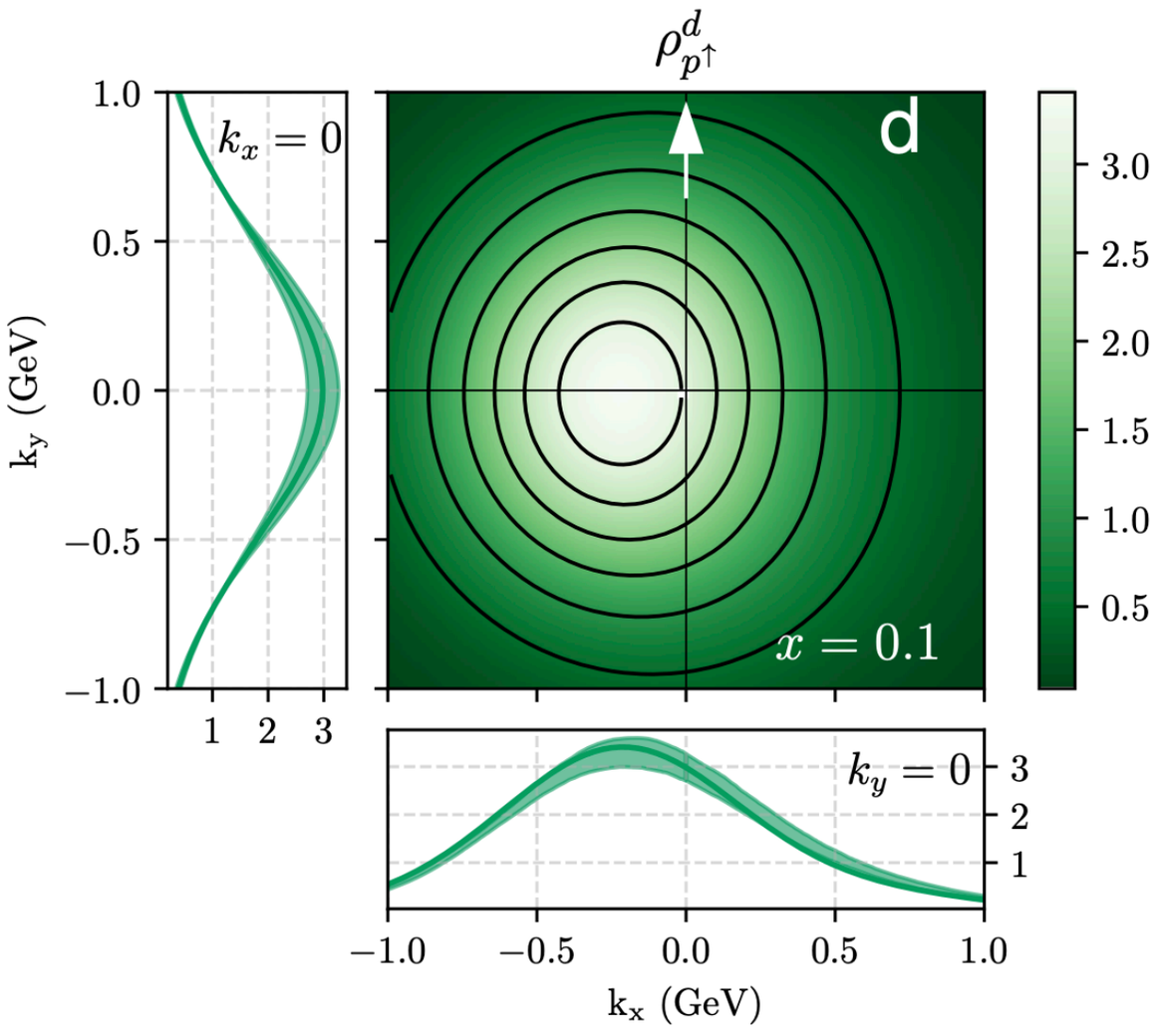
- All of these distributions are related to physical observables through **factorisation**.

Momentum-space tomography

- TMDs retain the intrinsic transverse momentum of the parton:

$$\int \frac{dz^- d^2\mathbf{z}_T}{2(2\pi)^3}\, e^{ik\cdot z} \langle P,S|\,\bar{\psi}(0)\gamma^+ \mathcal{W}_{\sqsupset}(0,z)\,\psi(z)|P,S\rangle\Big|_{z^+=0} = f_1(x,k_T^2) - \frac{\epsilon_T^{ij} k_T^i S_T^j}{M}\, f_{1T}^{\perp}(x,k_T^2)$$





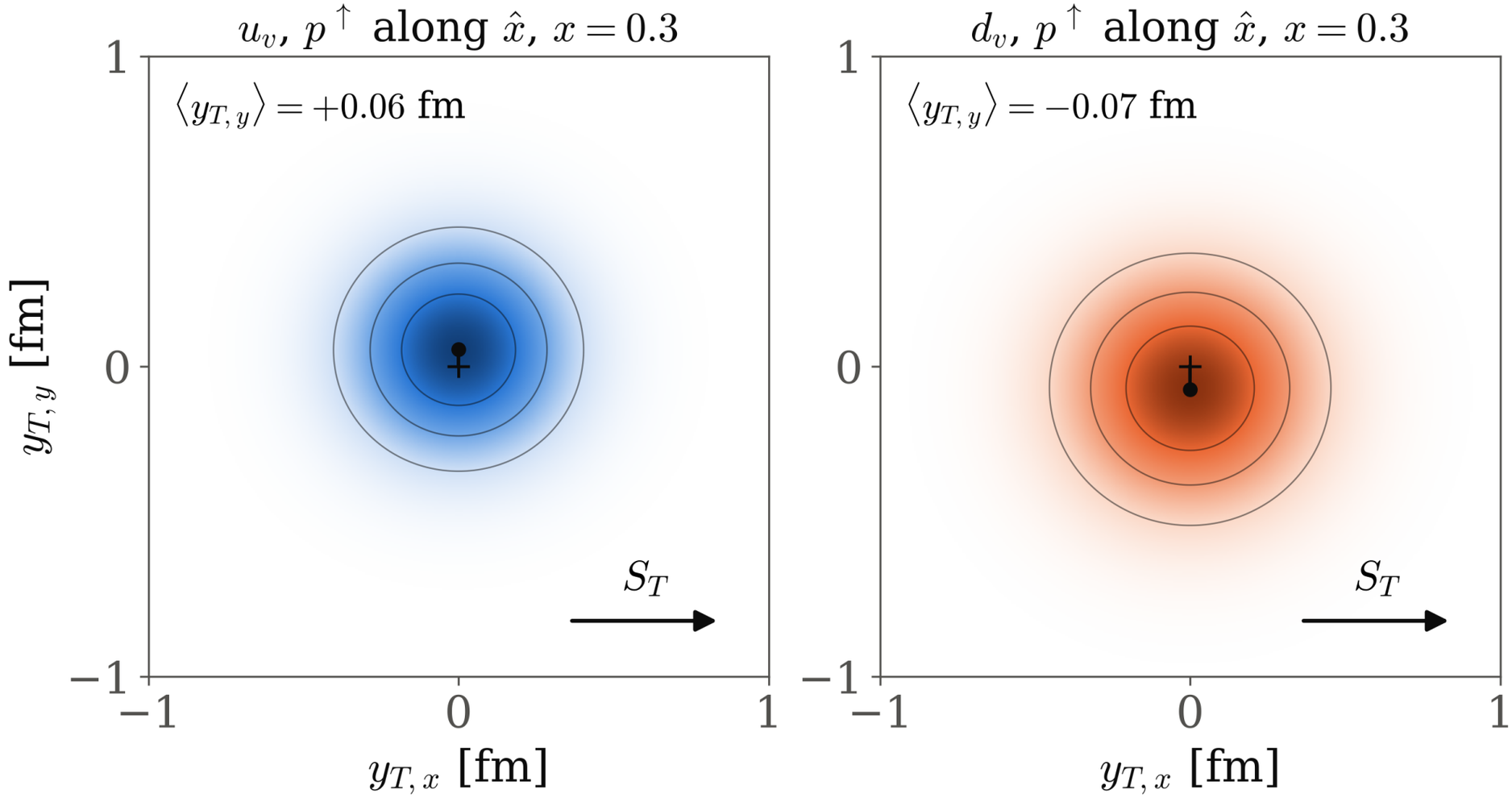
[*Phys.Lett.B* 827 (2022) 136961]

Position-space tomography

- In [*Phys.Rev.D* 62 (2000) 071503, Int.J.Mod.Phys.A 18 (2003) 173-208] it was proven that the quantity

$$f(x, \mathbf{y}_T) = \int \frac{d^2\boldsymbol{\Delta}_T}{(2\pi)} e^{i\boldsymbol{\Delta}_T \cdot \mathbf{y}_T} \mathrm{GPD}(x, \boldsymbol{\Delta}_T, \xi = 0)$$

- (a.k.a. impact-parameter PDFs) produce the PDFs as functions of the distance $\mathbf{y}_T$ from the c.o.m. of the hadron.



Position-space tomography

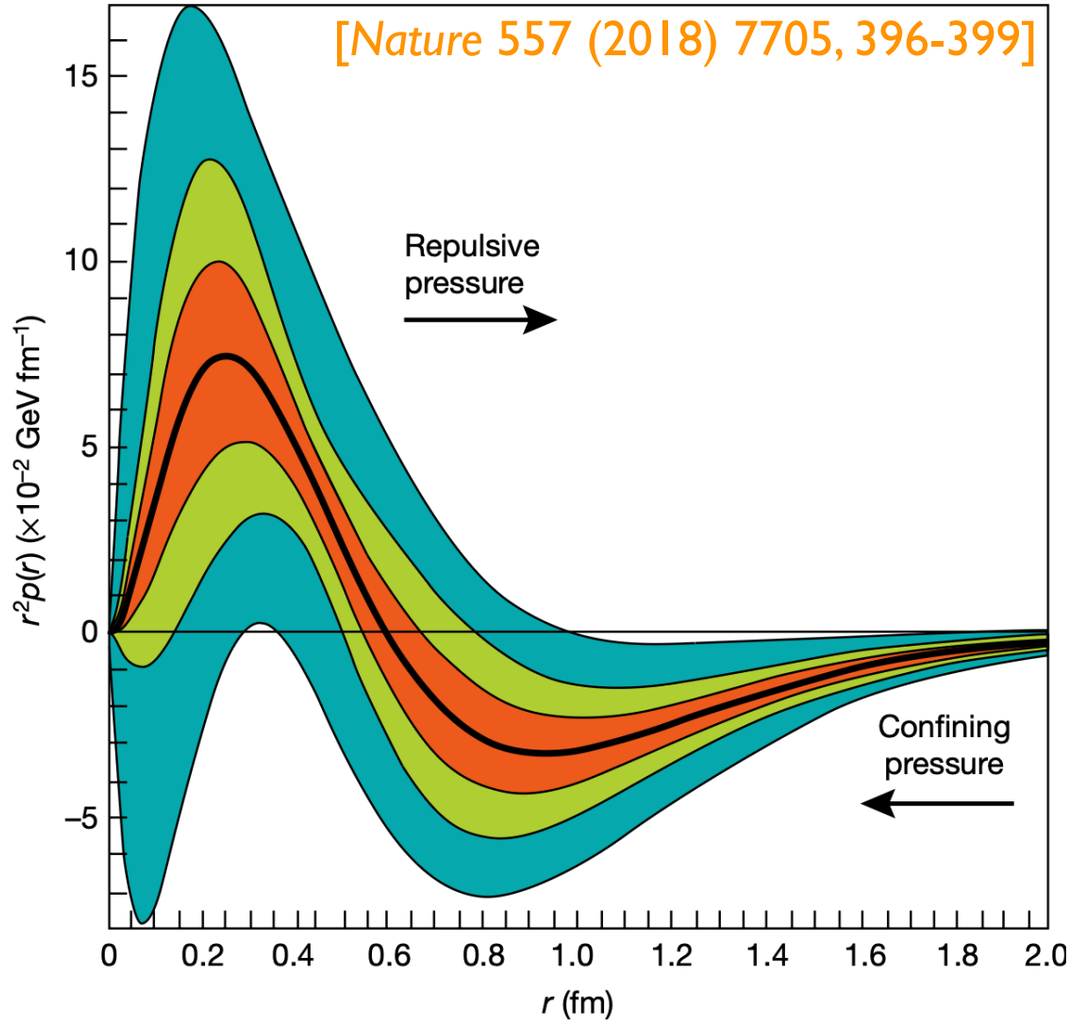
- GPDs are also related to the energy-momentum tensor (EMT):

$$\langle p'|T_a^{\mu\nu}|p\rangle = \bar{u}(p')\Big[A_a\frac{\gamma^{\{\mu}P^{\nu\}}}{2} + B_a\frac{iP^{\{\mu}\sigma^{\nu\}\rho}\Delta_\rho}{4M} + D_a\frac{\Delta^\mu\Delta^\nu - g^{\mu\nu}\Delta^2}{4M} + \bar{c}_a M g^{\mu\nu}\Big]u(p)$$

- Form factors are second moments of GPDs:

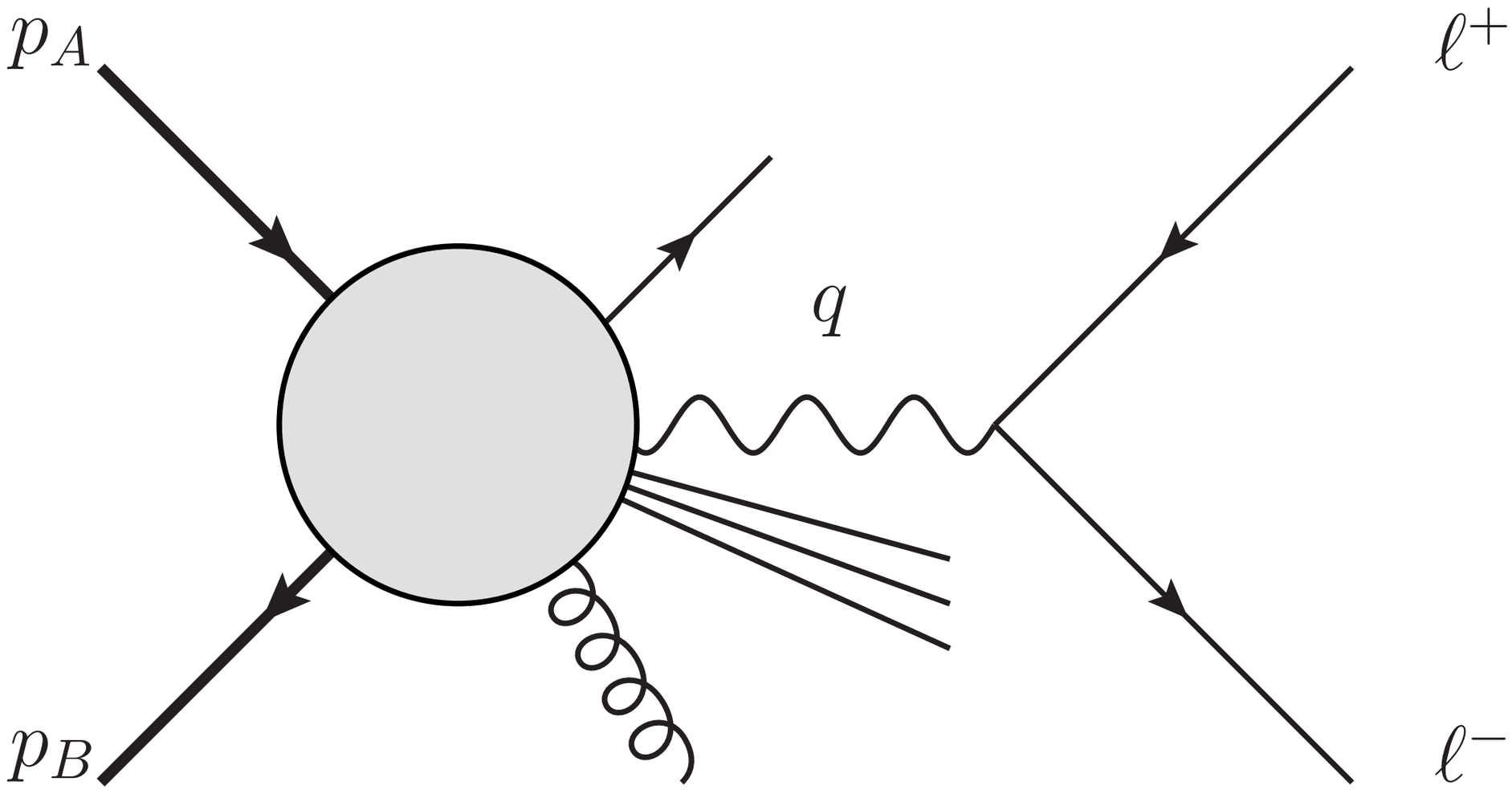
$$\int dx\, x\, H^q = A^q(t) + \xi^2 D^q(t), \qquad \int dx\, x\, E^q = B^q(t) - \xi^2 D^q(t)$$

- The D-term encodes the pressure and shear forces inside the proton:



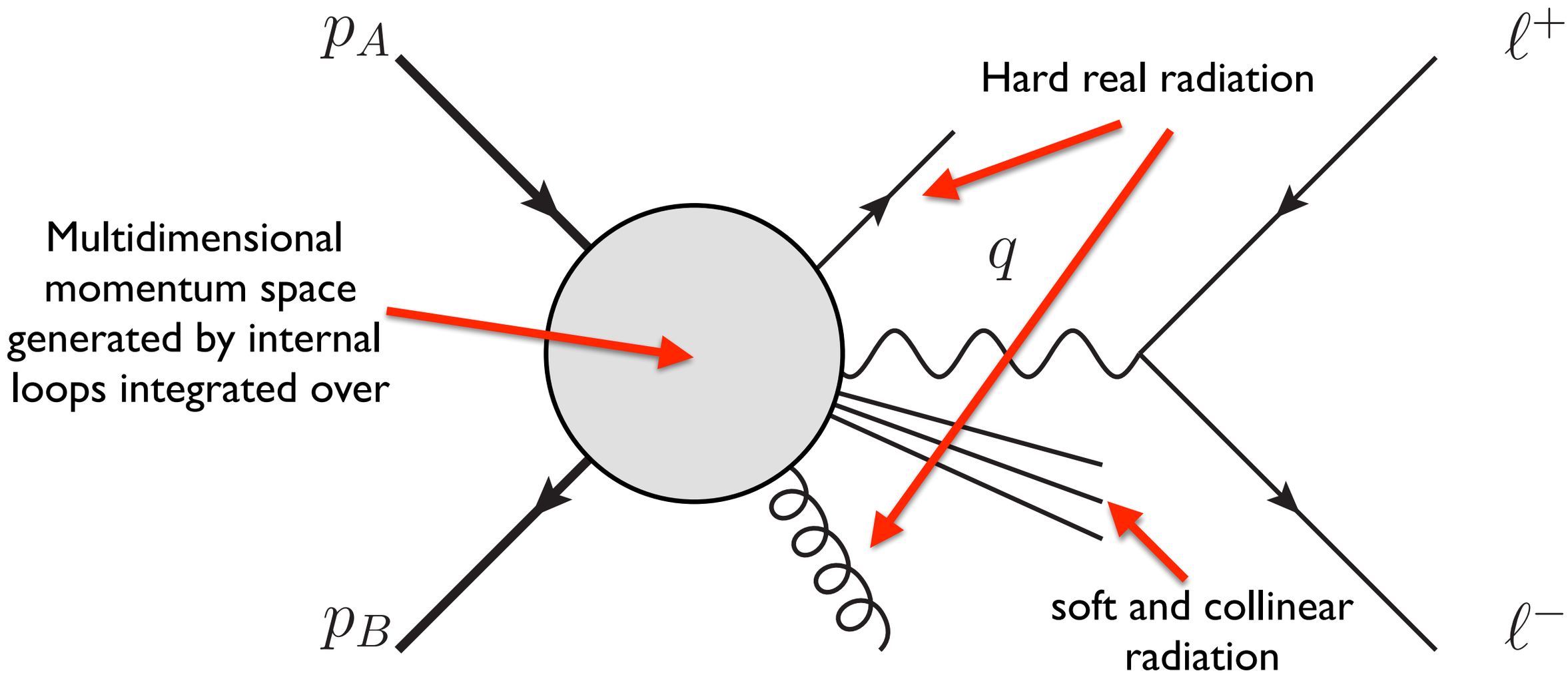
Factorisation

- Let us use Drell-Yan production (*i.e. inclusive* production of a lepton pair in hadron-hadron collisions) to sketch the main steps of factorisation:



Factorisation

- Let us use Drell-Yan production (*i.e. inclusive* production of a lepton pair in hadron-hadron collisions) to sketch the main steps of factorisation:

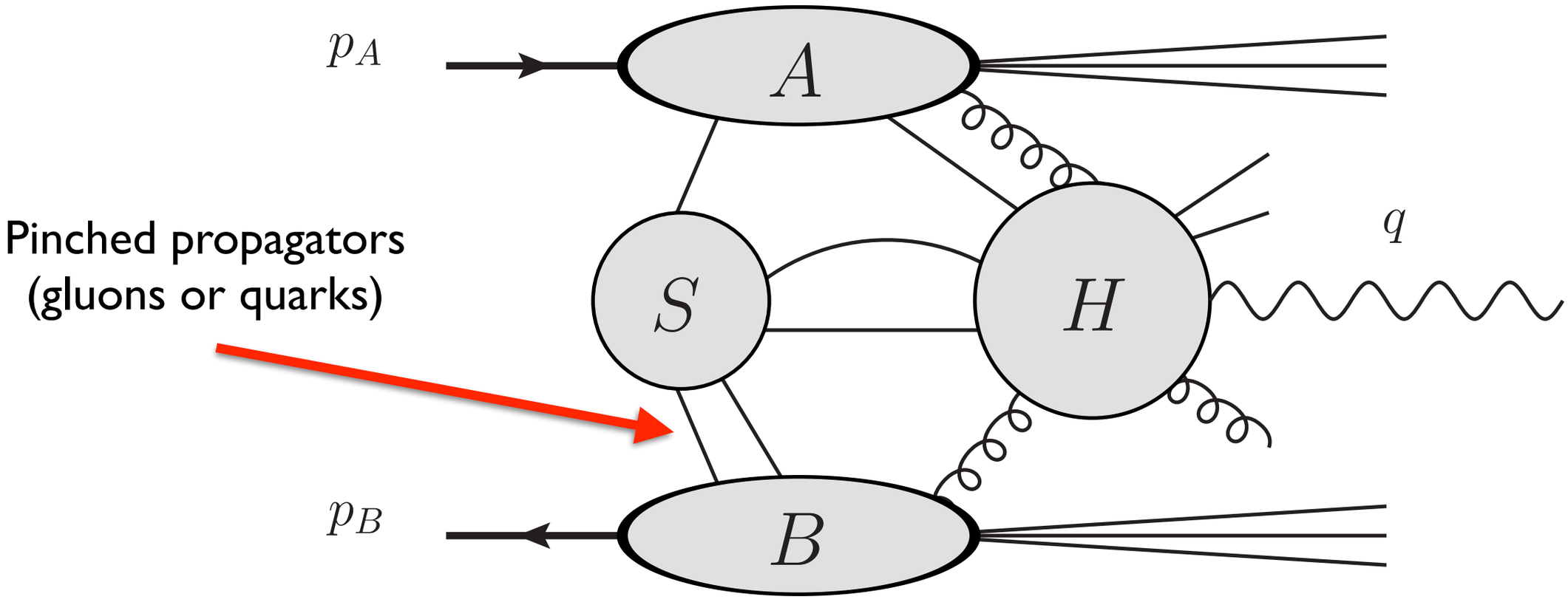


- Relevant scales:

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

Factorisation

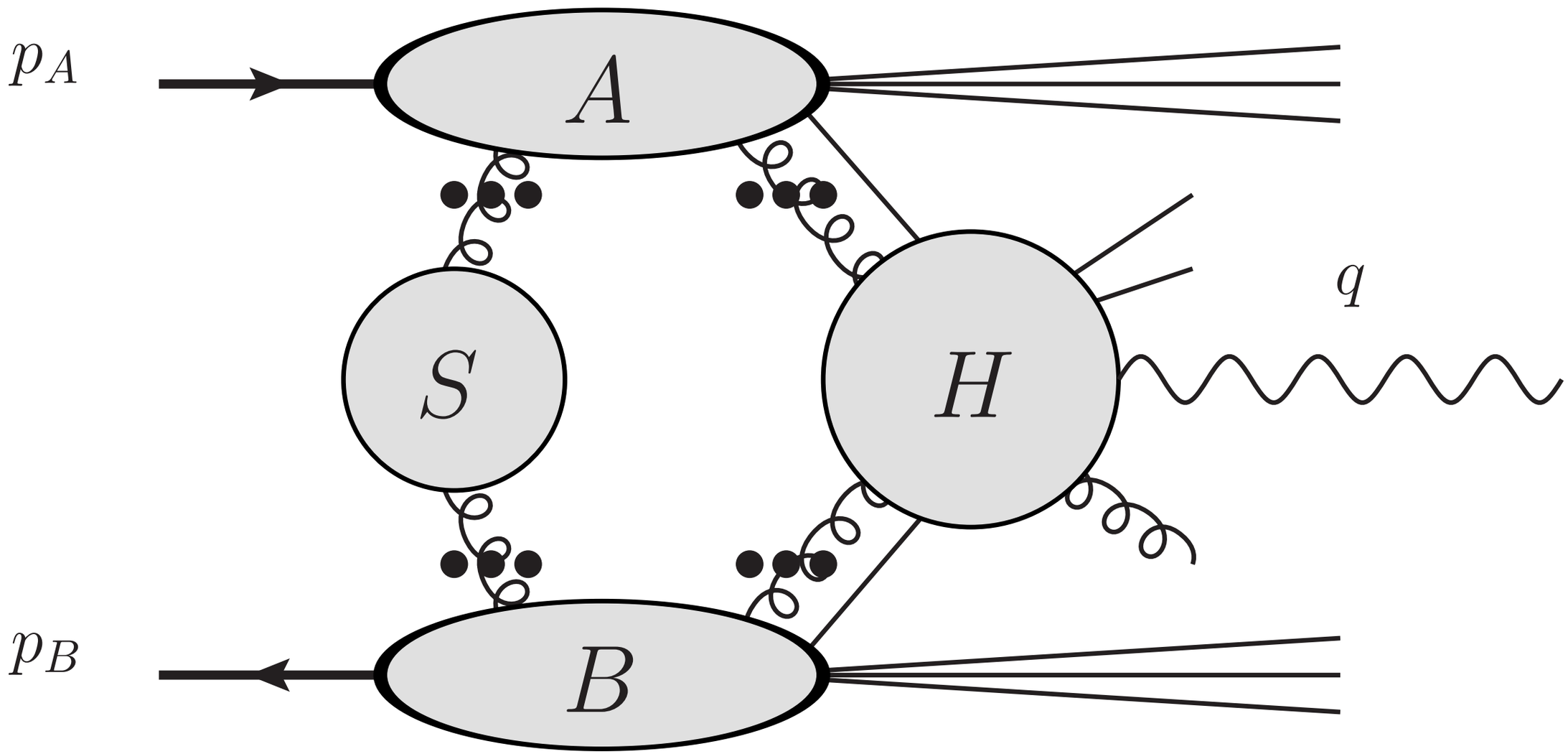
- Let us work in **massless** QCD and, for each *single* graph, identify the regions in the integration momentum space where the integrand is large:
  - **pinched** propagators ⇒ Landau criterion and **reduced graphs**,
    - soft modes (**S**),
    - (anti)collinear modes (**A** and **B**),
    - hard modes (**H**) (not pinched but large integration region).



Factorisation

[G. Sterman, Phys. Rev. D 17, 2773]

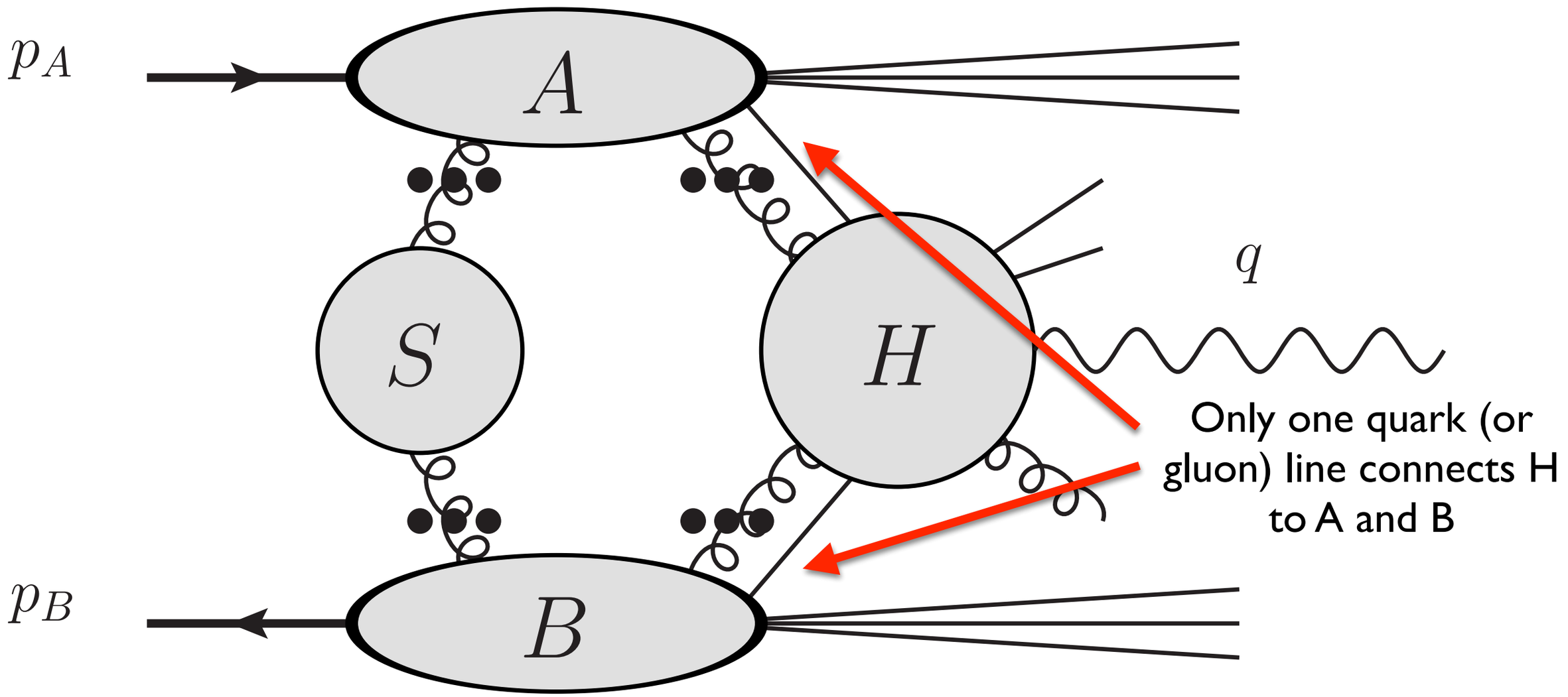
- Apply Libby-Sterman power counting to identify the **asymptote** in Q:
- The result (in a **covariant gauge**) is strikingly simple:



Factorisation

[G. Sterman, Phys. Rev. D 17, 2773]

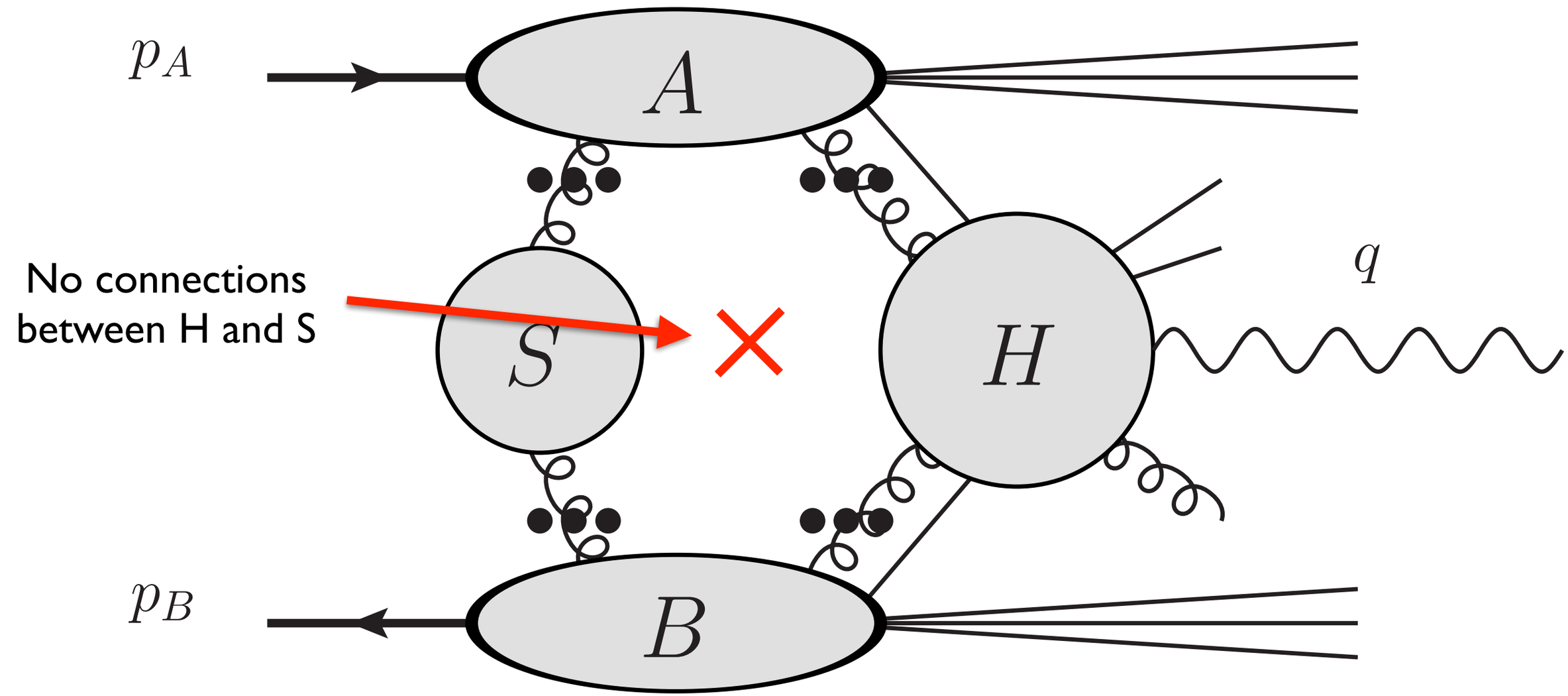
- Apply Libby-Sterman power counting to identify the **asymptote** in Q:
- The result (in a **covariant gauge**) is strikingly simple:



Factorisation

[G. Sterman, Phys. Rev. D 17, 2773]

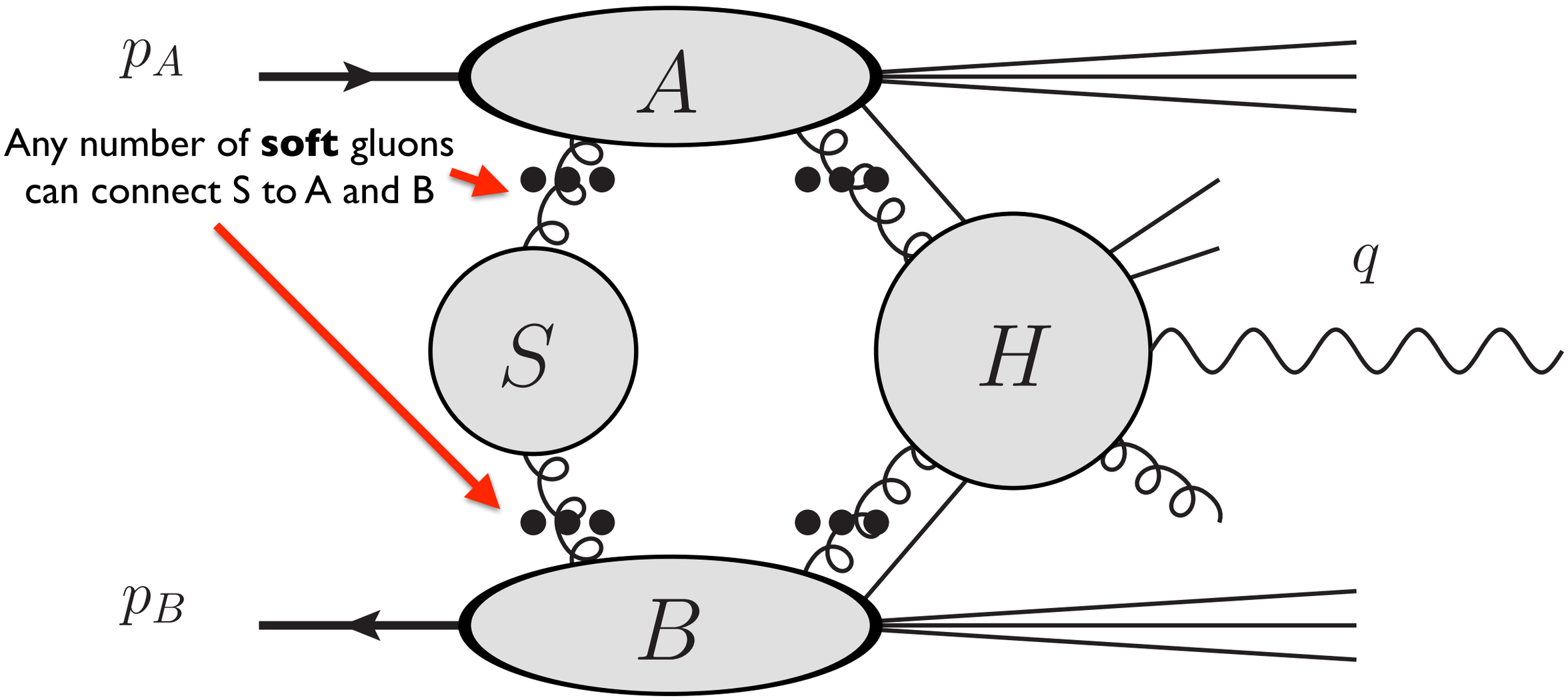
- Apply Libby-Sterman power counting to identify the **asymptote** in Q:
- The result (in a **covariant gauge**) is strikingly simple:



Factorisation

[G. Sterman, Phys. Rev. D 17, 2773]

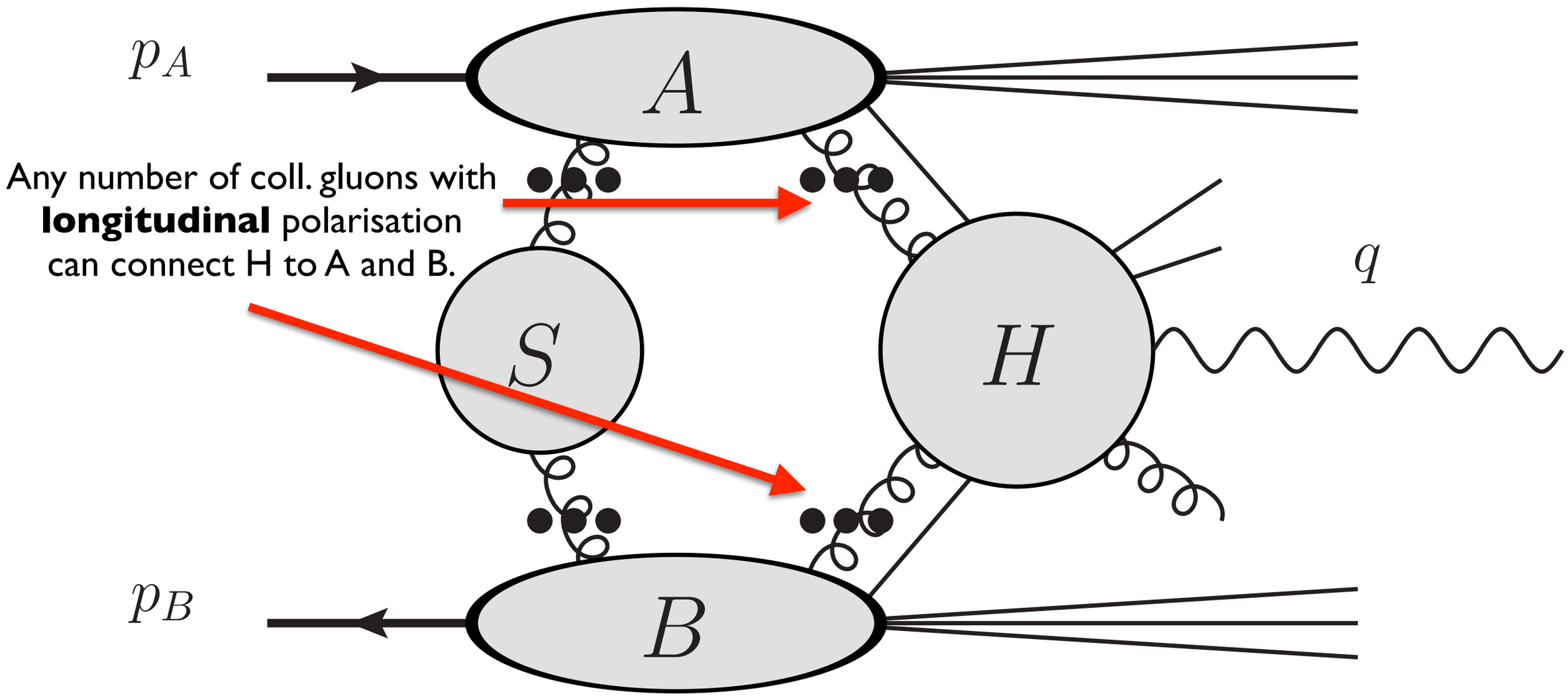
- Apply Libby-Sterman power counting to identify the **asymptote** in Q:
- The result (in a **covariant gauge**) is strikingly simple:



Factorisation

[G. Sterman, Phys. Rev. D 17, 2773]

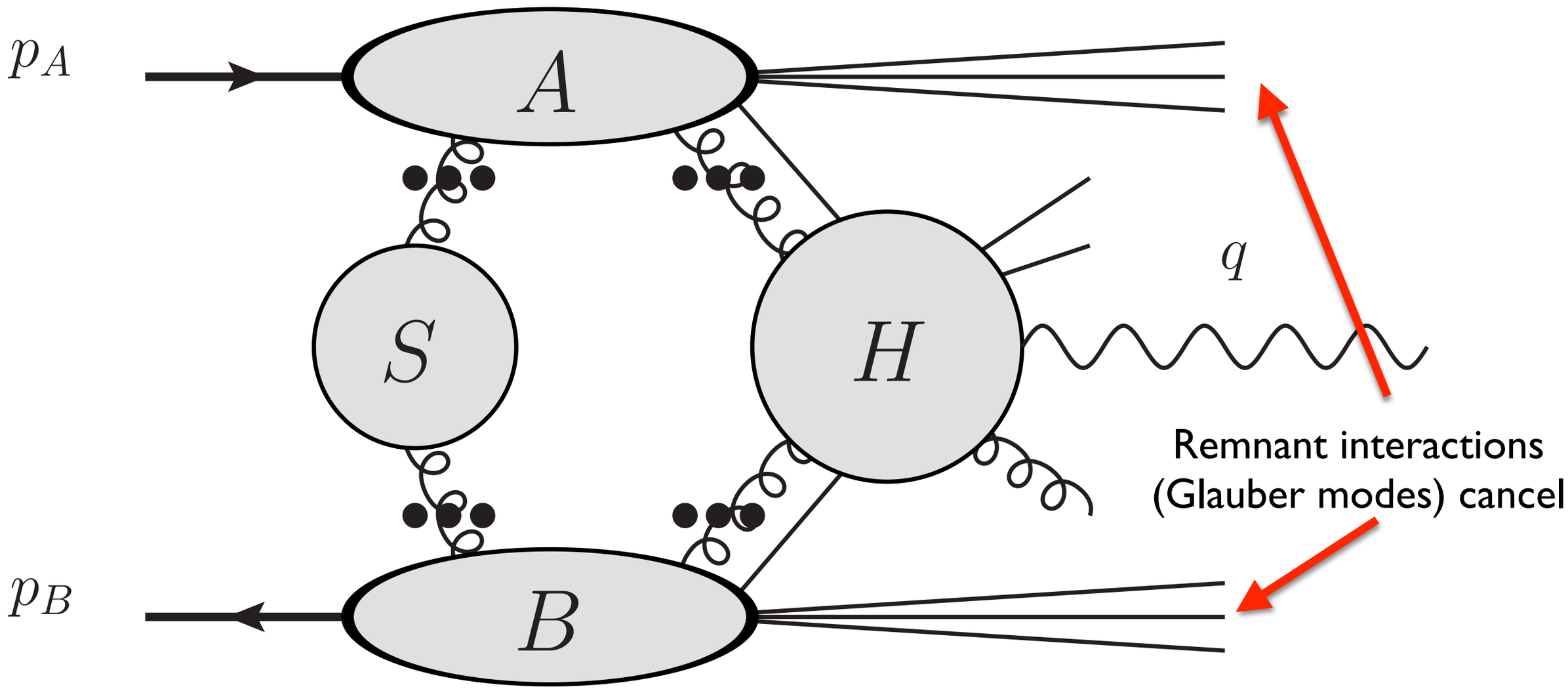
- Apply Libby-Sterman power counting to identify the **asymptote** in Q:
- The result (in a **covariant gauge**) is strikingly simple:



Factorisation

[G. Sterman, Phys. Rev. D 17, 2773]

- Apply Libby-Sterman power counting to identify the **asymptote** in Q:
- The result (in a **covariant gauge**) is strikingly simple:



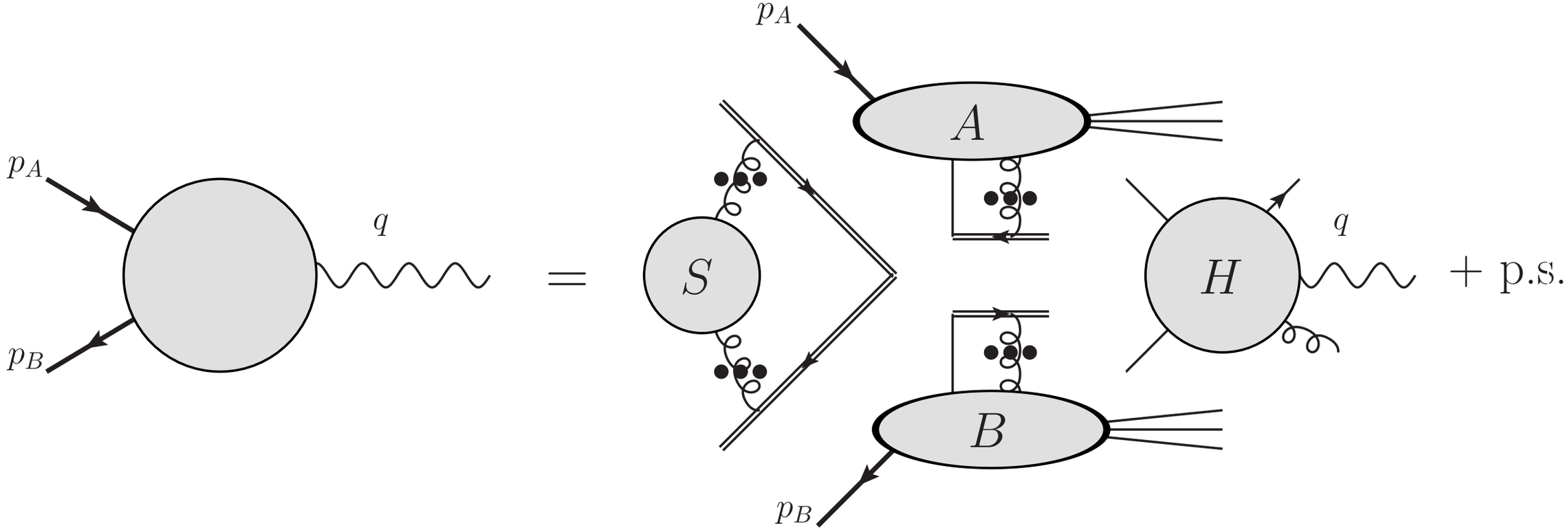
Factorisation

- Finally, factorisation is achieved by:
  - using **eikonal approximations** for each soft/collinear gluon attachment with momentum k to write, for example:

$$S^{\mu\cdots} g_{\mu\nu} A^{\nu\cdots} \sim (k \cdot S^{\cdots}) \frac{1}{n \cdot k} (n \cdot A^{\cdots})$$

  - this allows one to use **Ward identities** and introduces **Wilson lines**.

- A recursive application leads to factorisation of the amplitude:

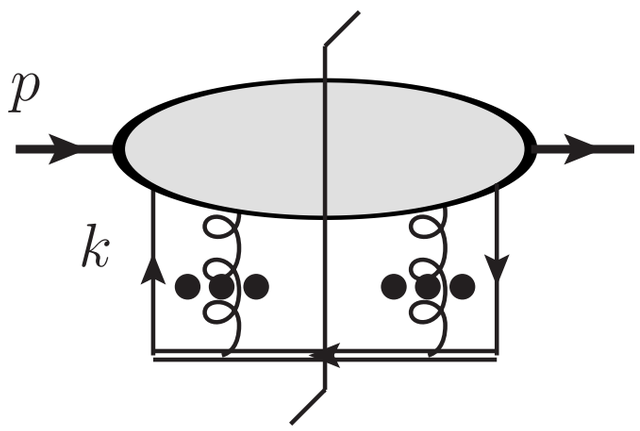


$$\mathcal{A} \sim H \cdot A \cdot B \cdot S$$

Factorisation

- Upon squaring, factorisation leads to the operator definition of:
  - gauge invariant (unsubtracted) parton-distribution function (**PDF**):

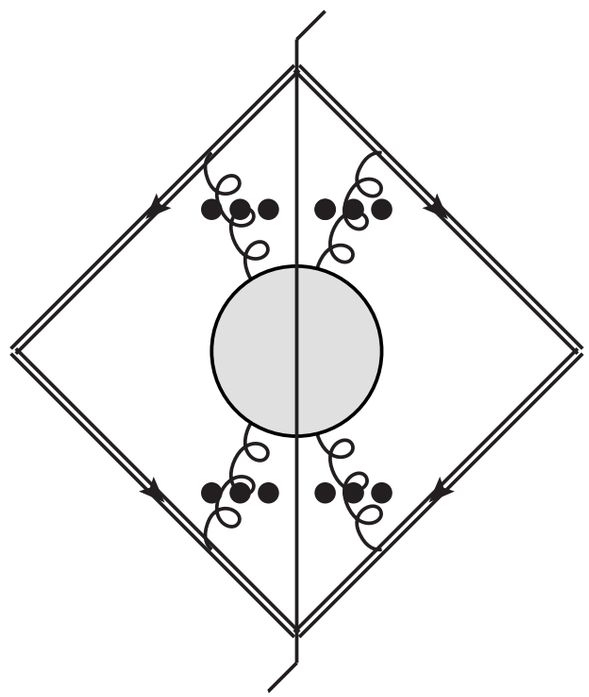
$$f^{(0)}(x, \mathbf{k}_T) \propto \int dk^- \quad \text{Tr}\Big[$$



$$\Big]$$

  - and **soft function:**

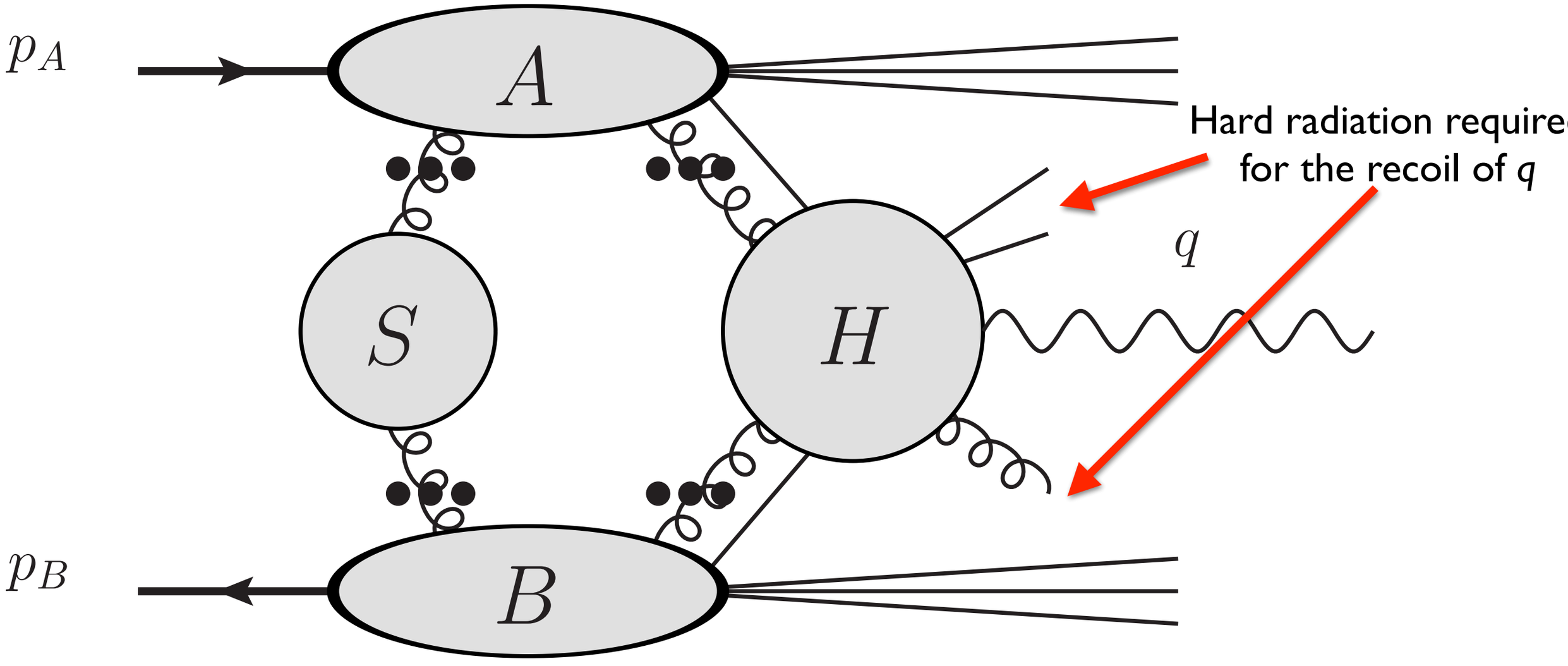
$$S^{(0)}(\mathbf{k}_T) \propto \int dk^+ dk^- \quad \text{Tr}\Big[\xrightarrow{k}$$



$$\Big]$$

Collinear factorisation

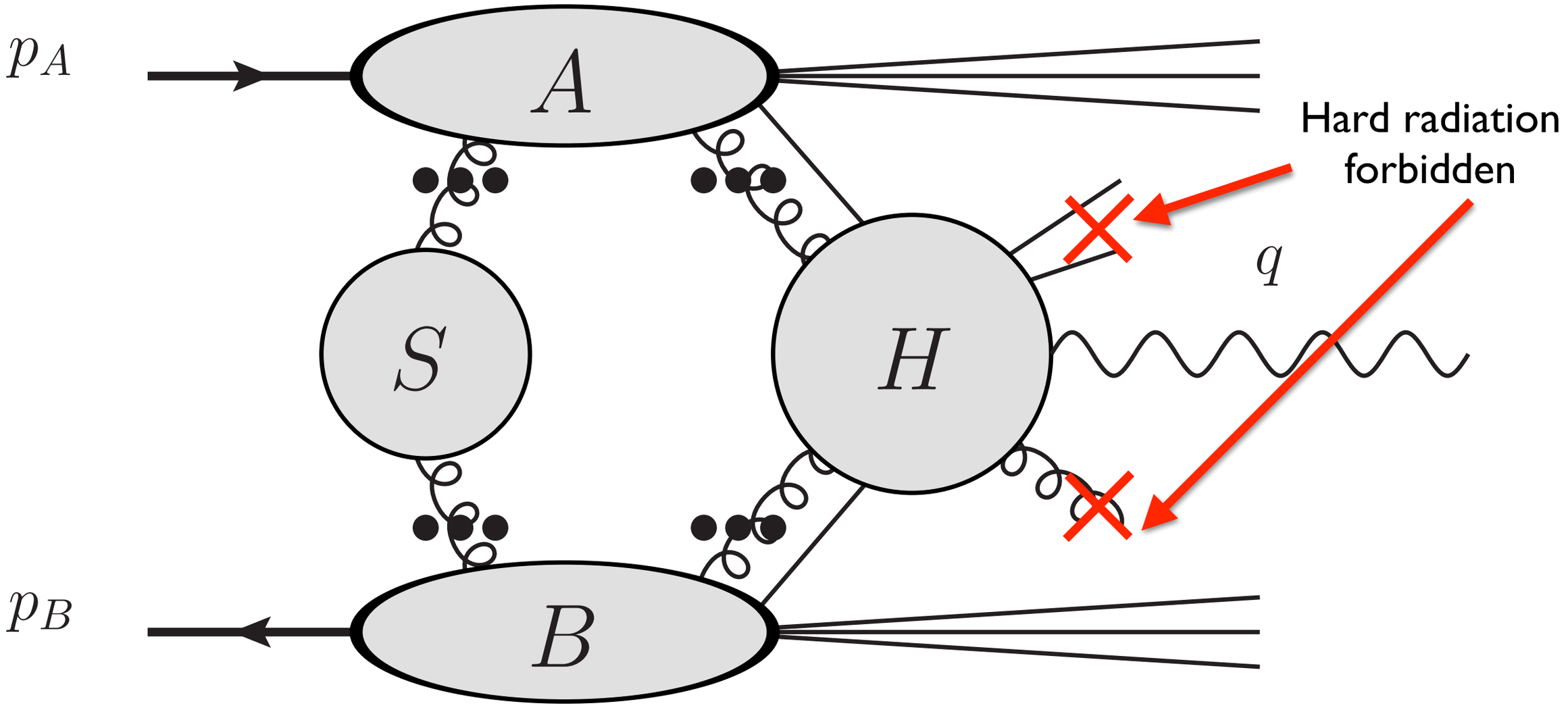
- Regime 1: $Q \simeq q_T$



- There must be real **hard radiation** to generate large q_T.
- ***H*** is a "complicated" function accounting for both **real** and **virtual** corrections.
- **Insensitivity** to intrinsic transverse momentum of partons ($k_T \simeq \Lambda_{\mathrm{QCD}}$).
- Soft radiation under control (*S* unitarises), collinear radiation resumed via DGLAP.

TMD factorisation

- Regime 2: $Q \gg q_T$



- Hard radiation inhibited, only **soft and/or collinear radiation** allowed.
- ***H*** is a "simple" scalar function that only accounts for **virtual** corrections.
- Sensitivity to parton **NP intrinsic transverse momentum** (when $q_T \lesssim \Lambda_{\rm QCD}$).
- Soft and collinear radiation resummed via TMD evolution and matching.

Factorisations in formulas

- The two regimes obey two different factorisation theorems:
  - for $q_\mathrm{T} \simeq \mathrm{Q}$ **collinear factorisation** at *fixed perturbative order* is appropriate:

$$\left(\frac{d\sigma}{dq_T}\right)_{\text{coll.}} = \int_{x_1}^{1} \frac{dy_1}{y_1} \int_{x_2}^{1} \frac{dy_2}{y_2} f_1(y_1,\mu) f_2(y_2,\mu) \frac{d\hat{\sigma}}{dq_T}\left(\frac{x_1}{y_1}, \frac{x_2}{y_2}, \mu, Q\right)$$

  - for $q_\mathrm{T} \ll \mathrm{Q}$ **transverse-momentum-dependent (TMD) factorisation** at *fixed logarithmic order* is appropriate:

$$\begin{aligned}\left(\frac{d\sigma}{dq_T}\right)_{\text{TMD}} &= \sigma_0 H(Q,\mu) \int d^2k_{T1} d^2k_{T2} F_1(x_1, k_{T1}, \mu, \zeta_1) F_2(x_2, k_{T2}, \mu, \zeta_2) \delta^{(2)}(q_T - k_{T1} - k_{T2}) \\ &= \sigma_0 H(Q,\mu) \int_0^\infty db_T\, b_T J_0(b_T q_T) F_1(x_1, b_T, \mu, \zeta_1) F_2(x_2, b_T, \mu, \zeta_2)\end{aligned}$$

  - Collinear and TMD factorisations can be **matched** to produce accurate results over the the full q_T spectrum.

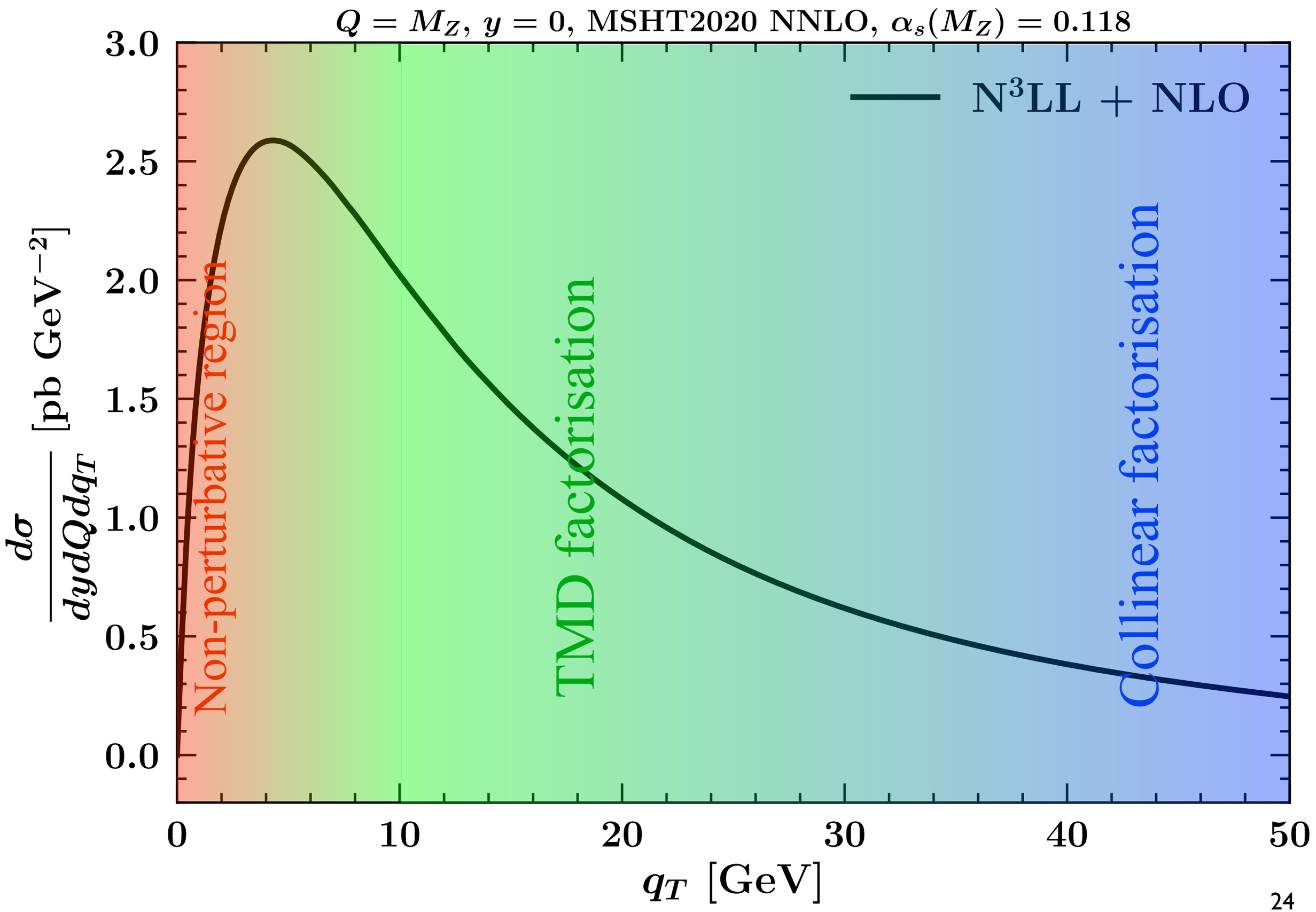
Combining factorisations

- Accurate predictions for all $q_T \gg \Lambda_{\rm QCD}$ can be obtained by **matching**:
  - different **recipes** for the matching exist.

$$\left(\frac{d\sigma}{dq_T}\right)_{\textbf{match.}} = \left(\frac{d\sigma}{dq_T}\right)_{\rm TMD} + \left(\frac{d\sigma}{dq_T}\right)_{\textbf{coll.}} - \left(\frac{d\sigma}{dq_T}\right)_{\textbf{d.c.}}$$

- In order for the matching to actually take place one needs:

$$\left(\frac{d\sigma}{dq_T}\right)_{\rm TMD} \xrightarrow[\textbf{f.o.}]{} \left(\frac{d\sigma}{dq_T}\right)_{\textbf{d.c.}} \xleftarrow[q_T \ll Q]{} \left(\frac{d\sigma}{dq_T}\right)_{\textbf{coll.}}$$


$Q = M_Z$, $y = 0$, MSHT2020 NNLO, $\alpha_s(M_Z) = 0.118$
N³LL + NLO
$\frac{d\sigma}{dydQdq_T}$ [pb GeV^{-2}]
3.0
2.5
2.0
1.5
1.0
0.5
0.0
Non-perturbative region
TMD factorisation
Collinear factorisation
0
10
20
30
40
50
q_T [GeV]

TMD evolution and resummation

- Removal of UV and rapidity divergences allows one to derive **two** evolution equations for the TMD, along with the cross derivative:

$$\begin{cases} \dfrac{d\ln F}{d\ln\mu} = \gamma(\mu,\zeta) \\ \\ \dfrac{d\ln F}{d\ln\sqrt{\zeta}} = K(\mu) \end{cases}, \quad \frac{d^2\ln F}{d\ln\mu\, d\ln\sqrt{\zeta}} = \begin{cases} \dfrac{d\gamma}{d\ln\sqrt{\zeta}} \\ \\ \dfrac{dK}{d\ln\mu} \end{cases} = \gamma_K(\alpha_s(\mu))$$

- Solving these equations requires fixing **two pairs of scales**:
  - **initial** scales: (μ_0,ζ_0)
  - **final** scales: (μ,ζ)
- The solution **resums** $\ln(\mu/\mu_0)$ and $\ln(\zeta/\zeta_0)$ to all orders reads in α_s:

$$F(\mu,\zeta) = R\left[(\mu,\zeta)\leftarrow(\mu_0,\zeta_0)\right] F(\mu_0,\zeta_0)$$

$$R\left[(\mu,\zeta)\leftarrow(\mu_0,\zeta_0)\right] = \exp\left\{K(\mu_0)\ln\frac{\sqrt{\zeta}}{\sqrt{\zeta_0}} + \int_{\mu_0}^{\mu}\frac{d\mu'}{\mu'}\left[\gamma_F(\alpha_s(\mu')) - \gamma_K(\alpha_s(\mu'))\ln\frac{\sqrt{\zeta}}{\mu'}\right]\right\}$$

- The scales (μ,ζ) have to be such that $\mu \sim Q$ and $\zeta_1\zeta_2 = Q^4$.
- How do we *sensibly* choose (μ_0,ζ_0)?

TMD evolution and resummation

[Collins, Soper, Sterman, *Nucl.Phys.B* 250 (1985) 199-224]

- A sensible choice of the scales is important to **allow truncated perturbation theory to be reliable**:
  - **no large unresummed logarithms** should be introduced,
  - each scale has to be set in the **vicinity of its "natural" value**.
- In TMD factorisation the relevant scales are $\boldsymbol{q_T}$ and $\boldsymbol{Q}$:
  - natural to expect $\mu_0 \sim \sqrt{\zeta_0} \sim q_T \sim b_T^{-1}$ and $\mu \sim \sqrt{\zeta} \sim Q$
- In the $\overline{\text{MS}}$ scheme, **central scales** are usually chosen to be:

$$\mu_0 = \sqrt{\zeta_0} = \frac{2e^{-\gamma_E}}{b_T} \equiv \mu_b \quad \text{and} \quad \mu = \sqrt{\zeta} = Q$$

- This particular choice **nullifies** all unresummed (fixed-order) logs.
- Modest variations around these values give an estimate of higher-order corrections.

TMD factorisation

- **At small values of b_T** TMDs can be related to *collinear* PDFs.

- The main observation is that TMDs obey a (non-local) **OPE**:

$$F(x, b_T, \mu_0, \zeta_0) = C(x, b_T, \mu_0, \zeta_0) \otimes f(x, \mu_0) + \mathcal{O}(b_T^2)$$

- The coefficient of leading-power matching are **perturbative**:

$$C(x, b_T, \mu_0, \zeta_0) = \sum_{n=0}^{\infty} \alpha_s^n C^{(n)}(x, b_T, \mu_0, \zeta_0)$$

- In conclusion, for b_T is small enough, one obtains the formula:

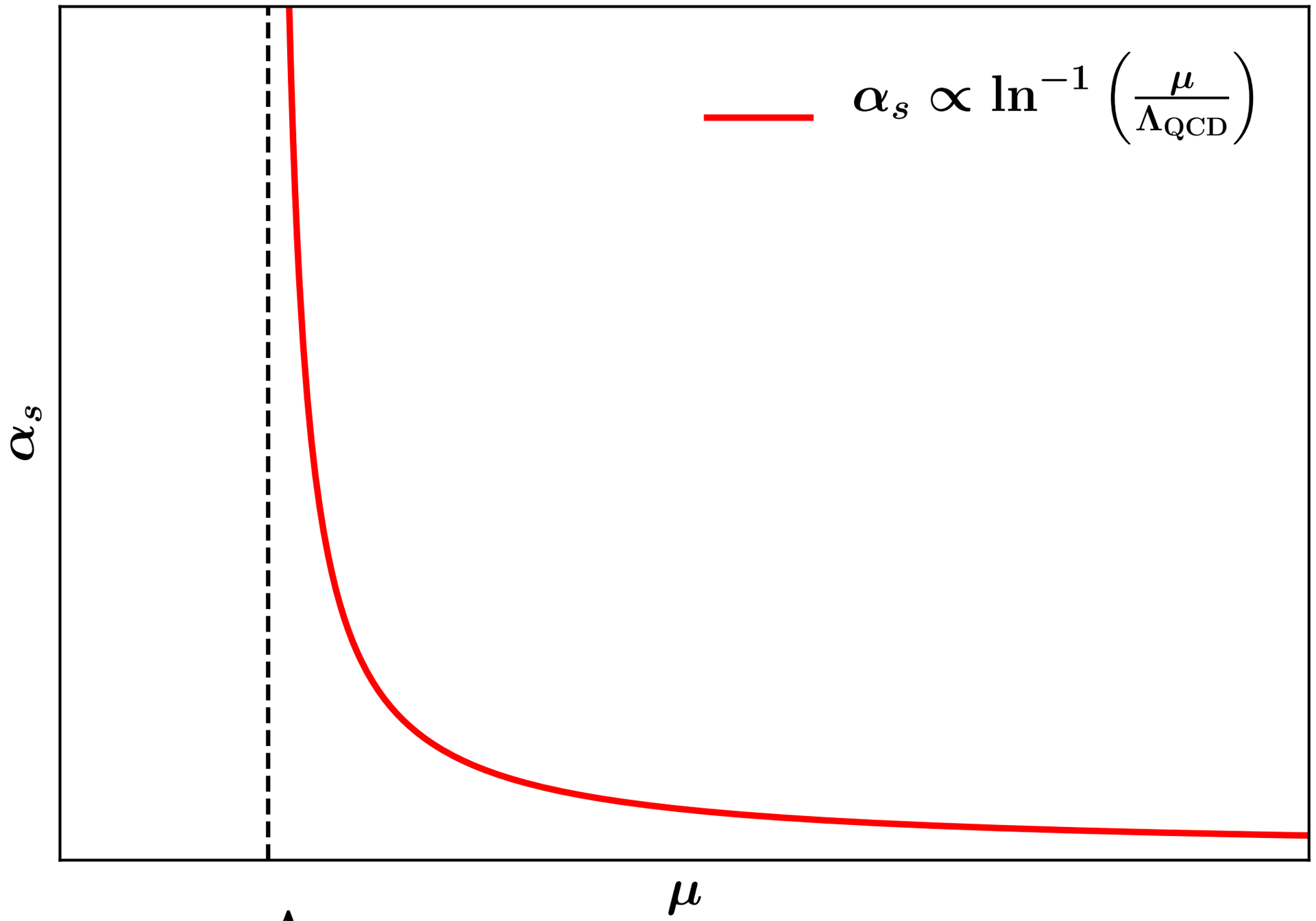
$$F(x, b_T, \mu, \zeta) = R\left[(\mu, \zeta) \leftarrow (\mu_0, \zeta_0); b_T\right] C(x, b_T, \mu_0, \zeta_0) \otimes f(x, \mu_0)$$

- This allows one to relate perturbative ingredients and the **collinear PDF** to obtain the TMD.

- Ready for phenomenology? Not quite. How about large b_T, *i.e* small k_T?

Non-perturbative effects

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



Non-perturbative effects

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

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

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

Non-pert.

Purely perturbative

- **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.

- Extracting TMDs boils down to determining

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

- Usually done through **fits to experimental data**.

Non-perturbative effects

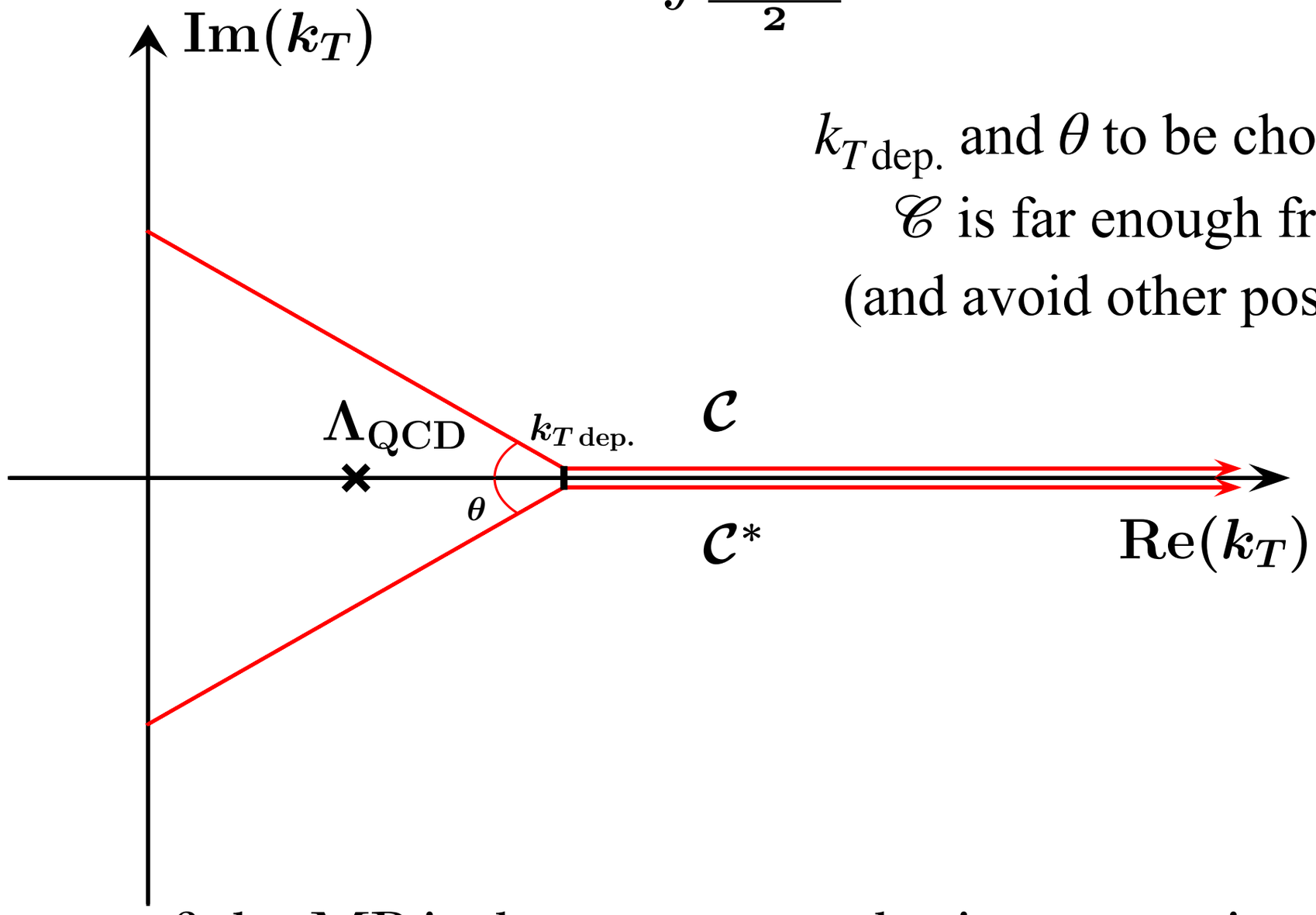
$$f_{\rm NP}^{(i)}(x, b_T; \zeta) = \exp\left[g_i(x, b_T) + g_K(b_T)\ln\left(\frac{\zeta}{Q_0^2}\right)\right]$$

- Given the properties above, there is not a huge latitude in defining $f_{\rm NP}$:
  - both 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:
    - $f_{\rm NP}$ can be sensibly different from one only for $b_T \gtrsim \Lambda_{\rm QCD}^{-1}$ ($k_T \lesssim \Lambda_{\rm QCD}$).
  - 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).
    - **Lattice** simulations can help.

Non-perturbative effects

2) The **Minimal Prescription**:

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



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

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

TMD factorisation

The operational version with the b_*-prescription

- At $q_T \ll Q$ (including $q_T \lesssim \Lambda_{\rm QCD}$), TMD factorisation reads:

$$\left(\frac{d\sigma}{dq_T}\right)_{\rm res.} = \sigma_0 H(Q,Q) \int_0^\infty db_T\, b_T J_0(q_T b_T) F_1(x_1, b_T, Q, Q^2) F_2(x_2, b_T, Q, Q^2)$$

- Defining $\mu_{b_*} = \dfrac{2e^{-\gamma_E}}{b_*(b_T)}$, the single TMD distributions are given by:

$$\begin{aligned} F(x, b_T, Q, Q^2) &= C(x, \mu_{b_*}, \mu_{b_*}^2) \otimes f(x, \mu_{b_*}) \\ &\quad \text{Matching onto collinear distributions} \\ &\times \exp\left\{ K(\mu_{b_*}) \ln \frac{Q}{\mu_{b_*}} + \int_{\mu_{b_*}}^{Q} \frac{d\mu'}{\mu'} \left[\gamma_F(\alpha_s(\mu')) - \gamma_K(\alpha_s(\mu')) \ln \frac{Q}{\mu'} \right] \right\} \\ &\quad \text{Perturbative evolution} \\ &\times \exp\left[g(x, b_T) + g_K(b_T) \ln\left(\frac{Q}{Q_0}\right) \right] \\ &\quad \text{TMD non-perturbative contribution} \end{aligned}$$

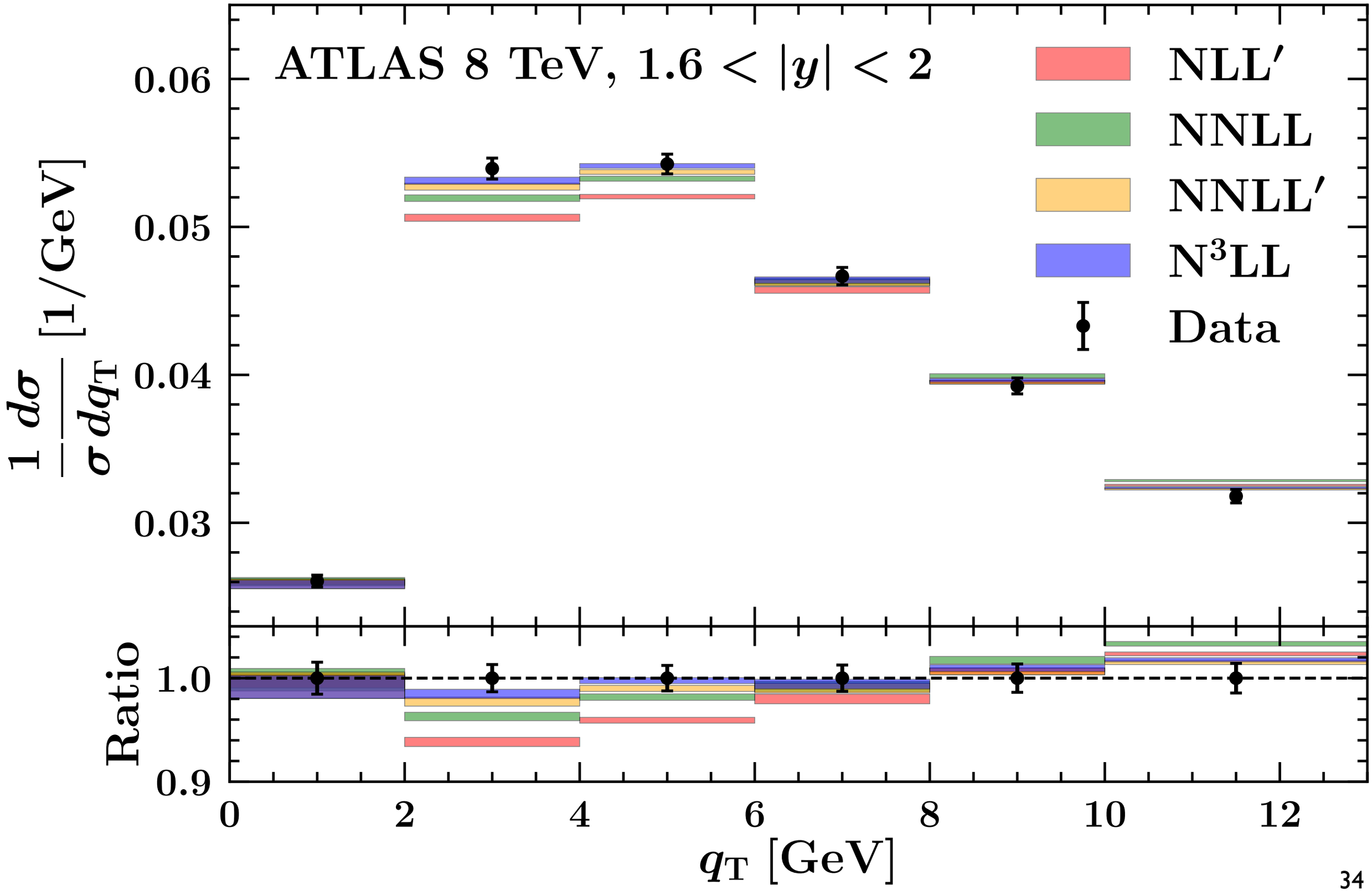
Logarithmic counting

$$\left(\frac{d\sigma}{dq_T}\right)_{\text{TMD}} = \sigma_0 H(Q) \int d^2\mathbf{b}_T e^{i\mathbf{b}_T\cdot\mathbf{q}_T} F_1(x_1, \mathbf{b}_T, Q, Q^2) F_2(x_2, \mathbf{b}_T, Q, Q^2)$$

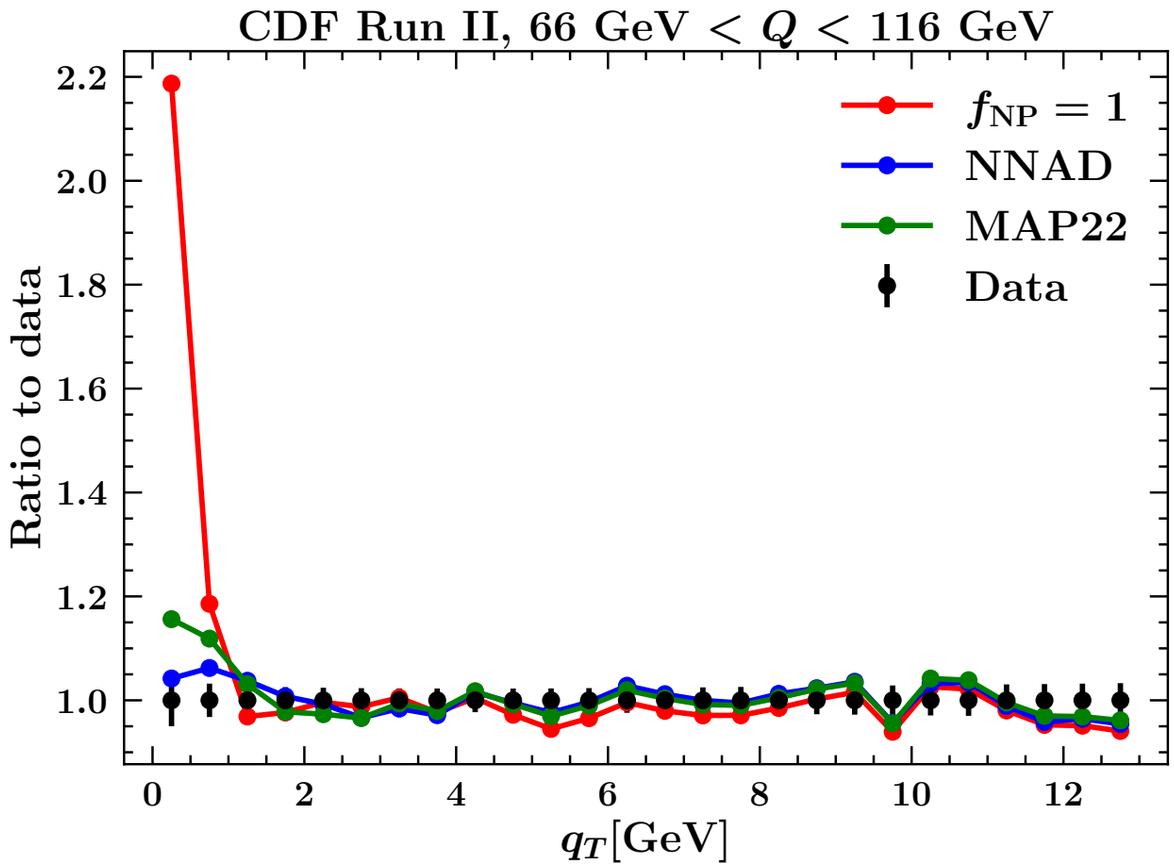
$$F = \exp\left\{ K \ln\frac{\sqrt{\zeta}}{\mu_b} + \int_{\mu_b}^{\mu} \frac{d\mu'}{\mu'}\left[\gamma_F - \gamma_K \ln\frac{\sqrt{\zeta}}{\mu'}\right]\right\} C \otimes f$$

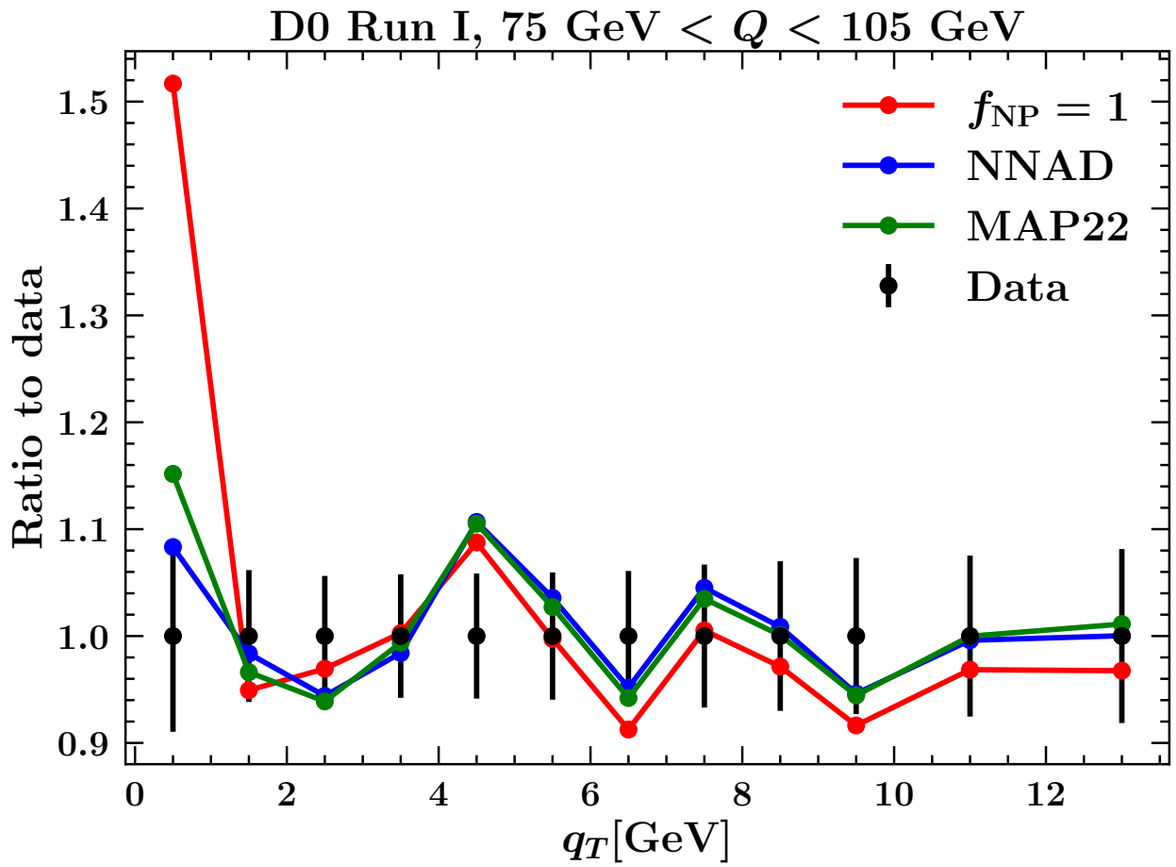
Accuracy	γ_K	γ_F	K	C	H	FFs/PDFs/α_s
LL	α_s	-	-	1	1	-
NLL	α_s^2	α_s	α_s	1	1	LO
NLL’	α_s^2	α_s	α_s	α_s	α_s	LO
N^{2}LL	α_s^3	α_s^2	α_s^2	α_s	α_s	NLO
N^{2}LL’	α_s^3	α_s^2	α_s^2	α_s^2	α_s^2	NLO
N^{3}LL	α_s^4	α_s^3	α_s^3	α_s^2	α_s^2	NNLO
N^{3}LL’	α_s^4	α_s^3	α_s^3	α_s^3	α_s^3	NNLO
N^{4}LL	α_s^5	α_s^4	α_s^3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

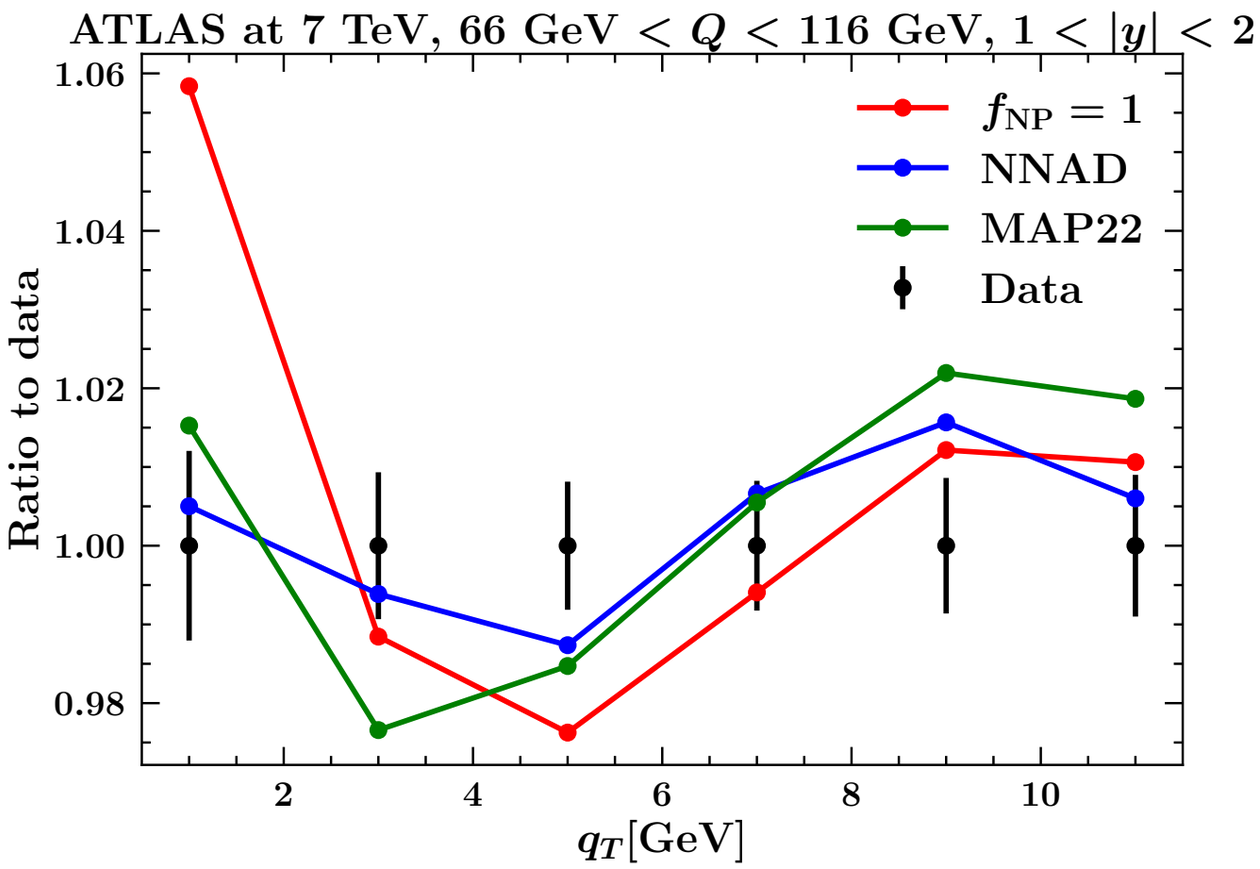
Perturbative convergence

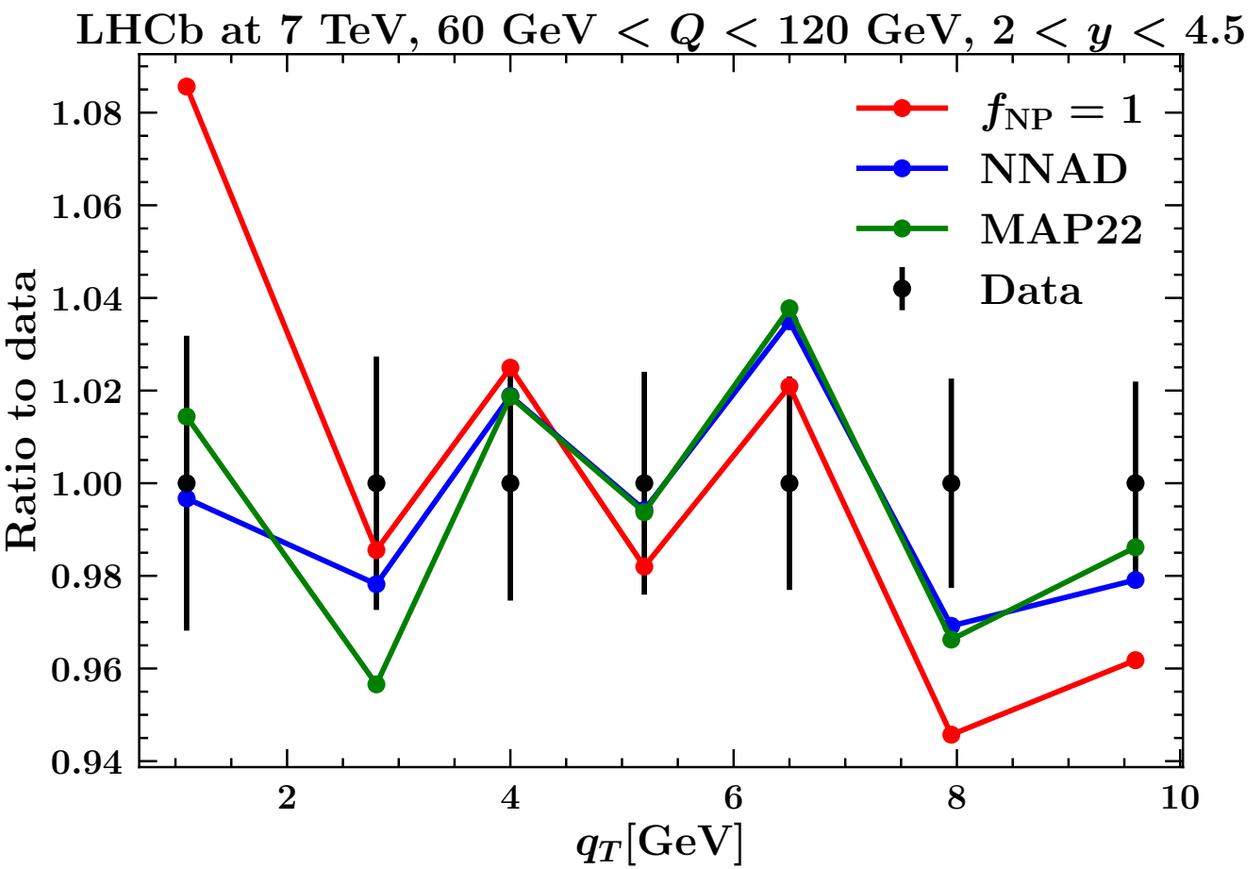


On the importance of $f_{\rm NP}$





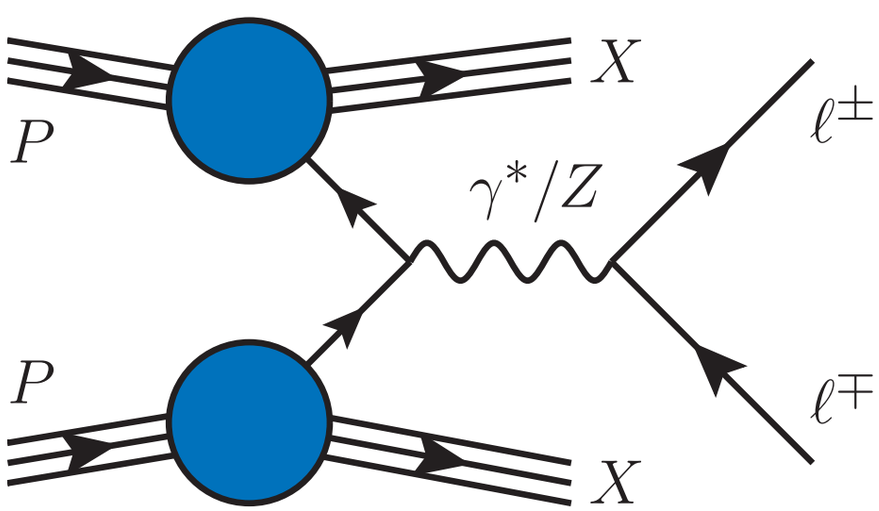




Factorising processes (quarks)

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

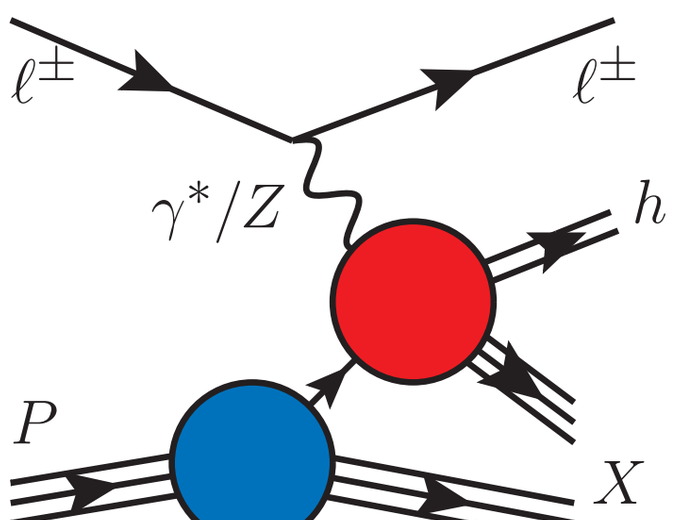
Drell-Yan



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

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

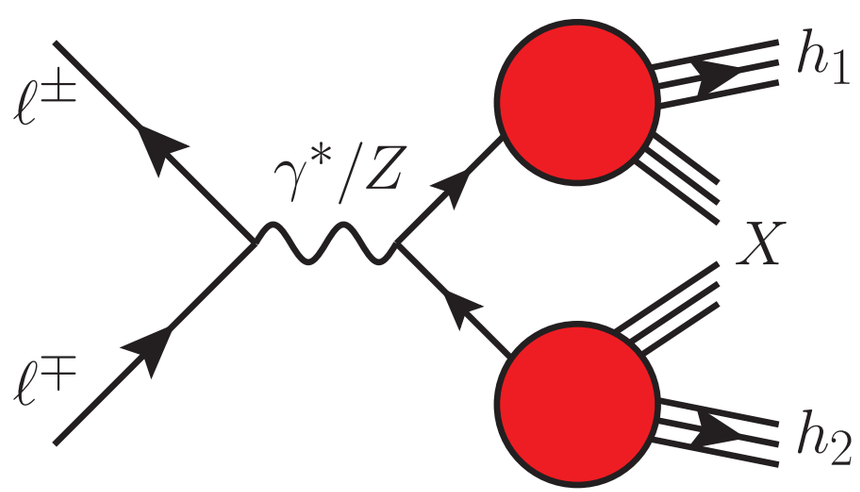
Semi-inclusive DIS



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

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

e^+e^- annihilation



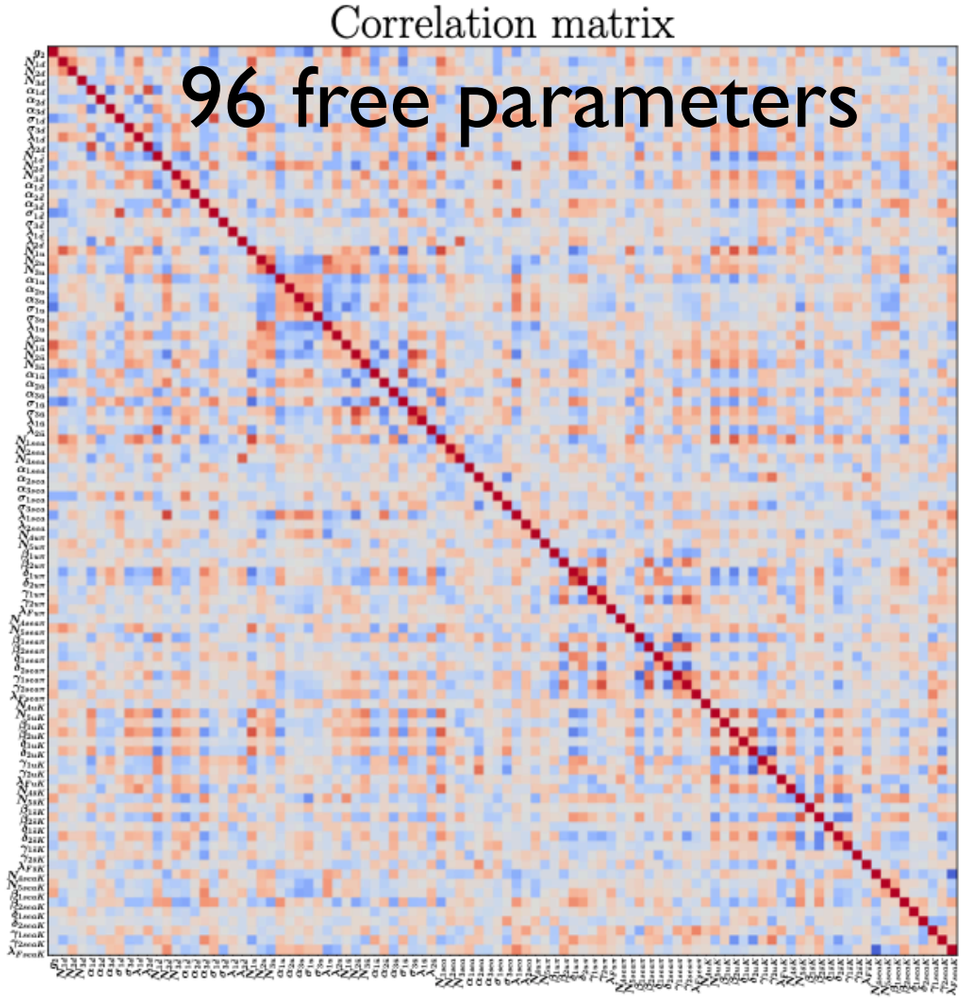
$$\ell^\pm \ell^\mp \to h_1 h_2 X$$

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

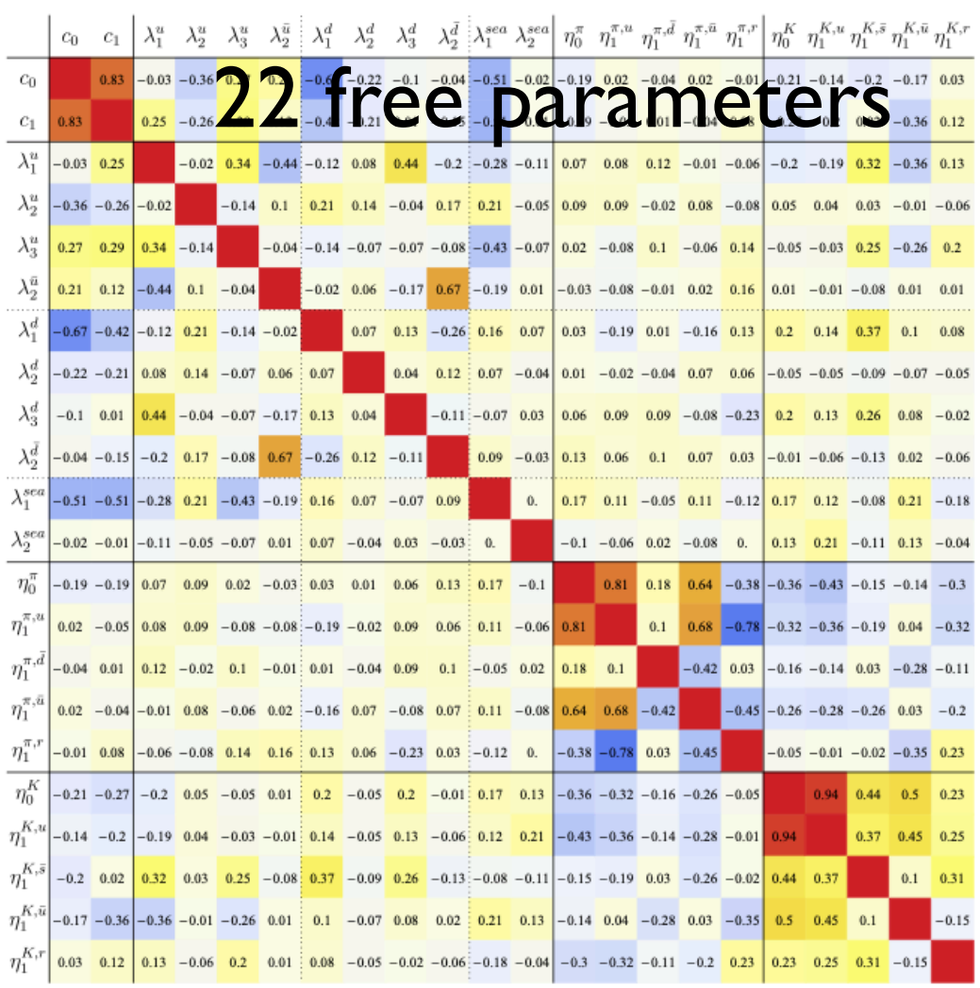
- Examples of other precesses:
  - thrust and p_{hT} distributions in single-hadron production in e^+e^-,
  - hadron-in-jet production,
  - …

Parametrising TMDs

- TMD parameterisations became involved as more data became available.



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



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

- We are in a stage in which **TMD parameterisations** may significantly **bias** the extraction.

Neural-Network TMDs

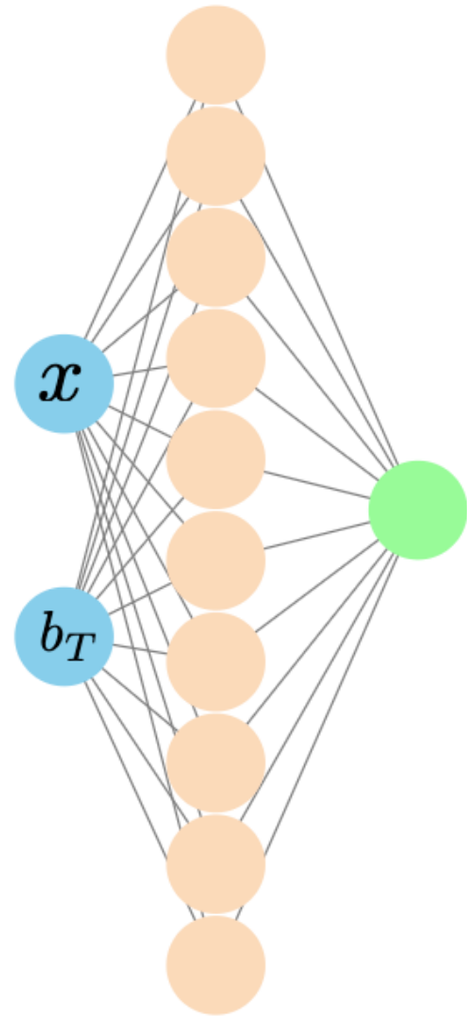
- Minimising the parametrisation bias calls for **neural networks**.
- In [*Phys. Rev. Lett.* 135 (2025) 2, 021904], we used:

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

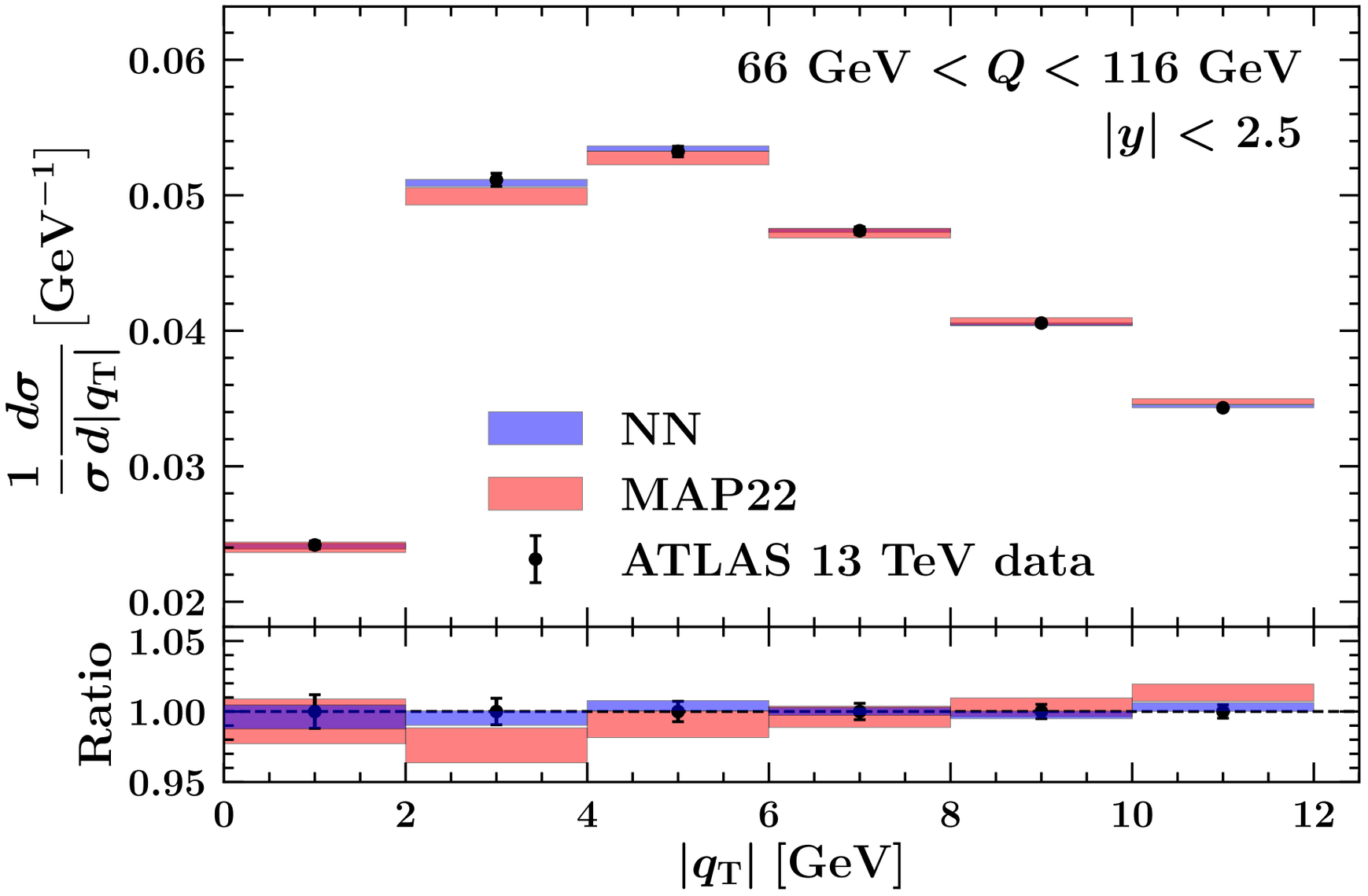
- NN architecture [2, 10, 1]: inputs x and b_T, 10 hidden nodes, one output.
- Sigmoid-like activation function:

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

- $f_{\rm NP} \to 1$ 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

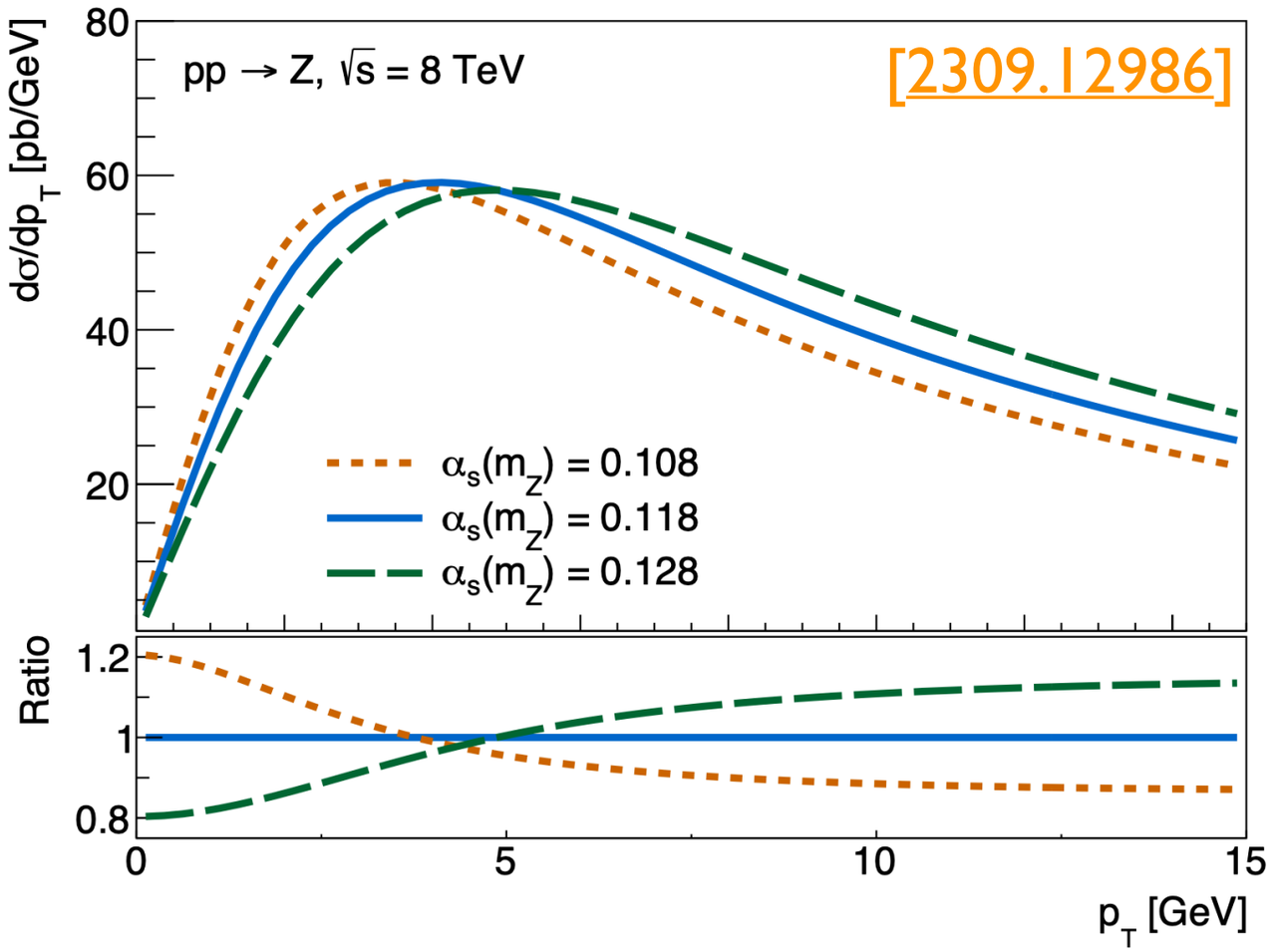


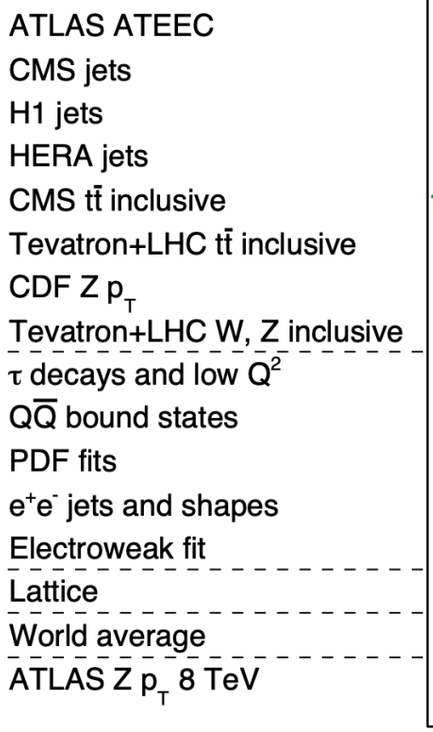
- NN performs significantly better on ATLAS data than MAP22:
  - **shape** is better described by NN,
  - Counterintuitively, NN uncertainties on predictions are smaller than MAP22 ones because the latter are driven by **systematic shifts**.

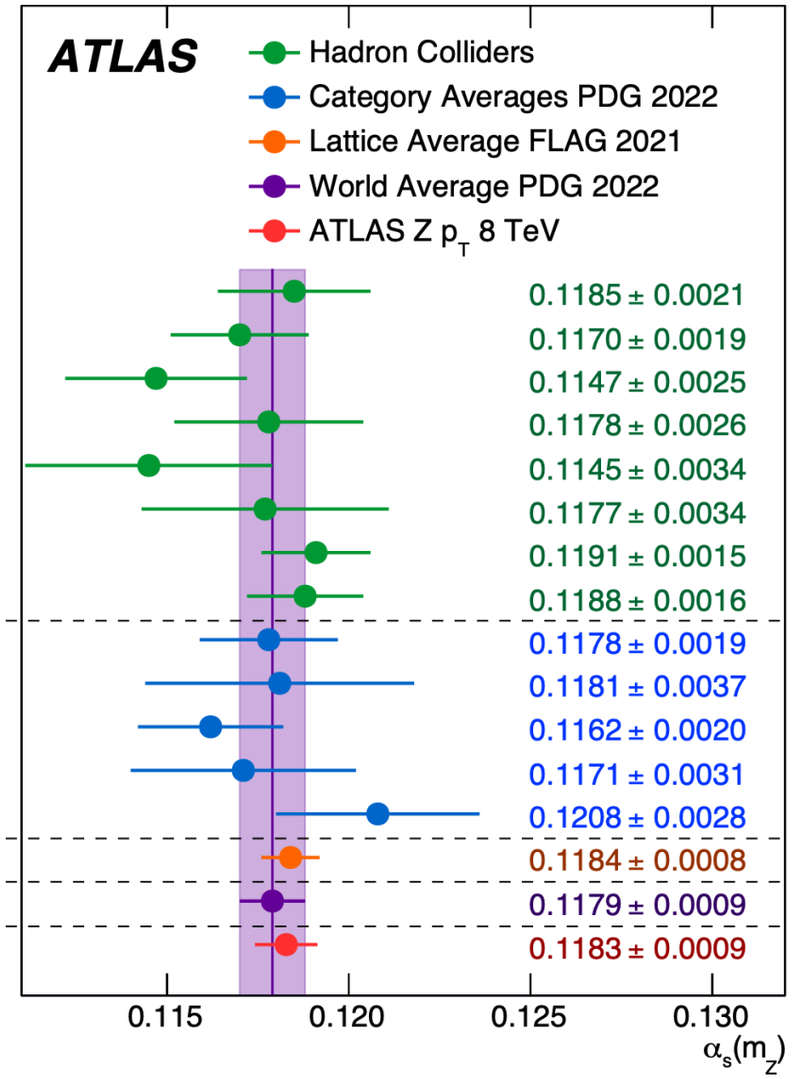
TMDs at the LHC

The strong coupling

- **Strong sensitivity** of the DY spectrum to $(\alpha_s(M_Z) = 0.1183 \pm 0.0009)$.







- Importantly from TMD perspective, this work required a determination of the non-perturbative component of TMDs.

TMDs at the LHC

The W mass

- A precise determination of the ***W* mass** plays an important role in testing the Standard Model and thus for **BSM** physics.

- This is a central task of the **LHC physics programme**.

- In order to minimise experimental systematic effects, the most promising procedure relies on the measurement of the W/Z ratio cross section:

  - the W mass is basically determined through template fits of:

$$\frac{d\sigma^W}{dq_T} = \left(\frac{d\sigma^W/dq_T}{d\sigma^Z/dq_T}\right)_{\text{exp.}} \left(\frac{d\sigma^Z}{dq_T}\right)_{\text{th.}}$$

- Therefore, an accurate and reliable prediction of the ***Z* spectrum** is essential.

- **TMD-based predictions** are currently playing an important role within the LHC electroweak working group along **other formalisms**.

Summary

- **GTMDs** are the "mother distributions": **GPDs** and **TMDs** descend from them and give complementary **3D images** of hadrons, in position and momentum space.

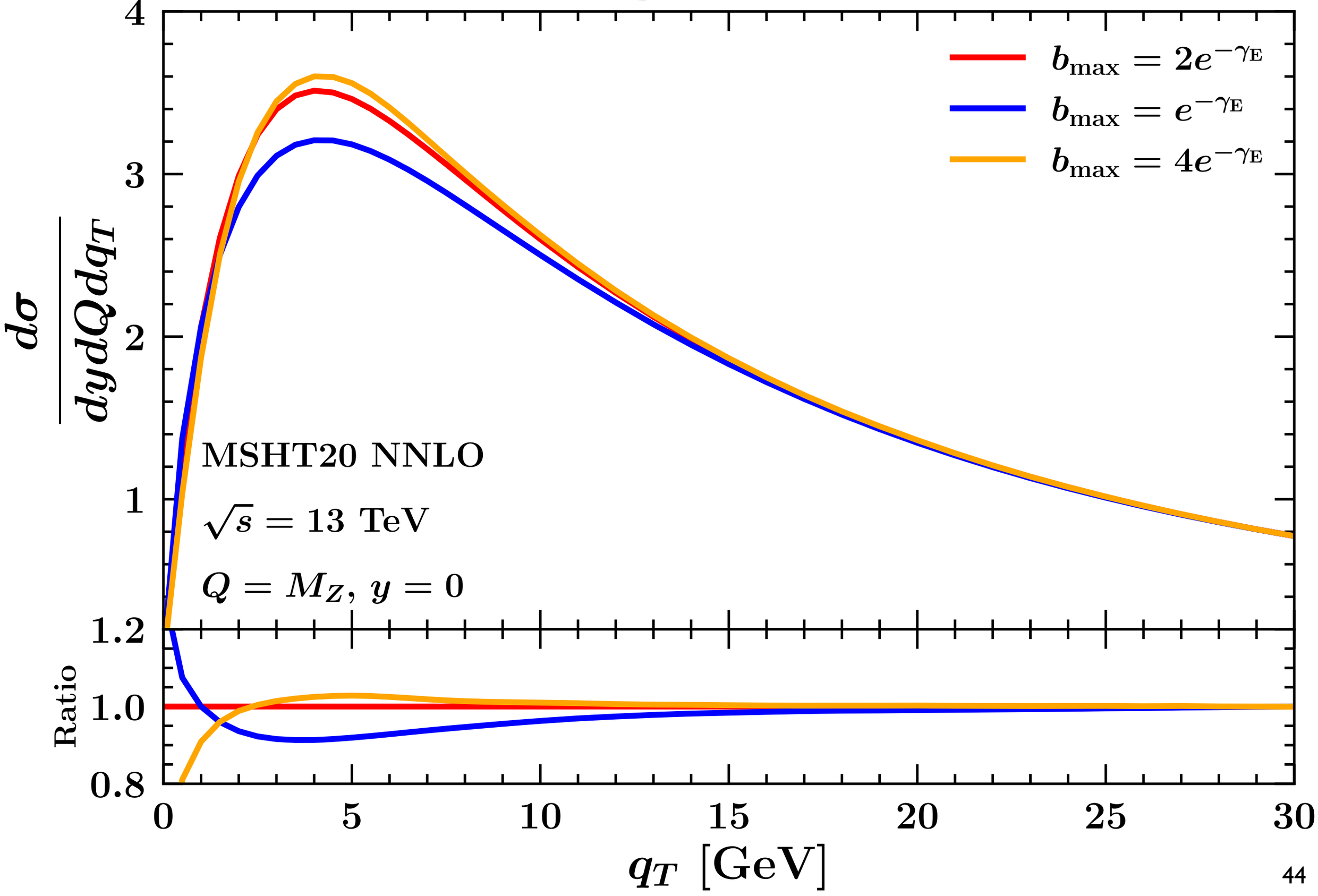
- **TMD factorisation** is the right framework for $q_T \ll Q$, where collinear factorisation breaks down. Evolution and matching **resum large logarithms** but **non-perturbative** corrections must be taken into account.

- As more data became available, parametrisations became more complex and may **bias** the extraction. **Neural networks** reduce this bias and describe LHC data better.

- **TMD physics** has important consequences **at the LHC**: for example α_s and W mass.

Backup

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



Importance of the x-dependence

- However $f_{\rm NP}$ cannot be too simple either!

- In [*JHEP* 07 (2020) 117] a fit of an ***x*-dependent** $f_{\rm NP}$ to a set 353 DY data points was done, obtaining a χ^2 per data point ($\chi^2/N_{\rm dat}$) equal to **1.02**:

- A **gaussian**, ***x*-independent** $f_{\rm NP}$ was also tested:

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

- This $f_{\rm NP}$ was used to fit the full data set as well as a subset in which rapidity-dependent data (from ATLAS) was excluded:

	Full dataset	No y-differential data
Global $\chi^2/N_{\rm dat}$	1.339	0.895

- **Gaussian ansatz is insufficient** to describe data accurately. Deterioration largely due to the rapidity-dependent data.

Resummation formalisms

- Different formulations of the q_T spectrum valid for $q_T \ll Q$ and $q_T \gg \Lambda_{\rm QCD}$:

$$\left(\frac{d\sigma}{dq_T}\right)_{\rm res.} \propto \begin{cases} e^{2S}\left[f_1 \otimes \mathcal{H} \otimes f_2\right] & : q_T \text{ resum.} \\ H \times B_1 \times B_2 \times S & : \text{SCET} \\ H \times F_1 \times F_2 & : \text{TMD} \end{cases} + \mathcal{O}\left[\left(\frac{\Lambda_{\rm QCD}}{q_T}\right)^n, \left(\frac{q_T}{Q}\right)^m\right]$$

- All of them **resum** large $\ln(q_T/Q)$.

- All **equivalent** for *factorising* processes (such as inclusive Drell-Yan).

- Dictionary:

$$\mathcal{H} = HC_1C_2$$

f_i : Collinear PDFs

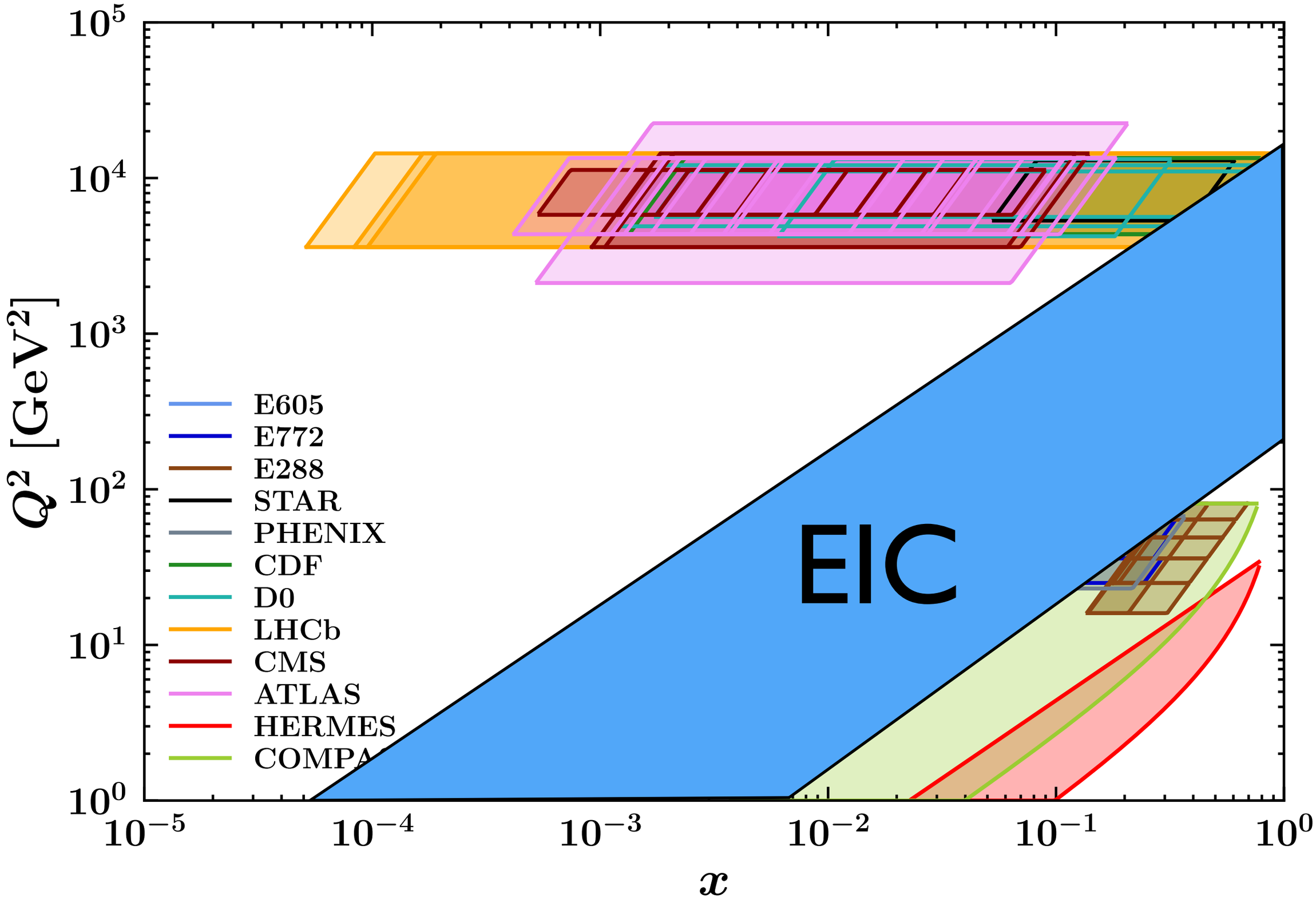
$$F_i = e^S C_i \otimes f_i$$

F_i : TMD PDFs

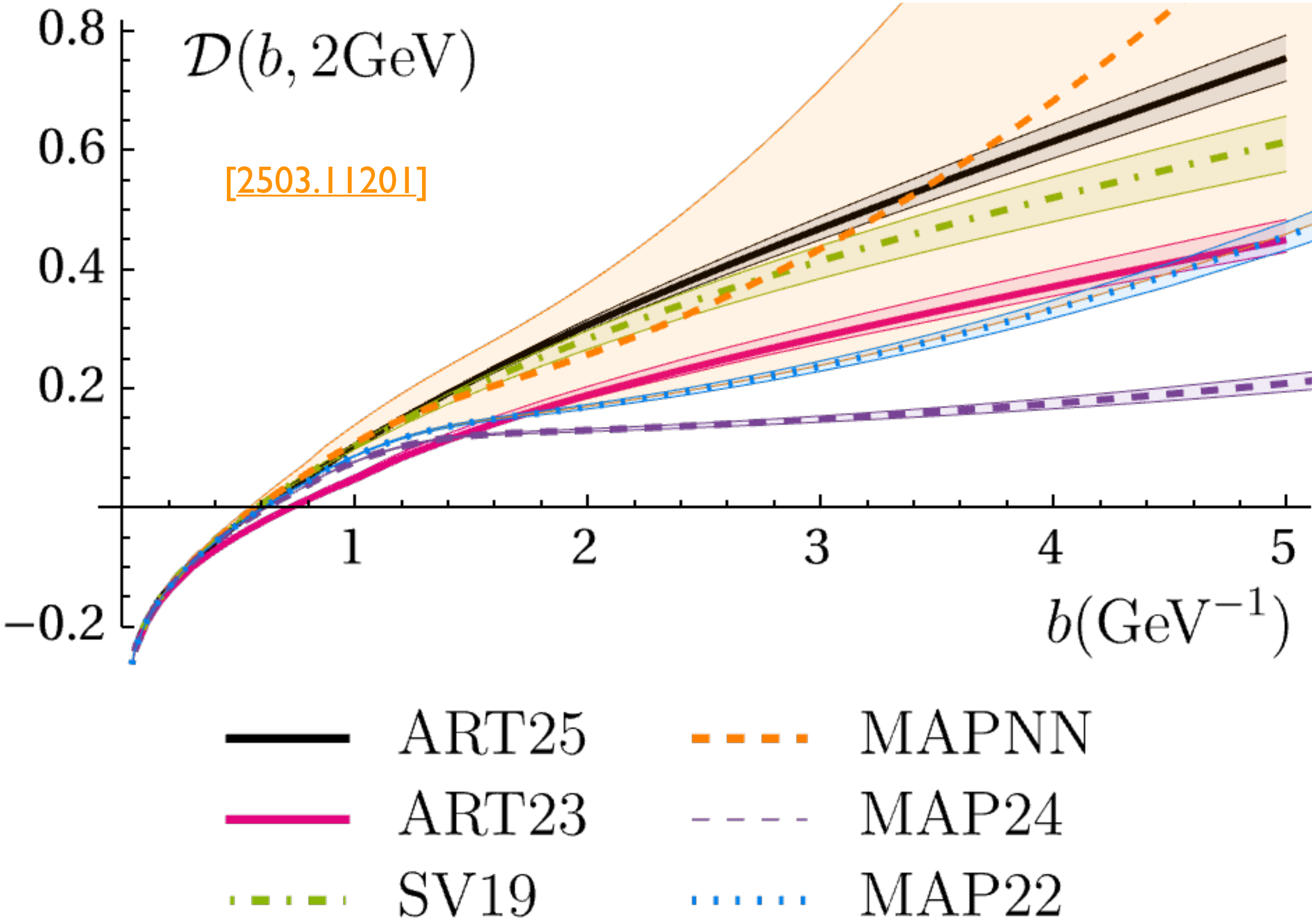
$$F_i = \sqrt{S} \times B_i$$

MAPTMD 2024

Kinematic coverage

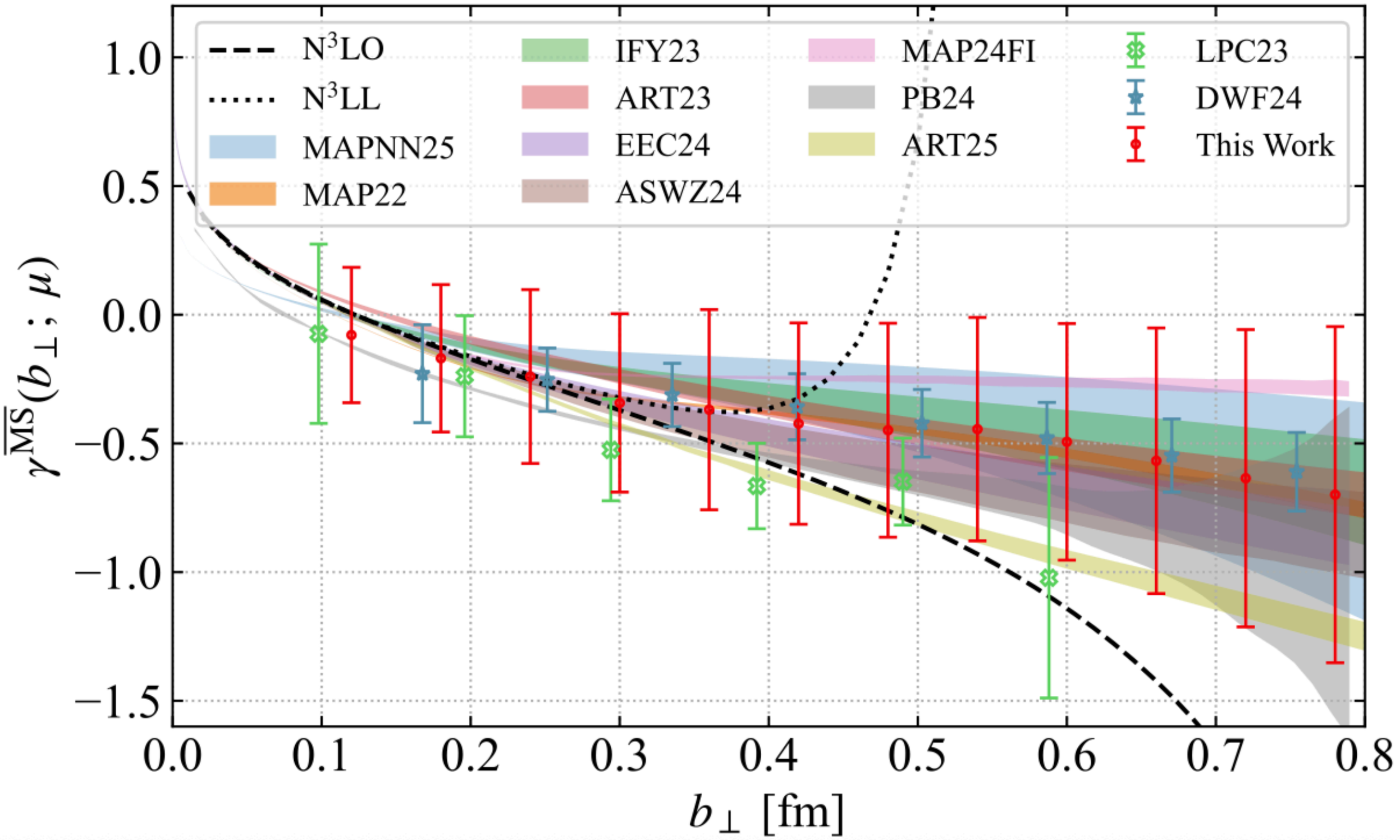


Collis-Soper kernel



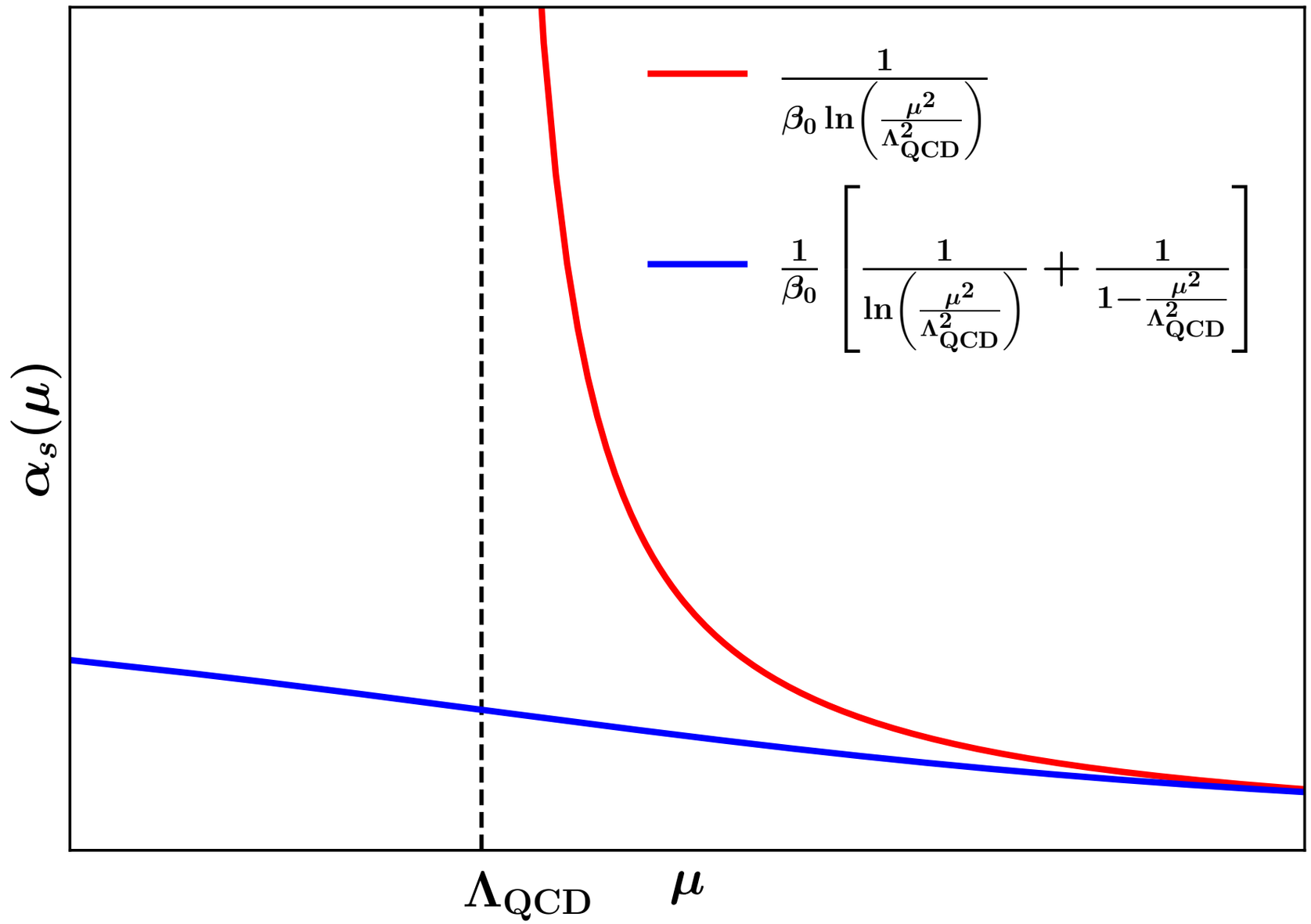
Collis-Soper kernel

[2504.04625]



Non-perturbative effects

- **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]:



Non-perturbative effects

- In the context of the resummation of logs of q_T, dispersive approaches have never been (seriously) explored.
- More popular approaches to the ***regularisation*** of the Landau pole are:

1) The **cutoff** method leads to corrections that scale as $(\Lambda_{\rm QCD}/q_T)^{\alpha}$ and in *direct space* (k_T) it amounts to:

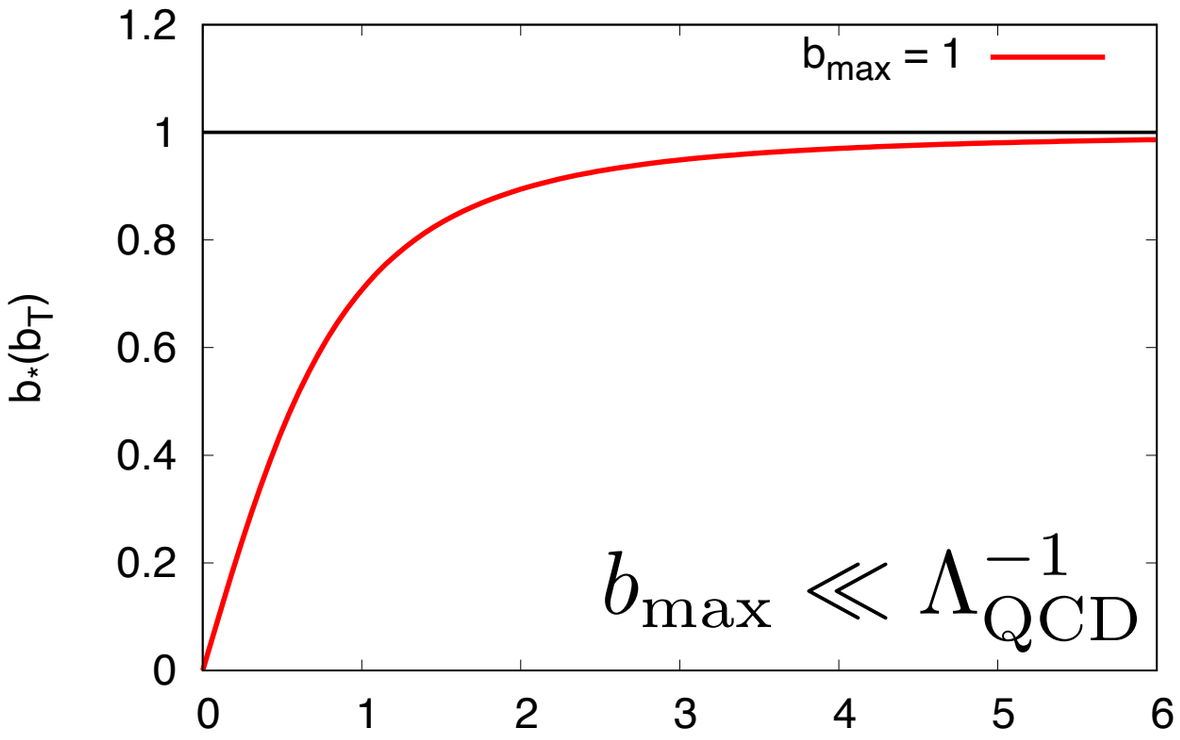
$$\alpha_s\left(k_{\rm T}\right) \to \alpha_s\left(\max[k_{\rm T,cutoff}, k_{\rm T}]\right) \,, \qquad k_{\rm T,cutoff} \gg \Lambda_{\rm QCD}$$

In *impact-parameter space* (b_T) it is typically implemented by means of the so-called **b∗-prescription**:

$$\alpha_s\left(\frac{2e^{-\gamma_{\rm E}}}{b_{\rm T}}\right) \to \alpha_s\left(\frac{2e^{-\gamma_{\rm E}}}{b_*(b_{\rm T})}\right)$$

A classical choice is:

$$b_*(b_{\rm T}) = \frac{b_{\rm T}}{\sqrt{1+b_{\rm T}^2/b_{\rm max}^2}}$$



Scales below $b_{\rm max}^{-1} \sim k_{T,\rm cutoff}$ are considered non-perturbative.

Non-perturbative effects

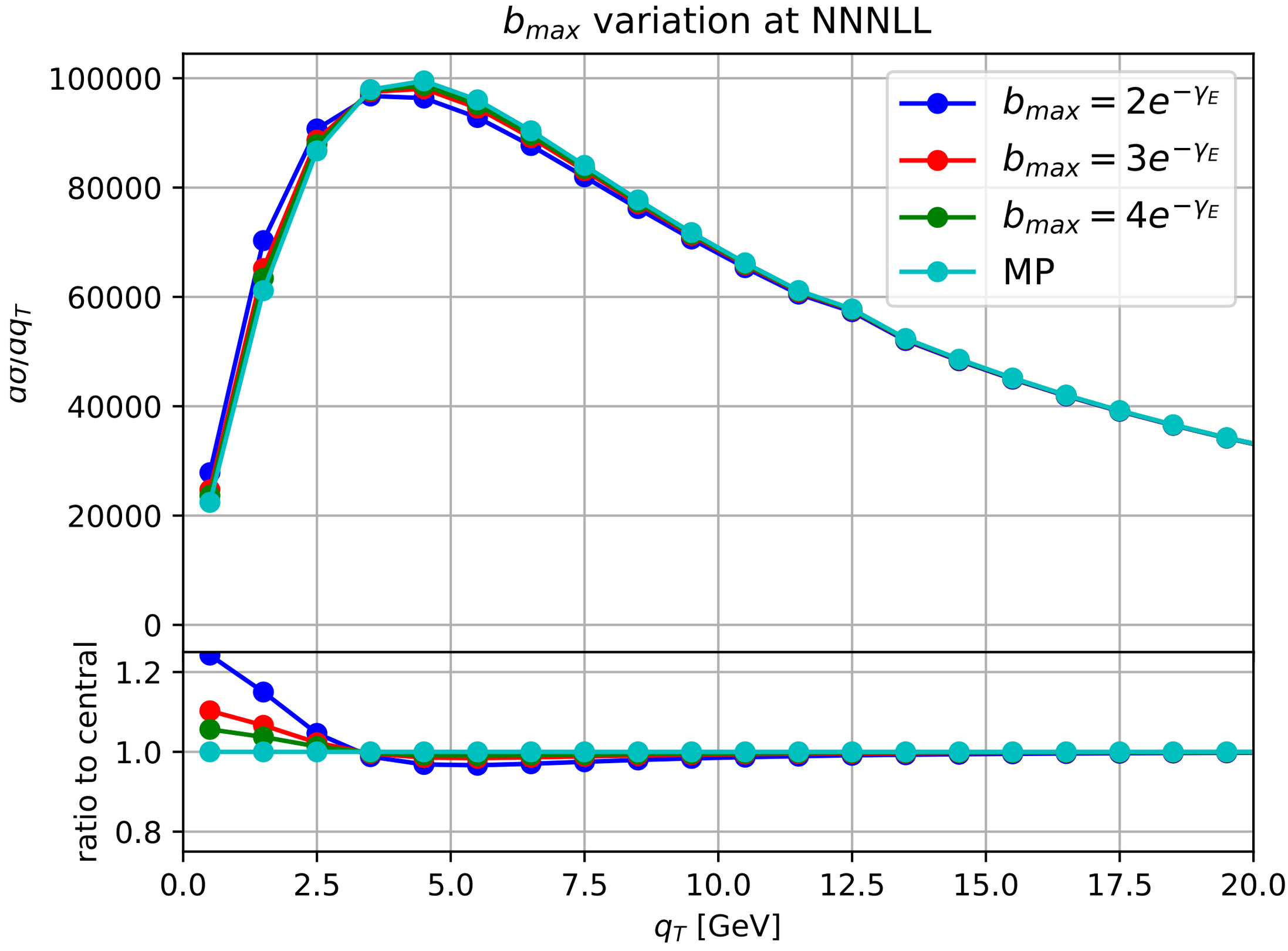
$$F_i(x, b_T; \mu, \zeta) = \left[\frac{F_i(x, b_T; \mu, \zeta)}{F_i(x, b_*(b_T); \mu, \zeta)}\right] F_i(x, b_*(b_T); \mu, \zeta) \equiv f_{\rm NP}^{(i)}(x, b_T; \zeta) F_i(x, b_*(b_T); \mu, \zeta)$$

Non-pert.

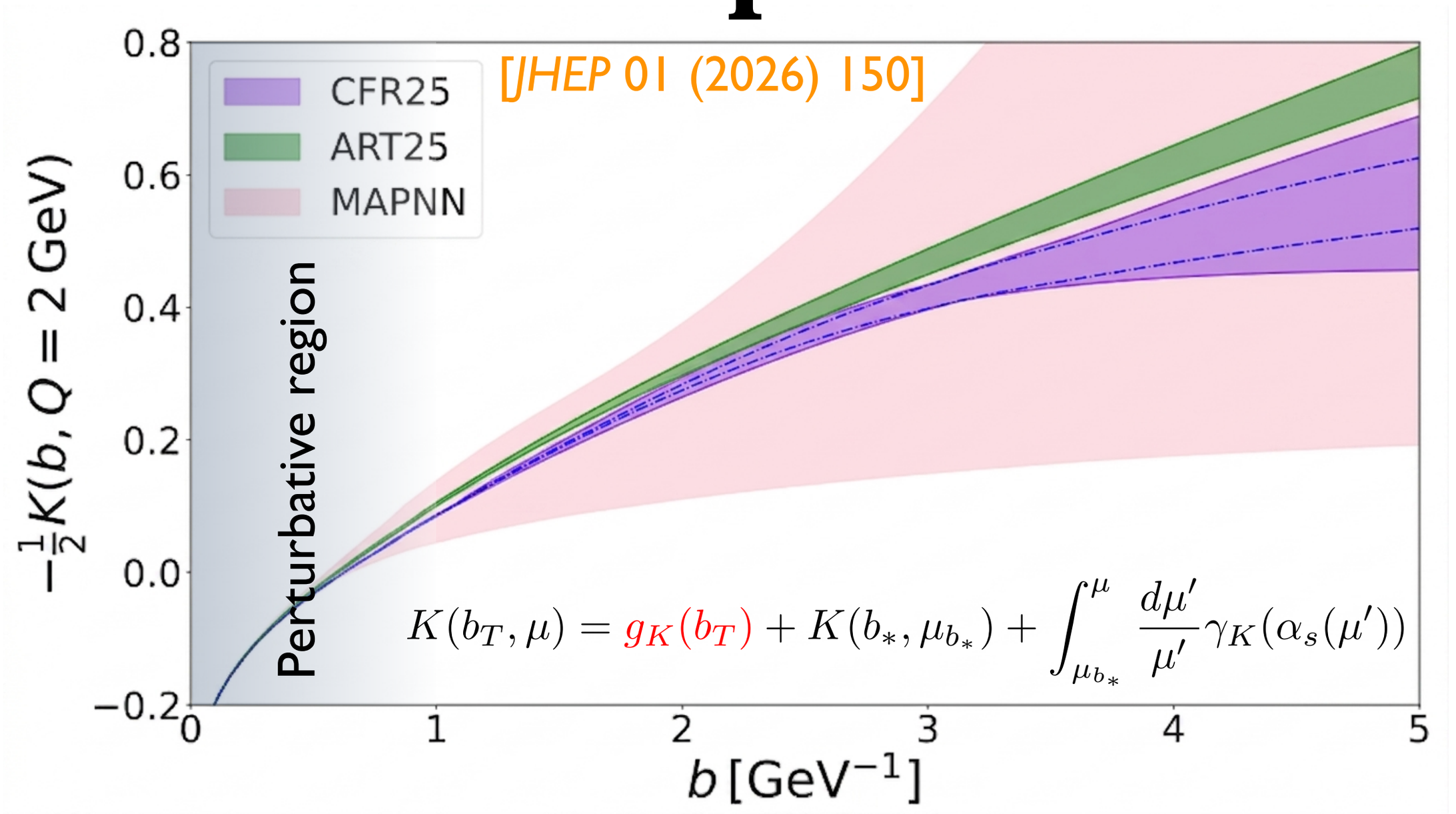
Purely perturbative

- Properties of $f_{\rm NP}$:
  - it parametrises $\mathcal{O}(\Lambda_{\rm QCD}/q_T)$ (non-perturbative) corrections,
  - it has to go to **one** as $b_{\rm T}$ goes to zero: reproduce the perturbative regime,
  - it has to got to **zero** as $b_{\rm T}$ becomes large: mimic the Sudakov suppression,
  - it does *not* depend on μ (ren. scale) but only on ζ (rapidity scale),
  - it is generally **flavour dependent**,
  - the ζ scaling is predictable, and flavour and x independent (Collins-Soper kernel).
- **Important**: $f_{\rm NP}$ is *not universal* as it depends on the specific b_* or, more in general, on the strategy used to regularise the Landau pole.

Non-perturbative effects



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].

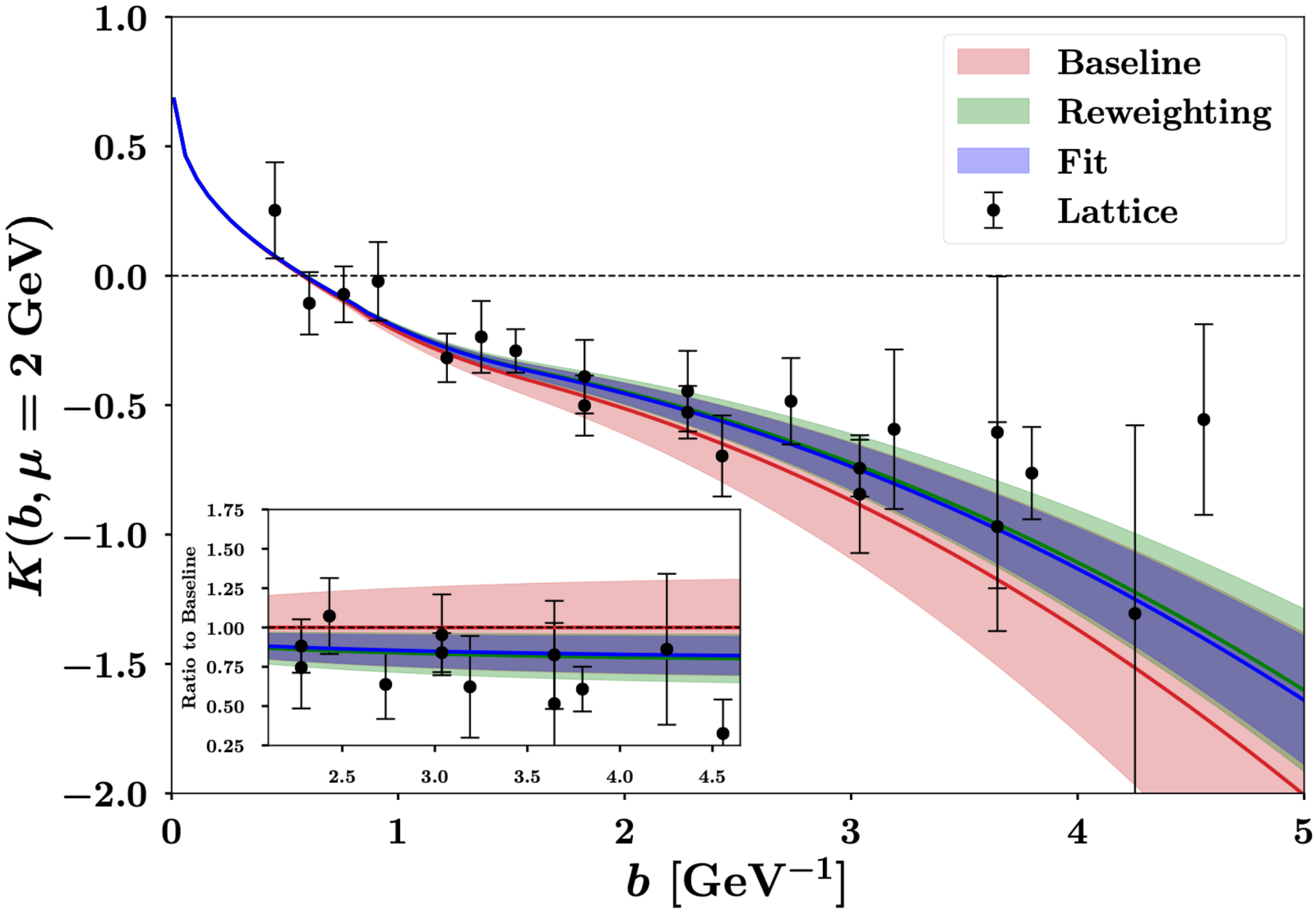
- We followed two different strategies:

  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.