

---

# Next-to-leading power endpoint DIS and PDF factorization

**Robert Szafron**

Brookhaven National Laboratory

with M. Beneke (TUM) and M. Schnubel (Nikhef)

[arXiv:2608.25502](#) (TUM-HEP-1600/26, Nikhef 2026-009)

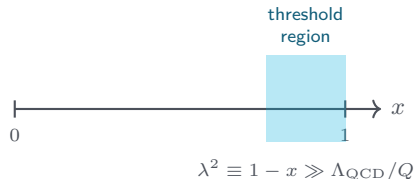
Precision QCD predictions for  $ep$  physics at the EIC (IV):  
High-Energy Physics at the Second Interaction Region  
CFNS, Stony Brook — 23 September 2026

# The question

What *is* a parton distribution function when  $x \rightarrow 1$ ,  
once you go one power beyond the leading one in  $(1 - x)$ ?

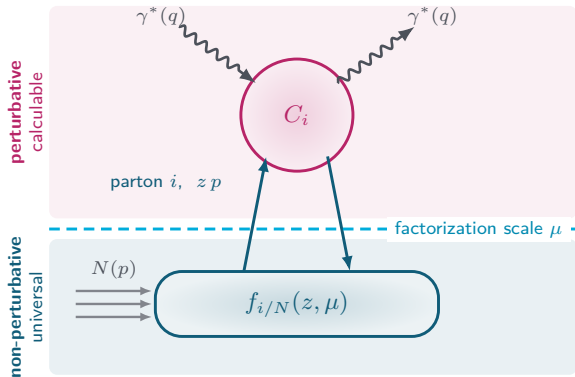
## Three claims:

1. Naive NLP factorization of threshold DIS **fails**:  
the new convolutions have **endpoint divergences**.
2. The PDF near the threshold becomes a **two-scale object at NLP**.
3. Once we cure the endpoint problem and resum the tower of  $\ln(1 - x)$ , the  $\overline{\text{MS}}$  PDF is **not** the natural scheme at the kinematic edge.



---

**Why threshold DIS beyond leading power**



$$F_2(x, Q^2) = x \sum_i \int_x^1 \frac{dz}{z} C_i\left(z, \frac{Q^2}{\mu^2}\right) f_{i/N}\left(\frac{x}{z}, \mu\right)$$

$C_i$  is calculable;  $f_{i/N}$  is universal, process-independent.

- ▶ The dashed line is a *choice*. Nothing physical happens at  $\mu$ .
- ▶ Move it up and the **coefficient** loses a logarithm that the **distribution** gains. The  $\mu$ -dependence has to cancel — **and that cancellation is DGLAP**.
- ▶ Everything sensitive to  $\Lambda_{\text{QCD}}$  ends up below the line, measured once and reused everywhere.

## RGE resummation

Single scale objects: each knows *one* scale, so each RGE resums physical logs.

## Threshold: a *second* factorization, inside $C$

As  $x \rightarrow 1$  a third scale appears between  $Q$  and  $\Lambda_{\text{QCD}}$ :

$$p_X^2 = Q^2 \frac{1-x}{x}.$$

The coefficient function is then no longer a single-scale object:

$$C(z) \supset \alpha_s^n \left[ \frac{\ln^{2n-1}(1-z)}{1-z} \right]_+.$$

Fixed order stops converging when  $\alpha_s \ln^2(1-x) \sim 1$ ; the tower has to be summed.

The cure is to factorize *again*, this time *inside*  $C$ : a **hard** function at  $Q^2$  times a **jet** function at  $Q^2(1-x)$ .

### How it was done

- ▶ **Sterman '87; Catani–Trentadue '89** — eikonal approximation and exponentiation in Mellin moments:  $\ln N$  exponentiates.
- ▶ **Korchensky, Marchesini '93** — the structure is the renormalization of *soft Wilson lines with a cusp*; the leading tower is governed by  $\gamma_{\text{cusp}}$ .
- ▶ **2003–06, SCET** — rederived as  $C = H(Q^2) J(Q^2(1-x))$ , each with its own RGE.

# Threshold DIS

As  $x \rightarrow 1$  the hadronic final state is squeezed,

$$p_X^2 = (p + q)^2 = Q^2 \frac{1 - x}{x} \ll Q^2,$$

and fixed-order coefficient functions develop

$$\alpha_s^n \frac{\ln^{2n-1}(1 - x)}{1 - x} \quad (\text{LP}).$$

— resummed at LP by the second factorization of the previous slide.

One power down, at *leading twist*, sits

$$\alpha_s^n \ln^{2n-1}(1 - x) \quad (\text{NLP}),$$

which is where factorization itself stops being automatic.

## Two independent expansions

**twist**  $\xi = \Lambda_{\text{QCD}}/Q$

**threshold**  $\lambda^2 = 1 - x$

This talk: **leading twist**, **NLP in  $1 - x$** .

NLP in  $1 - x$  is *not* higher twist — different expansion, different physics.

# The off-diagonal channel: the cleanest NLP laboratory

We take a scalar probe  $\phi^*$  coupled to  $G_{\mu\nu}^A G^{\mu\nu A}$ , so the gluon-initiated hard process appears already at tree level:

$$W_\phi = \frac{1}{8\pi Q^2} \int d^4x e^{iqx} \langle N(p) | [G^2](x) [G^2](0) | N(p) \rangle.$$

The **off-diagonal** contribution starts at NLP: the struck parton is a *quark*.

- ▶ leading non-vanishing order = NLP  $\Rightarrow$  no LP contamination to disentangle;
- ▶ displays *all* the structural features of NLP factorization;
- ▶ its anomalous dimension is the large- $x$  limit of  $P_{gq}$ .

## What was known

LL resummation of this channel was obtained from  $d$ -dimensional consistency conditions [Beneke, Garny, Jaskiewicz, Szafron, Vernazza, Wang '20], matching the all-order large- $x$  structure of the off-diagonal splitting functions [Vogt '10; Almasy, Soar, Vogt '11].

The SCET factorization theorem behind it was only *sketched*.

## This work

The complete operator-level factorization theorem, the definition of the PDF it implies, and the resummation in two schemes.

---

## A short guide to SCET

QCD, reorganized for processes with energetic, nearly light-like particles. Order by order in  $\lambda$  nothing is lost: the fields are split according to the momentum regions that actually contribute, and each *mode* becomes its own field with a definite power counting in a small parameter  $\lambda$ .

## Light-cone bookkeeping

Two light-like directions,  $n_+^2 = n_-^2 = 0$ ,  
 $n_+ \cdot n_- = 2$ :

$$p^\mu = n_+ p \frac{n_-^\mu}{2} + p_\perp^\mu + n_- p \frac{n_+^\mu}{2}, \quad p^2 = n_+ p n_- p + p_\perp^2.$$

A mode is defined by how  $(n_+ p, p_\perp, n_- p)$  scale with  $\lambda$ . **Two modes can share a virtuality and still be different modes.**

## Two things get expanded

1. **Operators.** A QCD current is matched onto SCET operators; the **hard** fluctuations are integrated out into Wilson coefficients,

$$J_{\text{QCD}} = C^{A0} J^{A0} + C^{B1} J^{B1} + \dots$$

each term one power of  $\lambda$  down.

2. **The Lagrangian.**  $\mathcal{L} = \mathcal{L}^{(0)} + \mathcal{L}^{(1)} + \mathcal{L}^{(2)} + \dots$  At leading power the low-energy gluons couple eikonally and can be *removed by a field redefinition*. That is what makes factorization happen in SCET language.

$$p^\mu = n_+ p \frac{n_-^\mu}{2}, \quad q^\mu = -Q \frac{n_-^\mu}{2} + Q \frac{n_+^\mu}{2}, \quad x = \frac{Q}{n_+ p}.$$

Two small parameters:

$$\lambda = \sqrt{1-x}, \quad \xi = \frac{\Lambda_{\text{QCD}}}{Q}.$$

On-shell modes,  $p = (n_+ p, p_\perp, n_- p)$ :

|                          |                                                  |
|--------------------------|--------------------------------------------------|
| collinear (struck quark) | $Q(1, \xi, \xi^2)$                               |
| anti-hardcollinear (jet) | $Q(\lambda^2, \lambda, 1)$                       |
| soft-collinear (remnant) | $Q(\lambda^2, \xi\beta, \xi^2(\beta/\lambda)^2)$ |
| hardcollinear            | $Q(1, \xi/\lambda, \xi^2/\lambda^2)$             |

## Getting the soft-collinear scaling right

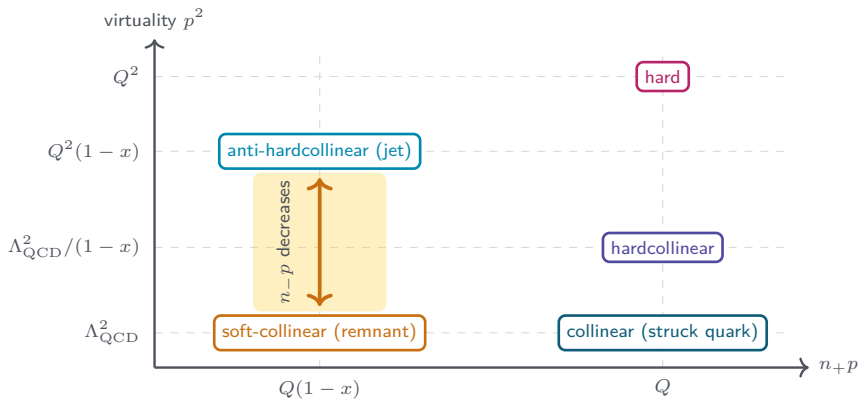
The only requirements are  $n_+ p_{sc} \sim Q(1-x)$  and  $p_{\perp sc}, n_- p_{sc} \rightarrow 0$  as  $\xi \rightarrow 0$ ; write  $\beta \sim \lambda^a$ .

$a = 1$  was assumed in [Becher, Neubert, Pecjak '06];  
 $a = 2$  appears for the QED electron PDF. But any  $a > 0$  gives virtuality  $\Lambda_{\text{QCD}}^2 (1-x)^a$  below  $\Lambda_{\text{QCD}}^2$  — impossible in a confining theory.

The derivation must work for  $a = 0$ .

The **hardcollinear** mode, virtuality  $\Lambda_{\text{QCD}}^2/(1-x)$ , is *irrelevant at LP* but **essential at NLP**. It is perturbative and  $Q$ -independent.

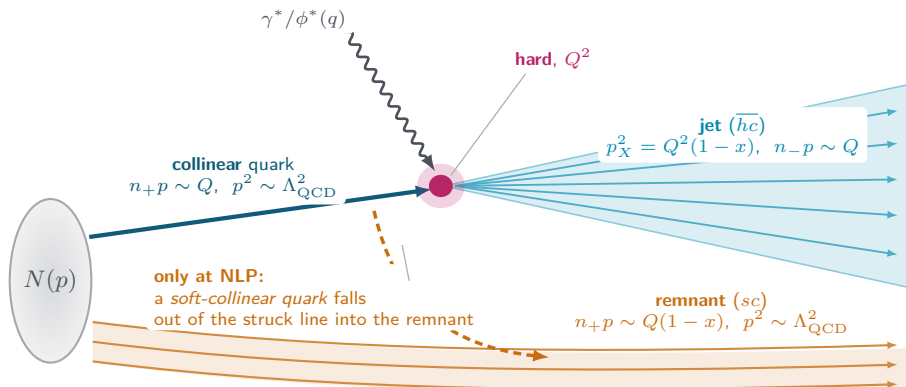
# The same modes



The **remnant** and the **jet** sit at the *same*  $n+p \sim Q(1-x)$  and differ only in  $n-p$ , which runs down the shaded ladder from  $Q$  to  $\Lambda_{\text{QCD}}^2/(Q(1-x))$ .

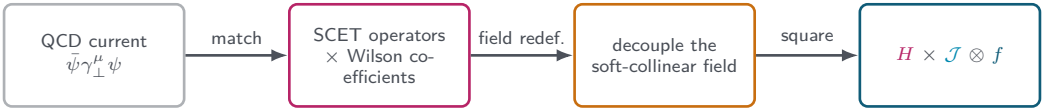
At LP the ladder is empty and can be ignored. At NLP it is populated, and the logarithmic integration along it *is* the endpoint divergence.

# The same modes, as an event



At leading power the story is short: a **collinear** quark is struck by a **hard** photon, everything it makes goes into the **jet**, and the **remnant** "decouples" from short distance physics. Every NLP effect in this talk is a way for those three to stop being independent.

# Leading power factorization theorem



The soft-collinear gluons couple eikonally, become Wilson lines, and the final state splits,  $|X\rangle = |X_{hc}\rangle |X_{sc}\rangle$ .

| factor                   | where it comes from                                                                                                              | lives at                 |
|--------------------------|----------------------------------------------------------------------------------------------------------------------------------|--------------------------|
| $H =  C^{A0} ^2$         | the squared Wilson coefficient of the matched current, purely virtual, everything of virtuality $Q^2$ , integrated out in step 1 | $Q^2$                    |
| $\mathcal{J}_{hc}^{(q)}$ | the anti-hardcollinear final state created at the vertex,<br>$\sum_{X_{hc}}  \langle X_{hc}   \bar{\chi}_{hc}   0 \rangle ^2$    | $Q^2(1-x)$               |
| $f_{q/N}$                | the nucleon matrix element left behind: a collinear quark field dressed by soft-collinear Wilson lines                           | $\Lambda_{\text{QCD}}^2$ |

Each factor depends on a single scale; each has its own RGE and running between the scales enables the resummation of large logs.

## Leading power: the threshold PDF is the $\overline{\text{MS}}$ PDF

Matching  $\bar{\psi}\gamma_{\perp}^{\mu}\psi \rightarrow \tilde{C}^{A0} \bar{\chi}_{hc}\gamma_{\perp}^{\mu}\chi_{hc}$  and decoupling soft-collinear gluons,

$$W^{\mu\nu} = -g_{\perp}^{\mu\nu} \frac{1}{2x} |C^{A0}(n+p, Q)|^2 Q^2 \int_x^1 d\xi \mathcal{J}_{hc}^{(q)}\left(Q^2 \frac{\xi-x}{x}\right) f_{q/N}(\xi).$$

No extra matching coefficient: at LP the threshold PDF coincides with the  $\xi \rightarrow 1$  limit of the  $\overline{\text{MS}}$  PDF, and  $\overline{\text{MS}}$  renormalization *commutes* with the  $\xi \rightarrow 1$  limit. The DGLAP kernel comes out expanded in the  $x \rightarrow 1$  limit.

**A single scale.** Everything below the jet lives at  $\Lambda_{\text{QCD}}^2$ .  
This is exactly what breaks at NLP.

# Opening up the PDF

Push the PDF operator to  $\xi \rightarrow 1$  and match the **collinear** quark field onto a *frozen-collinear* one: a field whose large momentum component can only have small fluctuations.

$$f_{q/N}^{\text{thr,LP}}(\xi) = n_+p \underbrace{K_{q/N}}_{\text{the nucleon}} \underbrace{|C_\Lambda|^2}_{\text{virtual}} \underbrace{S_{sc}^{\text{LP}}(n_+p(1-\xi))}_{\text{radiated}}$$

$K_{q/N}$  — the nucleon.

Matrix element of the frozen-collinear field. Non-perturbative;  $K_{q/q} = 1$  on a quark state.

$|C_\Lambda|^2$  — virtual collinear fluctuations.

$$\chi_c \rightarrow \overline{Y}_{sc,n_+}^\dagger C_\Lambda \chi_{fc},$$

the struck quark's collinear virtuality, integrated out. Non-perturbative.

$S_{sc}^{\text{LP}}$  — what was radiated.

A *vacuum* matrix element of the soft-collinear Wilson lines, measuring the momentum  $\Omega = n_+p(1-\xi)$  that did *not* stay with the quark.

All three sit at the *same* virtuality  $\Lambda_{\text{QCD}}^2$ , so the PDF is single-scale — and the product is exactly the  $\xi \rightarrow 1$  limit of the  $\overline{\text{MS}}$  PDF. **Both fail at NLP.**

---

## The two NLP mechanisms

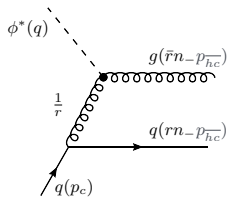


# Hard matching: the A0 and B1 operators

$$[G^A_{\mu\nu}G^{\mu\nu A}](0) = \frac{1}{g_s^2}\int dt\,d\bar t\,\tilde C^{\textcolor{red}{A}0}_g\,J^{\textcolor{red}{A}0} + \int dt\,d\bar t_1d\bar t_2\,\tilde C^{\textcolor{red}{B}1}_g\,J^{\textcolor{red}{B}1}$$

$$J^{\textcolor{red}{A}0}(t,\bar t) = 2g_{\mu\nu}\,n_-\partial\,\textcolor{blue}{\mathcal A}^{\nu A}_{\overline{hc\perp}}(\bar tn_-)\,n_+\partial\,\textcolor{blue}{\mathcal A}^{\mu A}_{hc\perp}(tn_+)$$

$$J^{\textcolor{red}{B}1}(t,\bar t_1,\bar t_2) = \bar\chi_{\overline{hc}}(\bar t_1n_-)\,\textcolor{blue}{\mathcal A}^{\mu A}_{\overline{hc\perp}}(\bar t_2n_-)\,T^A\gamma_{\perp\mu}\,\chi_{hc}(tn_+)$$



B1 carries one extra  $\textcolor{blue}{\mathcal A}_{\overline{hc\perp}}$ :  $\mathcal{O}(\lambda)$  in the amplitude,  $\mathcal{O}(\lambda^2) \sim \mathcal{O}(1-x)$  in  $W_\phi$ . At tree level, with  $r$  the quark share of  $n-p_{\overline{hc}} \equiv Q$ ,

$$\textcolor{red}{C}_g^{\textcolor{red}{B}1}(n_+p_c,rQ,\bar rQ) = -\frac{2\bar r}{r}$$

Note  $1/r$ !

# The anti-hardcollinear term, and its endpoint divergence

Decoupling soft-collinear gluons from the jet gives a factorization of the *expected* shape, with a generalized (composite-operator) jet function:

$$W_{\phi}^{\overline{hc}} = C_g^{\text{B1}*}(r') C_g^{\text{B1}}(r) \otimes^{r,r'} \mathcal{J}_{hc}^{(qg)}(\xi, r, r') \otimes^{\xi} f_{q/N}(\xi)$$

$$\mathcal{J}_{hc}^{(qg)}(k^2, r, r') = \frac{\alpha_s C_F}{4\pi} \frac{d-2}{\Gamma(1-\epsilon)} \delta(r-r') r \left( \frac{\mu^2 e^{\gamma_E}}{k^2 r \bar{r}} \right)^{\epsilon}$$

Put the tree-level pieces together:

$$\int_0^1 dr dr' C_g^{\text{B1}*}(r') C_g^{\text{B1}}(r) \mathcal{J}_{hc}^{(qg)} \propto \int_0^1 dr \frac{\bar{r}^{2-\epsilon}}{r^{1+\epsilon}} \implies \text{divergent at } r \rightarrow 0.$$

Regulated by  $\epsilon$ , but that is not a fix: one cannot renormalize  $C^{\text{B1}}$  and  $\mathcal{J}^{(qg)}$  separately, solve their RGEs, and *then* integrate over  $r$ .

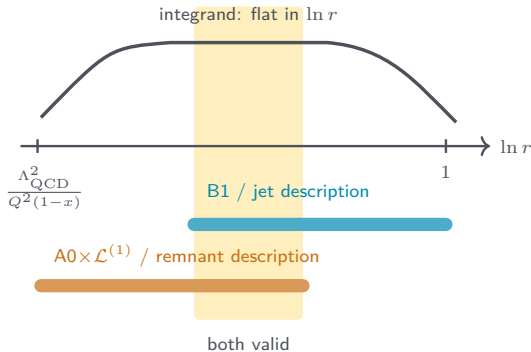
# What that divergence is actually telling us

The integrand is flat in  $\ln r$ . A flat integrand over a logarithmic range is not a pathology — it is a **logarithm**, and the range is set by where each description stops being true.

$C^{B1}$  was derived assuming the emitted quark is **energetic**: fine for  $r \sim 1$ , wrong once  $rQ$  drops to the remnant scale. Symmetrically, the **soft-collinear** description fails once  $\omega$  grows to  $\mathcal{O}(Q)$ .

Each term, integrated over all  $r$ , reaches beyond its natural scale. Dimensional regularization hides this: it does not tell you where the boundary is.

The overlap is the point: two descriptions are valid at once, so **the same physics sits in both terms**, and each diverges because neither knows where to stop.



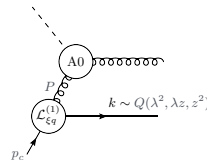
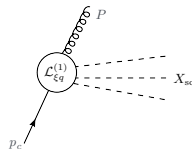
# The soft-collinear term: a subleading Lagrangian

The  $\mathcal{O}(\lambda)$  soft-quark interaction, boosted to  $x \rightarrow 1$  DIS kinematics, converts the struck **collinear** quark into a **hardcollinear** gluon plus a **soft-collinear** quark:

$$\mathcal{L}^{(1)}(y) = [\bar{\eta}_{sc} Y_{sc, n_-}](y_-) \mathcal{A}_{hc\perp}^{(0)}(y) \chi_c^{(0)}(y_+) + \text{h.c.}$$

with  $\eta_{sc} = -\frac{\not{n}_+}{2} \frac{1}{in_+ D_{sc}} i \not{D}_{\perp sc} \xi_{sc}$  the *small* component of the soft-collinear spinor.

Why the emitted quark must be soft-collinear: the struck quark carries almost all of the proton momentum and must hand nearly all of it to the gluon. The left over has  $n_+ p \sim Q(1-x)$ , the same order as the remnant.



Hard part is unchanged — the LP **A0** current.  
The suppression is entirely in the Lagrangian insertion.

# Soft *quarks*, not soft *gluons*

At leading power the only thing the jet and the remnant exchange is a soft *gluon*. Soft gluons are easy: they couple eikonally, they do not change the colour representation of the line they attach to, they exponentiate, and the result is the familiar Sudakov with  $\gamma_{\text{cusp}}$  and  $C_F$ .

A soft *quark* does none of that. It is a fermion, it cannot be written as a Wilson line, and it **changes the emitting line from a quark into a gluon**.

Three consequences, all visible later in the talk:

If you take one structural thing away from the NLP literature, it is this: **soft-quark exchange is where the new physics and the new difficulties both live**.

1. No  $1/(1-x)$  enhancement, which is exactly why this is NLP:  $\alpha_s^n \ln^{2n-1}(1-x)$  and not  $\alpha_s^n \left[ \ln^{2n-1}(1-x)/(1-x) \right]_+$ .
2. The exponent involves  $C_F - C_A$ : the *mismatch* between the colour charge of the line before and after the conversion, not a single cusp charge.
3. The answer is not a plain exponential of a cusp. In  $\overline{\text{MS}}$  it comes out as a Bernoulli series.

# The NLP gluon PDF becomes a two-scale object

The soft-collinear term has the LP *shape*,

$$W_{\phi}^{sc} = |C_g^{A0}|^2 \mathcal{J}_{hc}^{(g)}(\xi) \otimes^{\xi} f_{g/N}^{\text{thr,NLP}}(\xi),$$

except that the distribution is the *NLP term* in the threshold expansion of the gluon PDF. **That object itself factorizes:**

$$f_{g/N}^{\text{thr,NLP}}(\xi) = D_g^{\text{B1}*}(\omega') D_g^{\text{B1}}(\omega) \otimes^{\omega, \omega'} S_{sc, \text{sub}}^{\text{NLP}}(1 - \xi', \omega, \omega') \otimes^{\xi'} f_{q/N}^{\text{thr,LP}}(1 + \xi - \xi')$$

- ▶ The *gluon* distribution at NLP is a convolution with the *quark* distribution at LP.
- ▶  $D_g^{\text{B1}}$  is a hardcollinear matching coefficient at virtuality  $\Lambda_{\text{QCD}}^2/(1-x) \gg \Lambda_{\text{QCD}}^2$ : **perturbative**, yet sitting *inside* the PDF, because its  $k_{\perp} \sim \Lambda_{\text{QCD}}/\sqrt{1-x}$  is still below  $\mu$ .

## Consequence

Two scales inside the PDF  $\Rightarrow$  large  $\ln(1-x)$  *within* the PDF, **not controlled by DGLAP**. They must be summed by other means.

# Where the PDF's second scale comes from

The struck quark carries almost all of the proton's momentum and has to hand nearly all of it to the gluon.

What remains is a soft-collinear quark with

$n_+p \sim Q(1-x)$  and  $k_\perp \sim \Lambda_{\text{QCD}}$ ; the gluon is off-shell by

$n_+p n_-k \sim \Lambda_{\text{QCD}}^2/(1-x)$ , so the hardcollinear fluctuations have

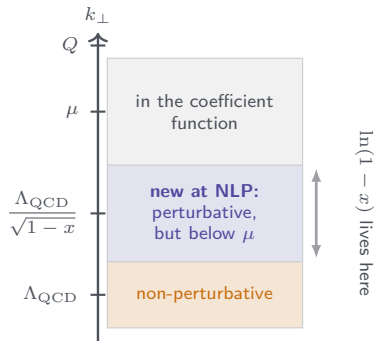
$$k_{\perp,hc} \sim \frac{\Lambda_{\text{QCD}}}{\sqrt{1-x}}.$$

Now apply the rules of collinear factorization literally:

- ▶  $k_\perp < \mu$ , so by definition this is **inside the PDF**;
- ▶  $k_\perp \gg \Lambda_{\text{QCD}}$ , so it is **calculable**.

A perturbative sub-structure, living inside a non-perturbative object.

DGLAP moves  $\mu$ ; it knows nothing about this second,  $x$ -dependent scale.



# Endpoint divergence

At lowest order the (large- $\omega$ , hence perturbative) soft-collinear function is

$$\mathcal{S}_{sc}^{\text{NLP,LO}}(\Omega, \omega, \omega') = \frac{\alpha_s C_F}{4\pi} \frac{2(d-2)}{\Gamma(1-\epsilon)} \omega \delta(\omega - \omega') \left( \frac{\mu^2 e^{\gamma_E}}{\Omega \omega} \right)^\epsilon,$$

so that

$$\mathcal{S}_{sc}^{\text{NLP,LO}}(\Omega) \propto \int_0^\infty \frac{d\omega}{\omega} \left( \frac{\mu^2 e^{\gamma_E}}{\Omega \omega} \right)^\epsilon \implies \text{UV divergent at } \omega \rightarrow \infty.$$

Keeping  $\epsilon \neq 0$ , the pole of the bare PDF is

$$f_{g/q}^{\text{thr,NLP}}(\xi) = \frac{\alpha_s C_F}{2\pi} \frac{1}{\epsilon} + \mathcal{O}(\epsilon^0) = \frac{1}{\epsilon} [P_{gq}(x)]_{x \rightarrow 1},$$

$$P_{gq}(x) = \frac{\alpha_s C_F}{2\pi} \frac{1 + (1-x)^2}{x}.$$

The off-diagonal splitting kernel emerges from an endpoint divergence in the **longitudinal** ( $n-p$ ) integration, *not* from a transverse-momentum divergence.

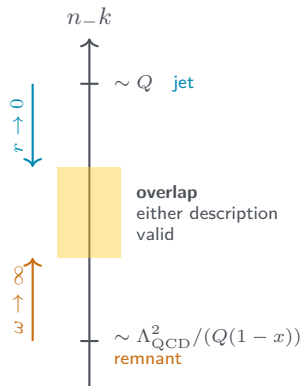
# One divergence, seen from two sides

- ▶ B1: divergent as  $r, r' \rightarrow 0$ .
- ▶  $A0 \times \mathcal{L}^{(1)}$ : divergent as  $\omega, \omega' \rightarrow \infty$ .

Both variables measure the *same* thing: the  $n_-k$  component.  
At fixed  $n_+k \sim Q(1-x)$ , lowering  $n_-k$  from  $\mathcal{O}(Q)$  to  $\mathcal{O}(\Lambda_{\text{QCD}}^2/(Q(1-x)))$  turns the **anti-hardcollinear** mode *smoothly* into the **soft-collinear** one.

In the overlap region the two descriptions must agree  $\Rightarrow$  **endpoint (re)factorization relations**.

This is a factorization in  $n_-k$ : **which mode owns the parton**; not the transverse-momentum factorization of ordinary collinear factorization. It appears only at threshold, and only from NLP.



# The NLP factorization theorem — in $d$ dimensions

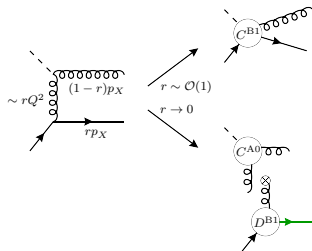
$$\begin{aligned}
 W_{\phi}^{\text{NLP}} = & \frac{1}{4x} \int_0^1 dr dr' C_g^{\text{B1}*}(n_+p, r'Q, \bar{r}'Q) C_g^{\text{B1}}(n_+p, rQ, \bar{r}Q) \\
 & \times \int_x^1 d\xi \mathcal{J}_{\frac{qg}{hc}}^{(qg)}\left(\frac{Q^2(\xi-x)}{x}, r, r'\right) f_{q/N}^{\text{thr,LP}}(\xi) \\
 & + |C_g^{\text{A0}}(n_+p, Q)|^2 \frac{Q^2}{x^2} \int_x^1 d\xi \mathcal{J}_{\frac{g}{hc}}^{(g)}\left(Q^2 \frac{\xi-x}{x}\right) f_{g/N}^{\text{thr,NLP}}(\xi)
 \end{aligned}$$

- ▶ Complete at NLP for the off-diagonal channel (antiquark term identical, with  $f_{\bar{q}/N}$ ).
- ▶ **Strictly defined only for  $d \neq 4$ :** the  $r^{(\prime)}$  and  $\omega^{(\prime)}$  convolutions are endpoint divergent; the poles cancel only in the *sum*, *after* the convolutions.

Correct, but there are no four-dimensional, single-scale component functions that could undergo RG evolution.

---

## Refactorization and the NLP PDF



$C_g^{B1}$  depends on two scales,  $\tilde{Q}^2 = n_+ p_c Q$  and  $r\tilde{Q}^2$ . The derivation assumed  $r \sim \mathcal{O}(1)$ ; but  $r$  is then integrated down to 0, where the internal propagator is *not* hard and the emitted quark is **soft-collinear**.

The right object in the endpoint region is  $C_g^{A0} D_g^{B1} / r$ , and matching the two gives

$$\llbracket C_g^{B1}(n_+ p_c, rQ, \bar{r}Q) \rrbracket = -2 C_g^{A0}(n_+ p_c, Q) \frac{D_g^{B1}(r\tilde{Q}^2)}{r}$$

A second relation does the same job for the jet function:

$$\llbracket \mathcal{J}_{\frac{hc}{hc}}^{(qg)}(Q\Omega, r, r') \rrbracket = \frac{Q}{2} \int d\Omega' \mathcal{J}_{\frac{hc}{hc}}^{(g)}(Q\Omega') \llbracket S_{sc,sub}^{NLP}(\Omega - \Omega', \omega, \omega') \rrbracket_{\omega^{(\prime)} = r^{(\prime)} Q}$$

$\llbracket \cdot \rrbracket$  denotes the leading endpoint behaviour:  $r^{(\prime)} \rightarrow 0$ , or large  $\omega^{(\prime)}$ .

# Endpoint subtraction

## The familiar version

At NLO the real-emission integral diverges in the soft and collinear limits, and the virtual has a pole. Nobody calls this a crisis: you build a counterterm matching the singular limit exactly, subtract it from the real, add it back to the virtual.

$$\sigma = \underbrace{\int (R - S)}_{\text{finite in } d=4} + \underbrace{\int S + V}_{\text{finite}}$$

Nothing about  $\sigma$  changed. Each piece became usable.

## The version in this talk

The **B1** term diverges as  $r \rightarrow 0$ , the **A0** term as  $\omega \rightarrow \infty$ , and the refactorization relation says *the same function describes both* in the overlap. We subtract it from each term on complementary ranges; the two subtractions add up to a *scaleless* integral.

$$W = \underbrace{\left[ B1 - \llbracket B1 \rrbracket \big|_{r < \Lambda/Q} \right]}_{\text{finite}} + \underbrace{\left[ A0 - \llbracket A0 \rrbracket \big|_{\omega > \Lambda} \right]}_{\text{finite}}$$

The one new feature: the boundary  $\Lambda$  is a *choice*, like a subtraction scheme. It cancels in the sum — and because it decides which term carries the large logarithms, choosing it well is what makes the resummation easy.

# Endpoint rearrangement: a four-dimensional theorem

The combination

$$\begin{aligned} 0 &= (\text{A0 with } \llbracket \mathcal{S}_{sc,sub}^{\text{NLP}} \rrbracket) \\ &= (\text{B1 with } \llbracket \mathcal{C}^{\text{B1}} \rrbracket, \llbracket \mathcal{J}^{(qg)} \rrbracket) \end{aligned}$$

is **scaleless** — and the two forms are equal by the refactorization relations.

Subtract it from the theorem, split at a scale  $\Lambda$ ,

$$\Lambda_{\text{QCD}}^2 / (Q(1-x)) \ll \Lambda \leq Q,$$

$$R_1 : \omega, \omega' < \Lambda, \quad R_2 = \text{complement},$$

and give  $R_1$  to the B1 term,  $R_2$  to the A0 term.

## Result

- ▶  $r, r' \rightarrow 0$  regulated by the  $R_1$  subtraction;
- ▶  $\omega^{(\prime)} \rightarrow \infty$  regulated by the  $R_2$  subtraction;
- ▶  $d \rightarrow 4$  can now be taken;
- ▶  $\Lambda$  cancels between the two terms.

## The lever

$\Lambda$  decides *which term carries the logarithms*. For  $\Lambda = Q$  all LL are in the A0 term. This is convenient because the B1 hard and jet functions obey integro-differential RGEs beyond LL.

# A new scheme for the NLP PDF

$$f_{g/N}^{\text{NLP},\textcircled{R}}(\Lambda, \xi) = \frac{1}{2} \int d\xi' \left\{ \int_0^\Lambda \frac{d\omega}{\omega} \frac{d\omega'}{\omega'} D_g^{\text{B1}} D_g^{\text{B1}*} S_{sc,\text{sub}}^{\text{NLP}} \right. \\ \left. + \int_{R_3} \frac{d\omega}{\omega} \frac{d\omega'}{\omega'} D_g^{\text{B1}} D_g^{\text{B1}*} \left[ S_{sc,\text{sub}}^{\text{NLP}} - \llbracket S_{sc,\text{sub}}^{\text{NLP}} \rrbracket \right] \right\} f_{q/N}^{\text{thr,LP}}(1 + \xi - \xi')$$

$R_3$ : exactly one of  $\omega, \omega'$  above  $\Lambda$ . First line cut off at  $\Lambda$ ; second line regulated by subtraction. Manifestly endpoint-finite.

- Renormalization and convolution now **commute**.
- One shows  $Z_{gq}^{\text{NLP},\textcircled{R}} = 0$ : the scheme renormalizes through its component functions alone.

This is a genuinely different PDF

$f_{g/N}^{\text{NLP},\textcircled{R}}$  differs from the  $\overline{\text{MS}}$  NLP PDF **already at LL**.

$\overline{\text{MS}}$  does the divergent convolution in  $d$  dimensions and subtracts the total pole afterwards, which does not distinguish UV from endpoint divergences, and obscures the  $x \rightarrow 1$  physics.

---

## Resummation and the scheme relation

# LL resummation in the refactorization scheme

All component RGEs take the cusp form

$$\frac{dh}{d \ln \mu} = C_h \gamma_{\text{cusp}} \ln \frac{-Q_X^2}{\mu^2} h,$$

with  $\gamma_{\text{cusp}} = \alpha_s/\pi + \dots$ , and *dynamical* matching scales

$$\begin{aligned} \mu_Q^2 &\sim Q^2, & \mu_{hc}^2 &\sim \frac{Q^2}{N}, \\ \mu_{D^{\text{B1}}}^2 &\sim n_+ p \omega, & \mu_{sc}^2 &\sim \frac{n_+ p \omega}{N}. \end{aligned}$$

Set at every value of  $\omega$ , so all LL sit in the evolution kernels.

$\Lambda = Q$ : everything collapses

$$W_\phi^{\text{NLP,LL}} = \frac{Q^2}{x^2} |C_g^{\text{A0}}(n_+, p, Q)|^2 \mathcal{J}_{hc}^{(g)} f_{g/N}^{\text{NLP}, \text{R}}(Q)$$

Written out, with  $S_X \equiv S(\mu_X, \mu)$ , the exponents are

$$\begin{aligned} 4C_A(S_Q - S_{hc}) & \quad \text{hard} \times \text{jet} \\ 2(C_F - C_A)(S_D + S_{D'} - S_{sc} - S_{sc'}) & \quad \text{in the } \omega \text{ integrals} \end{aligned}$$

- ▶  $\mu$ -dependence, including that of the PDF, cancels to LL.
- ▶ The initial condition of  $S_{sc, \text{sub}}^{\text{NLP}}$  is non-perturbative.
- ▶ The characteristic colour structure is  $C_F - C_A$ : a soft-*quark*, not soft-gluon, Sudakov.

## The $\overline{\text{MS}}$ route: resumming poles instead of logarithms

Keep  $d \neq 4$ , resum the leading  $1/\epsilon$  poles of the *bare* functions. The off-diagonal counterterm packages into a closed form,

$$Z_{gq}^{\text{NLP}} = \frac{1}{2NL_N} \frac{C_F}{C_F - C_A} \exp\left[\frac{\alpha_s C_F}{\pi} \frac{L_N}{\epsilon}\right] F_{\text{pole}}\left[-\epsilon L_N, \frac{\alpha_s}{\pi} (C_F - C_A) L_N^2\right],$$

and the single pole gives the all-order LL anomalous dimension

$$\gamma_{gq}^{\text{LL,NLP}}(N) = -\frac{\alpha_s C_F}{\pi} \frac{1}{N} \mathcal{B}_0\left(\frac{\alpha_s}{\pi} (C_F - C_A) L_N^2\right), \quad \mathcal{B}_0(z) = \sum_{n \geq 0} \frac{B_n}{(n!)^2} z^n.$$

This sums  $\alpha_s^{n+1} \ln^{2n} N$  in  $P_{gq}(N)$  to all orders, with the colour structure  $C_F(C_F - C_A)^n$  and the Bernoulli series, reproducing the known large- $x$  DLA result for the off-diagonal splitting functions.

$L_N = \ln N$ ; Bernoulli numbers  $B_n$ ; cf. Vogt '10, Almasy–Presti–Vogt '11, Beneke et al. '20.

Now derived from an operator-level factorization theorem, rather than from  $d$ -dimensional consistency alone.

# The two PDFs differ — and the difference is calculable

Write  $K = \frac{1}{2NL_N} \frac{C_F}{C_F - C_A}$  and  $E(z) = \exp \left[ \frac{\alpha_s}{\pi} (C_F - C_A) \left( L_N \ln z + \frac{L_N^2}{2} \right) \right]$ . Then

$$\hat{V}_{gq}(\mu) = K \left\{ E\left(\frac{\mu^2}{\mu_0^2}\right) - \mathcal{B}_0\left(\frac{\alpha_s}{\pi}(C_F - C_A)L_N^2\right) \right\} \quad (\overline{\text{MS}})$$

$$\hat{V}_{gq}^{\textcircled{\text{R}}}(\Lambda, \mu) = K \left\{ E\left(\frac{\mu^2}{\mu_0^2}\right) - E\left(\frac{\mu^2}{\Lambda_Q^2}\right) \right\} \quad (\text{refactorization})$$

$$\Delta f_g^{\text{NLP}} = f_{g/N}^{\text{NLP}, \textcircled{\text{R}}}(\Lambda, \mu) - f_{g/N}^{\text{thr}, \text{NLP}}(\mu)$$

$$= \Delta V_{gq}^{\text{NLP}}(\Lambda) f_{q/N}^{\text{thr}, \text{LP}}(\mu)$$

$$\Delta V_{gq}^{\text{NLP}}(\Lambda) = \hat{V}_{gq}^{\textcircled{\text{R}}}(\Lambda, \mu) - \hat{V}_{gq}(\mu)$$

The Bernoulli numbers are an  $\overline{\text{MS}}$  artifact

They appear only in  $\overline{\text{MS}}$ . The refactorization-scheme expression is a plain exponential.

## What this means for PDFs at large $x$

---

- ▶ At LP the threshold PDF *is* the  $\overline{\text{MS}}$  PDF. At NLP it is not a single-scale object.
- ▶ Logarithms of  $1 - x$  live *inside* the PDF and are not controlled by DGLAP evolution.
- ▶  $\overline{\text{MS}}$  is not the most adequate definition at the kinematic edge, but the conversion factor is perturbative and universal, so nothing in current fitting practice is invalidated.
- ▶ The large- $x$  corner is precisely where PDFs are least constrained and where this structure is largest.
- ▶ No numerics yet: this work establishes the factorization theorem and the LL structure.

# Summary

---

A complete, operator-level NLP factorization theorem for threshold DIS in the off-diagonal channel — and the definition of the PDF that goes with it.

- ▶ Two mechanisms: the power-suppressed  $B1$  current, and  $A0 \times \mathcal{L}^{(1)}$  with a soft-collinear quark emitted into the remnant.
- ▶ Their convolutions are endpoint divergent — at  $r \rightarrow 0$  and at  $\omega \rightarrow \infty$ . The same divergence, seen from two sides.
- ▶ Resolved by refactorization relations; a subtracted soft-collinear function with a well-defined anomalous dimension.
- ▶ The threshold PDF is a two-scale object at NLP.
- ▶ LL resummation in two schemes; the  $\overline{MS}$  result with its Bernoulli series is recovered, and the scheme relation is computed to double-logarithmic accuracy.
- ▶ Correct soft-collinear scaling and the role of the hardcollinear mode clarified along the way.

---

Next: NLL resummation (in preparation).

**Thank you**

---

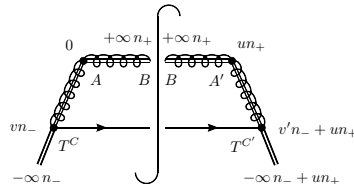
## Backup

$$\begin{aligned} \text{incoming} \quad Y_{sc,n_-}(x) &= \mathcal{P} \exp \left[ ig_s T^C \int_{-\infty}^0 ds n_- A_{sc}^C(x + sn_-) e^{s0^+} \right] \\ \text{outgoing} \quad \bar{Y}_{sc,n_+}^\dagger(x) &= \mathcal{P} \exp \left[ ig_s T^C \int_0^\infty ds n_+ A_{sc}^C(x + sn_+) e^{-s0^+} \right] \end{aligned}$$

- ▶ Adjoint version  $\mathcal{Y}_{sc,n_-}^{AB}$  with  $[T_{\text{ad}}^C]^{AB} = -if^{CAB}$ .
- ▶  $\bar{Y}_{sc,n} \neq Y_{sc,n}$ : the outgoing line is not the conjugate of the incoming one.
- ▶  $Y_n T^A Y_n^\dagger = T^B \mathcal{Y}_n^{BA}$ , and similarly with  $Y_n \rightarrow \bar{Y}_n^\dagger$ .
- ▶ Unitarity/reality of the adjoint line is *not* used — rapidity regulators can violate both.

# Backup: why soft subtraction, and how it differs from zero-bin

- ▶ The bare NLP soft-collinear matrix element mixes UV and IR structure; a *soft rearrangement* defines  $S_{sc,sub}^{NLP}$  with a well-defined UV anomalous dimension.
- ▶ Only then does each factor have a single scale and a standard RGE — the prerequisite for resummation.
- ▶ Conceptually related to LP zero-bin subtraction, but *not* the same: here the subtraction separates the soft-collinear remnant from the anti-hardcollinear jet in  $n_-k$ .



## Backup: $F[w, a]$ and the Bernoulli function

---

$$F[w, a] = \frac{w e^{a/w}}{e^w - 1}, \quad w = -\epsilon L_N, \quad a = \frac{\alpha_s}{\pi} (C_F - C_A) L_N^2$$

$$F_{\text{pole}}[w, a] = \sum_{k \geq 1} \frac{1}{w^k} \sum_{n \geq 0} \frac{B_n a^{n+k}}{n! (n+k)!}, \quad \mathcal{B}_0(z) = \sum_{n \geq 0} \frac{B_n}{(n!)^2} z^n$$

- ▶  $F$  splits into a pole part and a finite part; the pole part cancels against  $Z_{gq}^{\text{NLP}}$ .
- ▶  $w/(e^w - 1)$  and  $e^{\hat{S}_{FA}}$  are separately finite as  $\epsilon \rightarrow 0$ .
- ▶  $\mathcal{U}_{qq}^{\text{LP}}(\mu_0) = \exp\left[-\frac{\alpha_s C_F}{\pi} \left(L_N \ln \frac{\mu^2}{\mu_0^2} + \frac{L_N^2}{2}\right)\right]$ .

## Backup: the NLP conversion factor $V_{gq}$

---

Writing the bare threshold PDFs as

$$f_{i/N}^{\text{thr,NLP}}(\xi) = \sum_{j=g,q,\bar{q}} \int d\xi' V_{ij}(\xi') f_{j/N}^{\text{LP}}(1 - (\bar{\xi} - \bar{\xi}')),$$

the factorization gives an *operator definition*

$$V_{gq}(\xi) = \frac{1}{2} \int_0^\infty \frac{d\omega}{\omega} \frac{d\omega'}{\omega'} D_g^{\text{B1}}(-n_+ p \omega) D_g^{\text{B1}*}(-n_+ p \omega') S_{\text{sc,sub}}^{\text{NLP}}(n_+ p(1 - \xi), \omega, \omega'),$$

with  $\mathcal{Z}_{ij}^{\text{NLP}} = V_{ij} \mathcal{Z}_{jj}^{\text{LP}}$  for  $i \neq j$ .

$V_{ij}$  is *not* perturbatively calculable: the soft-collinear function lives at virtualities  $\sim \Lambda_{\text{QCD}}^2$  and must be evaluated in the non-perturbative vacuum. What *is* calculable is its large- $\omega$  behaviour.

## Backup: mode summary

| mode               | scaling $(n_+, p_\perp, n_-)$                    | virtuality                       | role                     |
|--------------------|--------------------------------------------------|----------------------------------|--------------------------|
| collinear          | $Q(1, \xi, \xi^2)$                               | $\Lambda_{\text{QCD}}^2$         | struck quark; in the PDF |
| anti-hardcollinear | $Q(\lambda^2, \lambda, 1)$                       | $Q^2(1 - x)$                     | the final-state jet      |
| soft-collinear     | $Q(\lambda^2, \xi\beta, \xi^2(\beta/\lambda)^2)$ | $\Lambda_{\text{QCD}}^2$         | proton remnant           |
| hardcollinear      | $Q(1, \xi/\lambda, \xi^2/\lambda^2)$             | $\Lambda_{\text{QCD}}^2/(1 - x)$ | NLP only; perturbative   |

- ▶  $\lambda = \sqrt{1 - x}$ ,  $\xi = \Lambda_{\text{QCD}}/Q$ ,  $\beta \sim \lambda^a$  with  $a \geq 0$  unspecified.
- ▶ No standard hardcollinear final-state mode: kinematics forbids it, and the virtual loops are scaleless.
- ▶ Soft modes are unconstrained in phase space  $\Rightarrow$  scaleless after squaring.