Physical and spurious thresholds in NNLO SIDIS hard functions.

The XXIII Annual Workshop on Soft-Collinear Effective Theory

Fabian Wunder

Seoul, Korea, March 4th, 2026

HELMHOLTZ

EBERHARD KARLS UNIVERSITÄT TÜBINGEN

INSTITUTE for NUCLEAR THEORY

DESY.

Overview.

> Why NNLO SIDIS?

> A closer look at the analytic structure of the hard function.

> What do we gain from removing spurious thresholds?

Talk based on:

Juliane Haug and FW (2025): *Single-valued representation of unpolarized and polarized semi-inclusive deep inelastic scattering at next-to-next-to-leading order*. In: Phys. Rev. D 112, 114036.

Scan & Cite

This work was supported by DFG through the Research Unit FOR 2926 (Project No. 409651613).

Clarification: What do we mean by SIDIS?

Two versions of semi-inclusive deep inelastic scattering:

SIDIS	also SIDIS
$\frac{\mathrm{d}\sigma_{ep\to eh+X}}{\mathrm{d}Q^2\,\mathrm{d}x\,\mathrm{d}z\,\mathrm{d}^2\mathbf{q}_\mathrm{T}}$	$\frac{\mathrm{d}\sigma_{ep\to eh+X}}{\mathrm{d}Q^2\,\mathrm{d}x\,\mathrm{d}z}$
> $\mathbf{q}_\mathrm{T}$-differential SIDIS (?!) > exclusive SIDIS (?!)	> $\mathbf{q}_\mathrm{T}$-integrated SIDIS (?!) > inclusive SIDIS (?!)

No consensus who has to specify their terminology.

For this talk: SIDIS = differential only in Q^2, x, and z.

Motivation: Why NNLO SIDIS?

- $e\,p \to e\,h + X$ key part of EIC physics program
- important for the extraction of (polarized) PDFs and FFs ($\to$ flavor separation)
- demand for higher-order corrections

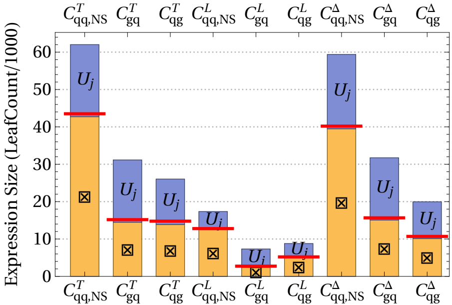
Recent developments

Threshold resummation	x and $z \to 1$	N^{2}LL, N^{3}LL [Abele '21,'22], N^{4}LL [Goyal '25]
	x or $z \to 1$	N^{2}LL [Forte '26]
Fixed order	NNLO QCD	unpol. [Goyal '23, Bonino '24]
		pol. [Bonino' 24, Goyal '24]
		W/Z (unpol.+pol) [Bonino '25, Bonino '25]
	QCD$\otimes$QED	[Goyal '25]

- Pol. PDF extraction (BDSSV24, MAPPDFpol1.0) uses N^{2}LL, will push to full NNLO.
- **Idea:** Analytically simplify published structure functions for simpler use

Motivation: What is there to simplify?

- **Observation:** Published structure functions contain unphysical case distinctions between four regions in (x, z)-plane of SIDIS
- We identify and remove this spurious analytical structure by expressing the structure functions in terms of single-valued polylogarithms (SVPs)
- **Size reduction: 30-60%**

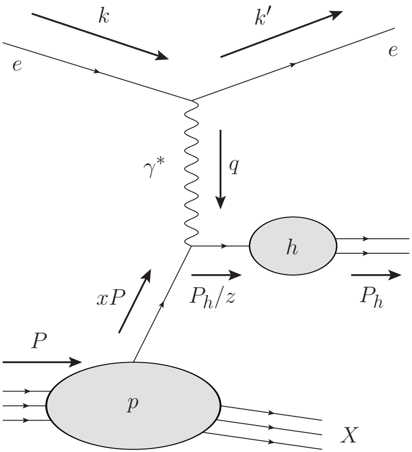


This talk:

How can we understand the simplification?

$\rightarrow$ Will lead us to discuss spurious branch-cuts/thresholds in real-virtual corrections of hard functions.

SIDIS in a nutshell.



- **SIDIS variables:**

$$x = \frac{Q^2}{2P \cdot q}, \quad z = \frac{P \cdot P_h}{P \cdot q}\,.$$

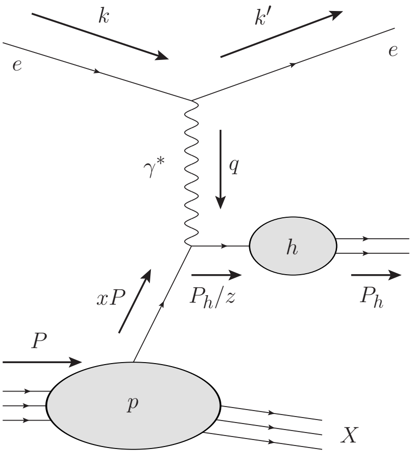
- **unpolarized cross-section:**

$$\frac{\mathrm{d}^3\sigma^h}{\mathrm{d}x\mathrm{d}y\mathrm{d}z} = \frac{4\pi\alpha^2}{Q^2}\left[\frac{1+(1-y)^2}{2y}\mathcal{F}_{\mathrm{T}}^h(x,z,Q^2) + \frac{1-y}{y}\mathcal{F}_{\mathrm{L}}^h(x,z,Q^2)\right]$$

- **Collinear factorization:**

$$\mathcal{F}_{\mathrm{T/L}}^h(x,z,Q^2) = \sum_{a,b}\int_x^1\frac{\mathrm{d}\hat{x}}{\hat{x}}\int_z^1\frac{\mathrm{d}\hat{z}}{\hat{z}}\, f_{a/p}\left(\frac{x}{\hat{x}},\mu_F^2\right) \times D_{h/b}\left(\frac{z}{\hat{z}},\mu_A^2\right)\mathcal{C}_{ba}^{\mathrm{T/L}}(\hat{x},\hat{z},\mu_R^2,\mu_F^2,\mu_A^2)$$

SIDIS in a nutshell - polarized.



> **SIDIS variables:**

$$x = \frac{Q^2}{2P\cdot q}, \qquad z = \frac{P\cdot P_h}{P\cdot q}\,.$$

> **polarized cross-section:**

$$\frac{\mathrm{d}^3\Delta\sigma^h}{\mathrm{d}x\mathrm{d}y\mathrm{d}z} = \frac{4\pi\alpha^2}{Q^2}(2-y)\,g_1(x,z,Q^2)$$

> **Collinear factorization:**

$$2g_1(x,z,Q^2) = \sum_{a,b}\int_x^1 \frac{\mathrm{d}\hat{x}}{\hat{x}}\int_z^1 \frac{\mathrm{d}\hat{z}}{\hat{z}}\,\Delta f_{a/p}\Big(\frac{x}{\hat{x}},\mu_F^2\Big)$$
$$\times\, D_{h/b}\Big(\frac{z}{\hat{z}},\mu_A^2\Big)\,\mathcal{C}^{\Delta}_{ba}\big(\hat{x},\hat{z},\mu_R^2,\mu_F^2,\mu_A^2\big)$$

Perturbative expansion of hard coefficient functions.

$$\mathcal{C}^{i}_{ba} = \mathcal{C}^{i,(0)}_{ba} + \frac{\alpha_s(\mu_R^2)}{2\pi}\,\mathcal{C}^{i,(1)}_{ba} + \left(\frac{\alpha_s(\mu_R^2)}{2\pi}\right)^2 \mathcal{C}^{i,(2)}_{ba} + \mathcal{O}(\alpha_s^3)$$

- up to NLO known for some time

  Furmanski, Petronzio (1982)
  De Florian, Stratmann, Vogelsang (1998)

- NNLO recently calculated

  Goyal, Lee, Moch, Pathak, Rana, Ravindran (2024)
  Bonino, Gehrmann, Löchner, Schönwald, Stagnitto (2024)

- NNLO channels: see right

- Rest of the talk: $\mathcal{C}^{i,(2)}_{ba}$

  Bonino et al. as baseline, but the presented steps work equally well for Goyal et al. results

$$C^{i,(2)}_{qq} = C^{i,(2)}_{\bar{q}\bar{q}} = e_q^2 C^{i,\text{NS}}_{qq} + \left(\sum_j e_{q_j}^2\right) C^{i,\text{PS}}_{qq}\,,$$

$$C^{i,(2)}_{\bar{q}q} = C^{i,(2)}_{q\bar{q}} = e_q^2 C^{i}_{\bar{q}q}\,,$$

$$C^{i,(2)}_{q'q} = C^{i,(2)}_{\bar{q}'\bar{q}} = e_q^2 C^{i,1}_{q'q} + e_{q'}^2 C^{i,2}_{q'q} + e_q e_{q'} C^{i,3}_{q'q}\,,$$

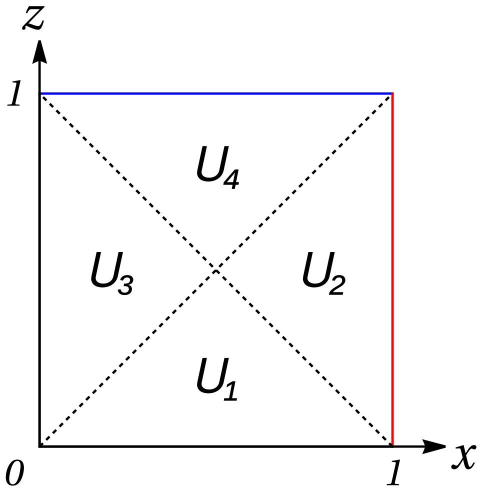
$$C^{i,(2)}_{\bar{q}'q} = C^{i,(2)}_{q'\bar{q}} = e_q^2 C^{i,1}_{q'q} + e_{q'}^2 C^{i,2}_{q'q} - e_q e_{q'} C^{i,3}_{q'q}\,,$$

$$C^{i,(2)}_{gq} = C^{i,(2)}_{g\bar{q}} = e_q^2 C^{i}_{gq}\,,$$

$$C^{i,(2)}_{qg} = C^{i,(2)}_{\bar{q}g} = e_q^2 C^{i}_{qg}\,,$$

$$C^{i,(2)}_{gg} = \left(\sum_j e_{q_j}^2\right) C^{i}_{gg}\,.$$

Case distinctions in structure functions.



Published NNLO structure functions have **case distinction** across different regions in the SIDIS plane $0 < x, z < 1$:

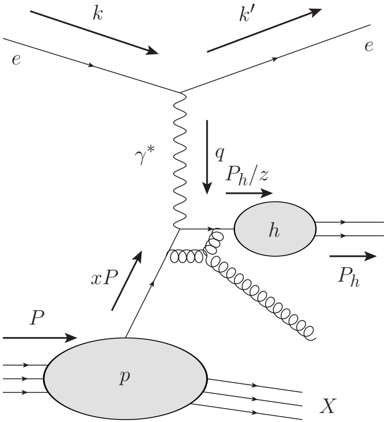
$$C_{ba}^{i,(2)} = C_{ba}^{i,(2),\boxtimes} + \sum_{i=1}^{4} U_j\, C_{ba}^{i,(2),U_j} + \sum_{i=1}^{2} R_j\, C_{ba}^{i,(2),R_j} + \sum_{j=1}^{2} T_j\, C_{ba}^{i,(2),T_j}$$

with $\boxtimes = \sum_{j=1}^{4} U_j$, $R_1 = U_1 + U_3$, $R_1 = U_2 + U_4$, $T_1 = U_3 + U_4$, $T_2 = U_1 + U_2$.

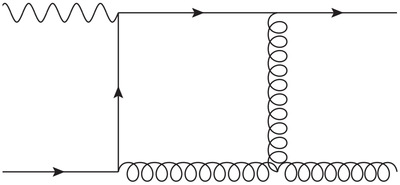
이 항은 도대체 어디서 튀어나온 거야?

Real-virtual branch-cuts.

Real-virtual correction (example):



Box contribution to $\mathcal{C}_{qq}^{i,(2)}$ (example):

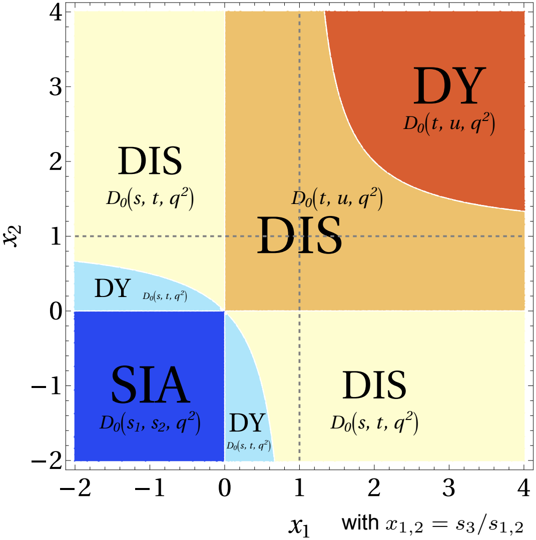


General analytic form of box integral:

$$D_0(s_1, s_2, q^2) \sim \left(\frac{\mu^2}{-s_2}\right)^{\varepsilon} \mathrm{F}_\varepsilon\left(\frac{s_3}{-s_1}\right) + \left(\frac{\mu^2}{-s_1}\right)^{\varepsilon} \mathrm{F}_\varepsilon\left(\frac{s_3}{-s_2}\right) - \left(\frac{\mu^2}{-q^2}\right)^{\varepsilon} \mathrm{F}_\varepsilon\left(\frac{-s_3 q^2}{s_1 s_2}\right)$$

with Mandelstam variables $s_{1,2,3} = s, t, u; \quad q^2 = \sum_i s_i$
and $\mathrm{F}_\varepsilon(z) = {}_2\mathrm{F}_1(1, -\varepsilon, 1-\varepsilon; z)$

Branch-cut structure of the one-loop box integral.



with $x_{1,2} = s_3/s_{1,2}$

- Contributions to D_0:

$$\underbrace{\left(\frac{\mu}{-s_i + \mathrm{i}0}\right)^{\varepsilon}}_{\text{branch-cut for } s_i<0} \qquad \underbrace{F_\varepsilon(x_i \pm \mathrm{i}0)}_{\text{b.-c. for } x_i>1}$$

branch-cut for $s_i<0$ → **physical threshold**

b.-c. for $x_i>1$ → **spurious threshold**

- BCs between 3 F_ε cancel
- replacement $F_\varepsilon(x) \to |x|^\varepsilon \mathfrak{F}_\varepsilon(x)$ removes spurious BCs.
- Single-valued version of F_ε:

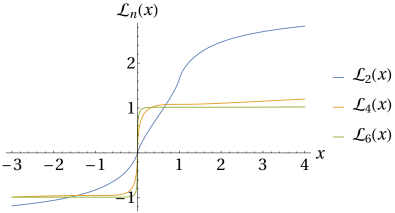
$$\mathfrak{F}_\varepsilon(x) = 1 - \ln\left|\frac{x}{x-1}\right| \sum_{n=1}^{\infty} \frac{\varepsilon^n \ln^{n-1}\left|\frac{1}{x}\right|}{n!} - \sum_{n=2}^{\infty} \varepsilon^n \mathcal{L}_n(x)$$

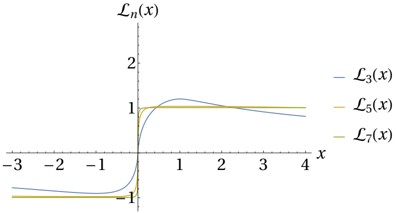
Intermezzo: Single-valued polylogarithms.

Definition:

$$\mathcal{L}_n(x) = \sum_{k=0}^{n-1} \frac{\ln^k\left|\frac{1}{x}\right|}{k!} \,\mathrm{Li}_{n-k}(x) + \frac{\ln^{n-1}\left|\frac{1}{x}\right|}{n!} \ln|1-x|\,.$$

- real-valued, continuous, and bounded on $\mathbb{R}$ (no branch-cut)
- satisfies “clean” versions of dilogarithmic functional equations, e.g.
  $\mathcal{L}_n(x) + (-1)^n \mathcal{L}\left(\frac{1}{x}\right) = \mathcal{L}_n(\mathrm{sgn}(x)\,\infty)$





Simplifying the coefficient functions.

- SIDIS NNLO coefficients contain dilogarithms with arguments
  $\frac{1-x}{z}\,,\;\frac{1-x}{1-z}\,,\;\frac{(1-x)x}{(1-z)z}\,,\;\frac{xz}{(1-x)(1-z)}\,,\;\frac{(1-z)x}{(1-x)z}\,,\dots$
- arguments become > 1 outside the $U_j \longrightarrow$ branch-cut

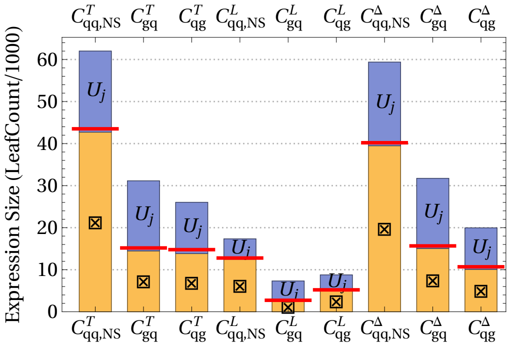
"spurious threshold"-logarithms

e.g. in region $x > z$: $(x-z)\log(x-z)$

- replace $\mathrm{Li}_2(x) \to \mathcal{L}_2(x) - \ln|x|\ln(1-x) + \frac{\ln|x|}{2}\ln|1-x|$ in each region
- use functional equations to map $\mathcal{L}_2(x)$ to universal form
- terms in all regions become the same, case distinctions can be removed
- **"naive" expectation:** size of U_j parts shrinks to 25%

Result is even better...

...because we can now sort the U_j into the ⊠ terms!



⊠ : original part w.o. case distinctions

U_j : original part with case distinctions

Red line : result after removal of case distinctions

Best case (distinction) scenario:

Extra size due to U_j virtually vanishes!

What do we gain from removing the spurious thresholds?

Shorter, more compact expressions:

- numerically faster and more stable close to spurious branch points $\to$ useful for PDF fits and event generators
- analytically simpler $\to$ useful for further processing e.g. calculating analytic Mellin transform

Upon closer inspection of numerics

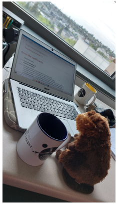
Numerical evaluation of coefficient functions bottle-necked by special function calls rather than expression length

What do we gain from removing the spurious thresholds?

Shorter, more compact expressions:

- numerically faster and more stable close to spurious branch points $\rightarrow$ useful for PDF fits and event generators
- analytically simpler $\rightarrow$ useful for further processing e.g. calculating analytic Mellin transform

$\rightarrow$ introduced C++ special functions library `BEAVER` (Better Evaluate A Very Efficient Rational) for a speed-up by a factor of $\sim$**5**



Upon closer inspection of numerics

Numerical evaluation of coefficient functions bottle-necked by special function calls rather than expression length

Function	Implementation	time/call	speed-up
$\ln(x)$	`std::log`	16.0 ns	1
$\ln(x)$	`vdt::fast_log`	6.7 ns	2.3
$\ln(x)$	`beaver::log`	5.0 ns	**3.2**
$\arctan(x)$	`std::atan`	18.2 ns	1
$\arctan(x)$	`vdt::fast_atan`	9.2 ns	2.0
$\arctan(x)$	`beaver::atan`	8.8 ns	**2.1**
$\mathrm{Li}_2(x)$	`CERNLIB::dilog`	124.4 ns	1
$\mathrm{Li}_2(x)$	`voigt::dilog`	43.6 ns	2.9
$\mathrm{Li}_2(x)$	`beaver::dilog`	11.3 ns	**11.2**
$\mathcal{L}_2(x)$	`beaver::svdilog`	14.1 ns	—
$\mathrm{Li}_3(x)$	`CERNLIB::trilog`	132.1 ns	1
$\mathrm{Li}_3(x)$	`voigt::trilog`	28.2 ns	4.7
$\mathrm{Li}_3(x)$	`beaver::trilog`	13.7 ns	**9.6**
$\mathrm{Ti}_2(x)$	`gsl_sf_atanint`	80.0 ns	1
$\mathrm{Ti}_2(x)$	`beaver::atanint`	11.9 ns	**6.7**

Analytic benefits: Towards analytic Mellin transform.

> Double Mellin-Transform of coefficient function

$$\tilde{\mathcal{C}}(N,M) = \int_0^1 \mathrm{d}x\, x^{N-1} \int_0^1 \mathrm{d}z\, z^{M-1} \mathcal{C}(x,z)\,.$$

> Example: Coefficient of $\mathcal{L}_2\left(\frac{1-x}{1-z}\right)$

$$\tilde{\mathcal{C}}_{gq}^{\mathsf{T}}\Big|_{\mathcal{L}_2}(N,M) = \int_0^1 \mathrm{d}x\, x^{N-1} \int_0^1 \mathrm{d}z\, z^{M-1} f(x,z) \mathcal{L}_2\left(\frac{1-x}{1-z}\right)$$

with

$$\begin{aligned} f(x,z) =& N_{\mathsf{C}}^2 \left[xz + \frac{x}{2z} - \frac{z}{2(1-x)} - \frac{1}{(1-x)z} - x + \frac{1}{1-x} + \frac{1}{2z} - 1 \right] \\ &+ \frac{1}{1-x} - \frac{z}{2(1-x)} - \frac{1}{N_{\mathsf{C}}^2} \left[xz + \frac{x}{2z} - \frac{z}{1-x} - \frac{1}{(1-x)z} - x \right. \\ &\left. + \frac{2}{1-x} + \frac{1}{2z} - 1 \right]. \end{aligned}$$

Towards analytic Mellin transform.

- Split integral in $0 < x < z$ and $z < x < 1$, use functional equation in the first region to map argument on unit interval, and change variables:

$$
\begin{aligned}
\tilde{\mathcal{C}}^{\mathsf{T}}_{gq}\Big|_{\mathcal{L}_2}(N,M) &= \int_0^1 \mathrm{d}z\, z^{M-1} \\
&\times \Bigg\{ \int_0^z \mathrm{d}t\, \frac{1-z}{(1-t)^2} \left(\frac{z-t}{1-t}\right)^{N-1} f\left(\frac{z-t}{1-t}, z\right) [\mathcal{L}_2(t) + \zeta_2] \\
&+ (1-z) \int_0^1 \mathrm{d}t\, (1-t-zt)^{N-1} f(1-t-zt, z) \mathcal{L}_2(t) \Bigg\}\,.
\end{aligned}
$$

- Replace $\mathcal{L}_2(t) = -G(0,1;t) + \frac{1}{2} G(0;t) G(1;t)$ (valid for $0 < t < 1$) and use `PolyLogTools` (for fixed integer N, M)

$$\tilde{\mathcal{C}}^{\mathsf{T}}_{gq}\Big|_{\mathcal{L}_2}(1,2) = N_{\mathsf{C}}^2\left(-\zeta_3 - \frac{1}{32} - \frac{31\pi^2}{240}\right) + \frac{1}{N_{\mathsf{C}}^2}\left(\frac{\zeta_3}{2} + \frac{1}{96} + \frac{13\pi^2}{720}\right) + \frac{\zeta_3}{2} + \frac{\pi^2}{9} + \frac{1}{48}\,,$$

$$\tilde{\mathcal{C}}^{\mathsf{T}}_{gq}\Big|_{\mathcal{L}_2}(1,3) = N_{\mathsf{C}}^2\left(-\frac{7\zeta_3}{16} - \frac{719}{13824} - \frac{611\pi^2}{5760}\right) + \frac{1}{N_{\mathsf{C}}^2}\left(\frac{\zeta_3}{8} + \frac{97}{6912} + \frac{53\pi^2}{2880}\right) + \frac{5\zeta_3}{16} + \frac{101\pi^2}{1152} + \frac{175}{4608}\,,$$

$$\tilde{\mathcal{C}}^{\mathsf{T}}_{gq}\Big|_{\mathcal{L}_2}(2,2) = N_{\mathsf{C}}^2\left(-\zeta_3 - \frac{1}{32} - \frac{59\pi^2}{2160}\right) + \frac{1}{N_{\mathsf{C}}^2}\left(\frac{\zeta_3}{2} + \frac{5}{192} - \frac{23\pi^2}{1080}\right) + \frac{\zeta_3}{2} + \frac{7\pi^2}{144} + \frac{1}{192}\,,$$

$$\tilde{\mathcal{C}}^{\mathsf{T}}_{gq}\Big|_{\mathcal{L}_2}(1,4) = N_{\mathsf{C}}^2\left(-\frac{11\zeta_3}{40} - \frac{2077}{34560} - \frac{2953\pi^2}{33600}\right) + \frac{1}{N_{\mathsf{C}}^2}\left(\frac{\zeta_3}{20} + \frac{17}{1080} + \frac{31\pi^2}{2100}\right) + \frac{9\zeta_3}{40} + \frac{117\pi^2}{1600} + \frac{511}{11520}\,,$$

$$\tilde{\mathcal{C}}^{\mathsf{T}}_{gq}\Big|_{\mathcal{L}_2}(2,3) = N_{\mathsf{C}}^2\left(-\frac{7\zeta_3}{16} - \frac{271}{13824} - \frac{5543\pi^2}{120960}\right) + \frac{1}{N_{\mathsf{C}}^2}\left(\frac{\zeta_3}{8} + \frac{47}{6912} + \frac{31\pi^2}{60480}\right) + \frac{5\zeta_3}{16} + \frac{29\pi^2}{640} + \frac{59}{4608}\,,$$

$$\tilde{\mathcal{C}}^{\mathsf{T}}_{gq}\Big|_{\mathcal{L}_2}(3,2) = N_{\mathsf{C}}^2\left(-\zeta_3 - \frac{55}{864} + \frac{\pi^2}{70}\right) + \frac{1}{N_{\mathsf{C}}^2}\left(\frac{\zeta_3}{2} + \frac{23}{432} - \frac{23\pi^2}{630}\right) + \frac{\zeta_3}{2} + \frac{\pi^2}{45} + \frac{1}{96}\,,$$

$$\vdots$$

$$\begin{aligned}\tilde{\mathcal{C}}^{\mathsf{T}}_{gq}\Big|_{\mathcal{L}_2}(25,25) = {}& N_{\mathsf{C}}^2\left(-\frac{163\zeta_3}{5200} - \frac{19160303432263921718814110\,8951}{821911803139027880982743040000\,00} - \frac{3437040783761160473972 39\pi^2}{14503528279871264585952 0000}\right)\\ &+ \frac{1}{N_{\mathsf{C}}^2}\left(\frac{\zeta_3}{5200} + \frac{26265010381314511533438 76273}{821911803139027880982743040000\,00} - \frac{116200904407023002701 9\pi^2}{14503528279871264585952 0000}\right)\\ &+ \frac{81\zeta_3}{2600} + \frac{1637540775974293814 93\pi^2}{6886765564991103792 0000} + \frac{68618930023423299213797107}{2984429205297849967257600 0000}\,.\end{aligned}$$

Conclusion and Outlook.

- understood the origin of spurious case distinctions of existing NNLO results for unpol.+pol. SIDIS in terms of spurious thresholds in real-virtual corrections
- significant compactification (by 30-60%) interesting for use within a fit (e.g. polarized PDFs at NNLO) or event generators (faster & more stable evaluation)
- facilitates further analytic manipulation, e.g. analytic Mellin transform
- presented idea directly applicable to all RV processes that contain box diagrams (e.g. DY, SIA, also beyond QCD)
- potentially interesting to explore if similar single-valued structure can be identified in multi-loop integrals with intricate branch-cut structure

Thank you!

Contact

Deutsches Elektronen-
Synchrotron DESY

https://theory-hamburg.desy.de/

Fabian Wunder
0009–0007–4136–7844
Theory Group
fabian.wunder@desy.de
+49–40–8998–4235