Heavy-quark initiated deep-inelastic scattering

First NNLO heavy-quark initiated coefficient functions for DIS

Kirill Kudashkin

as a former member of

N3PDF
Machine Learning • PDFs • QCD

and in close collaboration with

NNPDF

Shandong University[1], Università degli Studi di Milano Statale[2]

Based on arXiv:2601.02916, KK (accepted for publication in JHEP)

Quarks-2026, Petrozavodsk, Russia, May 18, 2026

[1]This work is supported by the National Science Foundation of China under grants No. 12275156 and No. 12321005.

[2]This work was supported in part by the European Union's Horizon 2020 Programme (grant agreement n.740006).

Outline

Paradigm Shift
(Transition Era)

Discovery Era $\longrightarrow$ Precision Era

Discovery Era

LEP Tevatron LHC

Maximize energy reach.

New kinematic regimes to find new particles (Top, Higgs, etc).

Precision Era

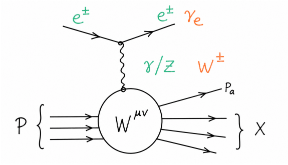
HL-LHC EIC LHeC FCC

Maximize luminosity/statistics $\rightarrow$ 1% theoretical uncertainties.

Scrutinize known sectors and complementary ones.

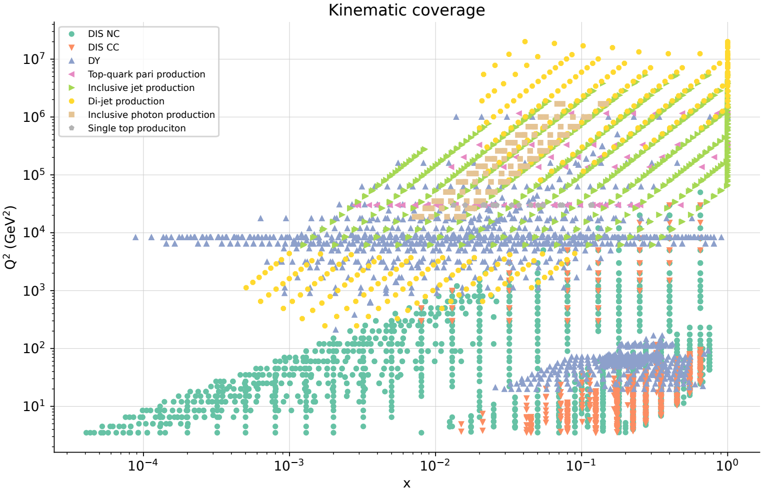
Deep inelastic scattering

Neutral/Charged current DIS



$W^{\mu\nu}$ is the hadronic tensor

- DIS dominates PDF constraints
- Broad kinematic coverage



Nearly 60% of data points ($\mathcal{O}(5000)$) are DIS events. [Roy Stegeman, 2022]

CC DIS: Experimental avenues

New experiments

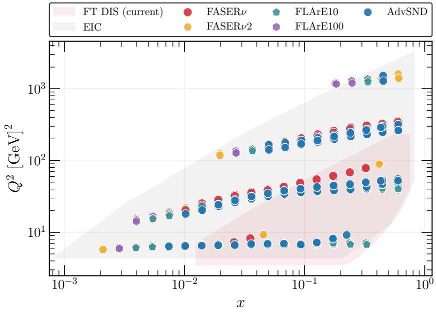
Detector	N_{ν_e}	$N_{\bar{\nu}_e}$	$N_{\nu_e} + N_{\bar{\nu}_e}$	N_{ν_μ}	$N_{\bar{\nu}_\mu}$	$N_{\nu_\mu} + N_{\bar{\nu}_\mu}$
FASERν	400 (62)	210 (38)	610 (100)	1.3k (200)	500 (90)	1.8k (290)
SND@LHC	180 (22)	76 (11)	260 (32)	510 (59)	190 (25)	700 (83)
FASERν2	116k (17k)	56k (9.9k)	170k (27k)	380k (53k)	133k (23k)	510k (76k)
AdvSND-far	12k (1.5k)	5.5k (0.82k)	18k (2.3k)	40k (4.8k)	16k (2.2k)	56k (7k)
FLArE10	44k (5.5k)	20k (3.0k)	64k (8.5k)	76k (10k)	38k (5.0k)	110k (15k)
FLArE100	290k (35k)	130k (19k)	420k (54k)	440k (60k)	232k (30k)	670k (90k)

Integrated event yields for the selected current and future experiments. [Juan M. Cruz-Martinez+,2024]

A bright future

- Larger luminosity: 10-100 × HERA
- Larger kinematic coverage: 2 × current FT DIS
- Higher statistics: $\mathcal{O}(10^6)$ new events

Electron-Ion Collider



The kinematic coverage in the $(x,\ Q^2)$ plane of muon-neutrino scattering at present and future collider experiments.
[Juan M. Cruz-Martinez+,2024]

New theory challenges

Higher luminosities opened new frontiers

Honorable mentions:

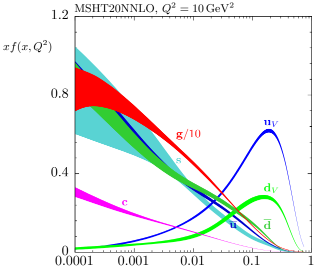
SIDIS ▪ nuclear PDFs ▪ precise W mass ▪ ...

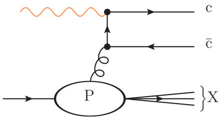
Personal interest:

Heavy-quark mass $\rightarrow$ Intrinsic charm

Extrinsic vs Intrinsic heavy flavor

Extrinsic

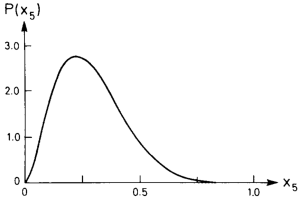


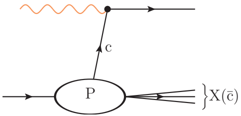


- Perturbative (origin)
- Dominant at small x
- Symmetric $c/\bar{c}$ & Sea
- Known up to N3LO fix-order QCD

[Witten, 1975]

Intrinsic



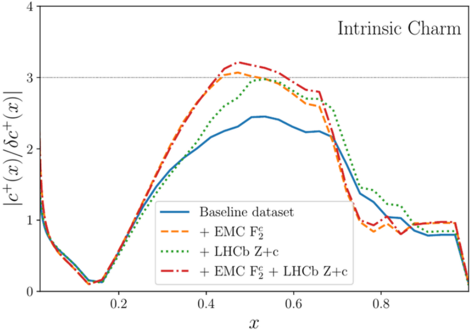


- Non-perturbative: $|\mathrm{P}\rangle = \mathcal{P}_{3q}\,|\mathrm{uud}\rangle + \mathcal{P}_{5q}\,|\mathrm{uudc\bar{c}}\rangle$
- Dominant at moderate $x > 0.1$
- Symmetric/asymmetric & Valence
- Known up to NLO fix-order QCD

[BHPS, 1980]

Status of intrinsic charm

Evidence for intrinsic charm quarks in the proton
[Nature 608 (2022) 7923 483, NNPDF, KK]



Intrinsic charm in QCD

- No theoretical (formal) problems
- PDFs are extracted from data (A global fit)
- Evident that the intrinsic $c^+ \neq 0$
  - ⋆ 2.5σ for baseline
  - ⋆ 3.0σ with LHCb Z+c and/or EMC F_2^c

A global fit depend on

✓ Experimental data

– Fitting methodology

⋆ Fixed order calculations

⋆ Resummations

From Evidence to Discovery!

Theoretical uncertainty $\delta c^+(x)$ has to be reduced!

From Evidence to Precision

Fixed-order QCD calculations for DIS:

Intrinsic channel known to $\mathcal{O}(\alpha_s)$ vs. Extrinsic channel known to $\mathcal{O}(\alpha_s^3)$

- Global PDF fits increasingly favor non-zero intrinsic charm
- Missing higher-order corrections in the intrinsic channel may bias future N^3LO global fits

This work addresses this gap by providing $\mathcal{O}(\alpha_s^2)$ corrections to HQI charged-current DIS.

Goals

✓ Compute $\omega_n^{(2),\mu\nu}$ for HQI CC DIS

⋆ Enable future resummation of large logarithms $\ln(Q^2/m^2)$

Factorization Theorem

DIS observables factorize into [Collins, Soper, Sterman, 1989],[Collins, 1998]:

$$W^{\mu\nu}(x_B, Q^2) = \sum_n \int_{x_B}^{1} \frac{d\xi}{\xi}\, \omega_n^{\mu\nu}\left(\frac{x_B}{\xi}, Q^2, m^2, \mu_F^2\right) f_n(\xi, \mu_F^2)$$

Perturbative part

Wilson coefficients

$$\omega_n^{\mu\nu} = \delta(1-\xi) + \sum_{k=1}^{\infty} \left(\frac{\alpha_s}{\pi}\right)^k \omega_n^{(k),\mu\nu}$$

Computed in perturbative QCD

Non-perturbative part

Parton distribution functions

$$f_n(\xi, \mu_F^2)$$

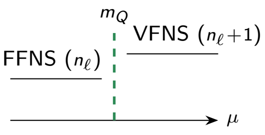
Universal proton structure

Goal of this work: compute $\omega_n^{(2),\mu\nu}$ for heavy-quark initiated CC DIS.

Variable Flavor Number Schemes (GM-VFNS)

Fixed vs. Variable Flavor

- The presence of an extra scale (within **FFNS**) results in large logarithms $\ln(m^2/Q^2)$.
- These are typically resummed via DGLAP-like approach $\rightarrow$ VFNS
- **FFNS**: n_f fixed, heavy quark $\approx$ massive particle in hard cross section.
- **VFNS**: n_f grows with scale; heavy quark becomes a *parton* above $\mu \sim m_Q$.

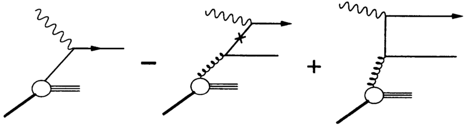


GM-VFNS Implementations

- **ACOT** (Aivazis-Collins-Olness-Tung)
  massive coefficient functions, subtraction terms, no massless limit.
- **FONLL** (Forte-Nason-Ridolfi)
  interpolates FFNS + massless VFNS, matches at threshold.

GM-VFNS = full mass dependence retained
both ACOT and FONLL are widely used.

This work: 3FNS scheme for CC DIS relevant for both ACOT and FONLL



Matching coefficient (example) [ACOT, 1991]:

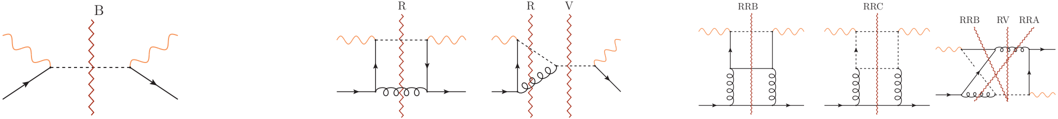
$$f_g^{Q,(1)} = \frac{\alpha_s}{\pi} \ln \frac{\mu^2}{m^2} P_{gQ}$$

Computational Toolchain

The computational pipeline from Feynman diagrams to NNLO results.

Computational Toolchain

- **QGRAF** generates all forward scattering Feynman diagrams for the partonic processes



- **Private code** handles many things automatically [Pak, 2012], [Hoff, 2015]
  1. Topology identification
  2. Cut classification via graph-coloring (needed for the optical theorem $\omega^{\mu\nu} = 2\,\text{Im}\,\mathcal{M}^{\mu\nu}$)

	$\mathcal{O}(\alpha_s^0)$	$\mathcal{O}(\alpha_s^1)$		$\mathcal{O}(\alpha_s^2)$				
	B	V	R	VV	RV	RRA	RRB	RRC
$\mathrm{X_{QCD}}$	q	q	gq	q	gq	ggq, $\tilde{\mathrm{g}}\tilde{\mathrm{g}}$q, $\bar{\mathrm{q}}$qq	$\bar{\mathrm{Q}}$Qq	Q$\bar{\mathrm{q}}$q
N	1	5		106				

  3. Sector mapping (next slides)

- **Preprocessing**: 106 out of 360 generic (forward scattering) Feynman diagrams (imposing charge symmetry + cuts).

Computational Toolchain

Algebraic reduction

- Dirac traces and color algebra performed with `FORM`
- Projection onto hadronic tensor structures
- Result expressed as scalar loop integrals

Partonic tensor

$$\omega^{\mu\nu} = -\omega_1 g^{\mu\nu} + \omega_2 p^\mu p^\nu + i\omega_3 \epsilon^{\mu\nu\alpha\beta} q_\alpha p_\beta + \cdots$$

$$\omega_i = \sum_j R_{ij}\, I_j$$

where R_{ij} are rational functions and I_j are scalar integrals.

Calculation complexity

- Two-loop, two-scale problem
- $\mathcal{O}(1000)$ scalar integrals
- 6 generic integral families
- Integrals regulated via the dimensional regularization $d = 4 - 2\epsilon$
- Traces with γ_5

Scalar Feynman integral

$$I = (\mu^2)^{2\epsilon} \int \frac{\mathrm{d}^d l_1}{(2\pi)^d} \frac{\mathrm{d}^d l_2}{(2\pi)^d} \frac{\mathcal{N}(l_1, l_2, p_1, p_2)}{D_1^{a_1} D_2^{a_2} D_3^{a_3} D_4^{a_4} D_5^{a_5} D_6^{a_6} D_7^{a_7}},$$

The output of this stage is a structured system of scalar integrals ready for IBP reduction.

Computational Toolchain

QGRAF → FORM → REDUZE2 → Libra → AMFlow, PolyLogTools

The full amplitude is reduced to a finite basis of master integrals.

Integration-by-parts identities

Loop integrals satisfy identities of the form [Chetyrkin, Tkachov, 1981]

$$0 = \int \prod_{i=1}^{L} d^d k_i \, \frac{\partial}{\partial k_j^\mu} \left(v^\mu \prod_{r=1}^{n} \frac{1}{D_r^{a_r}} \right)$$

which generate large linear systems relating different integrals.

$$I_k = \sum_{l=1}^{111} C_{kl} \, j_l$$

IBP reduction performed with REDUZE2.

Complexity reduction

$\mathcal{O}(1000)$
scalar integrals

⇓

111
master integrals

organized into

6 integral families

Computational Toolchain

Cut equations ⇒ Smaller systems ⇒ Simpler DEQs

Differential equations method

Master integrals satisfy systems of differential equations

$$\frac{\partial}{\partial y_i} j = M_i(y_1, y_2, \epsilon)\, j,$$

which are transformed to canonical form

$$\frac{\partial}{\partial y_i} J = \epsilon\, A_i(y_1, y_2)\, J.$$

Canonical differential equations allow iterative solution in ϵ.

[Kotikov, 1991; Remiddi, 1997; Henn, 2013]

Cut-based strategy

The optical theorem separates the problem into different cut sectors:

$$2\,\mathrm{Im}\, j_k = \theta(y_1 - 4y_2)\, j_k^{\mathrm{RRB}} + \theta(y_1 - y_2)\, j_k^{\mathrm{RRC}} + \theta(y_1)\, \left(j_k^{\mathrm{RRA}} + j_k^{\mathrm{RV}}\right)$$

Adjacency matrices are used to:

- identify cuts,
- construct cut-specific systems,
- reduce matrix size.

Each cut sector is transformed independently into canonical Fuchsian form with `Libra`.[3]

$$j^k = \mathrm{T}^k \cdot J^k, \quad \textit{with} \quad k = \mathrm{RRA},\ \mathrm{RRB},\ \mathrm{RRC}, \ldots$$

[3] I thank Roman N. Lee for assistance with his package [Lee, 2020].

Computational Toolchain

QGRAF → FORM → REDUZE2 → Libra → AMFlow, PolyLogTools

Iterative solution

The canonical system is solved through path-ordered exponentiation

$$J = P \exp \left[\epsilon \int dy \cdot A \right] .$$

Expanding in ϵ generates iterated integrals, which are evaluated recursively via Picard iteration.

Functional structure

The solutions are expressed in terms of Chen iterated integrals and then Goncharov polylogarithms [Goncharov, 2001]

$$G(a_1, \ldots, a_n; y) = \int_0^y \frac{dt}{t - a_1} G(a_2, \ldots, a_n; t).$$

Algebraic letters arise through

$$u = \sqrt{\frac{y_1}{y_1 - 4y_2}}.$$

Boundary conditions

Integration constants are determined from:

- high-precision numerical evaluations,
- AMFlow,
- PSLQ reconstruction.

Typical constants:

$$\pi^n, \quad \zeta_n, \quad \log^n 2, \quad \mathrm{Li}_n(1/2).$$

[Liu, 2022]

Final result

All boundary constants fixed

$$\Downarrow$$

Bare analytic NNLO coefficient functions

Renormalization

Strong coupling renormalization

Massive quark loops modify the gluon self-energy and violate on-shell conditions for external massless states.

We therefore employ the decoupling scheme

$$\Pi^{\mu\nu}(0, m^2) = 0,$$

leading to

$$a_0 = a(\mu^2)\left[1 + a(\mu^2)\left(-\frac{\beta_0}{2\epsilon} - n_H f(\epsilon)\frac{\beta_{0,Q}}{2\epsilon}\left(\frac{m^2}{\mu^2}\right)^{-\epsilon}\right)\right].$$

This setup is particularly relevant for GM-VFNS matching.

Wave-function and mass renormalization

Quark fields and masses are renormalized on-shell:

$$\psi_0 = Z_2^{1/2}\psi, \qquad m_0 = Z_m\, m.$$

[Melnikov+, 2000]

Beta-function coefficients

$$\beta_0 = \frac{11}{3}C_A - \frac{4}{3}T_R n_L,$$

$$\beta_{0,Q} = -\frac{4}{3}T_R.$$

$SU(N_c)$ color factors:

$$C_F = \frac{N_c^2 - 1}{2N_c}, \qquad C_A = N_c, \qquad T_R = \frac{1}{2}.$$

For QCD:

$$N_c = 3.$$

Physical picture

Heavy-quark loops decouple below threshold

Infrared Regularization

Massless final states

Soft and collinear singularities are regulated using plus-distributions:

$$y^{-1-b\epsilon} = -\frac{\delta(y)}{b\epsilon} + \sum_{k=0}^{\infty} \frac{(-b\epsilon)^k}{k!} D_k(y),$$

where

$$\int_0^1 g(y) D_k(y) = \int_0^1 g(y) \left[\frac{\log^k y}{y}\right]_+ = \int_0^1 \log^k y \left(\frac{g(y)-g(0)}{y}\right).$$

These distributions isolate endpoint singularities in phase-space integrals.

Massive final states

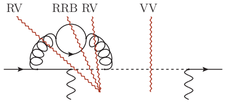
For massive quarks, the phase-space threshold shifts:

$$\int_{4y_2}^1 dy_1\, g(y_1)\, y_1^{-1-b\epsilon}.$$

The expansion generates Sudakov logarithms

$$\log^3\left(\frac{m^2}{Q^2}\right).$$

Sudakov structure



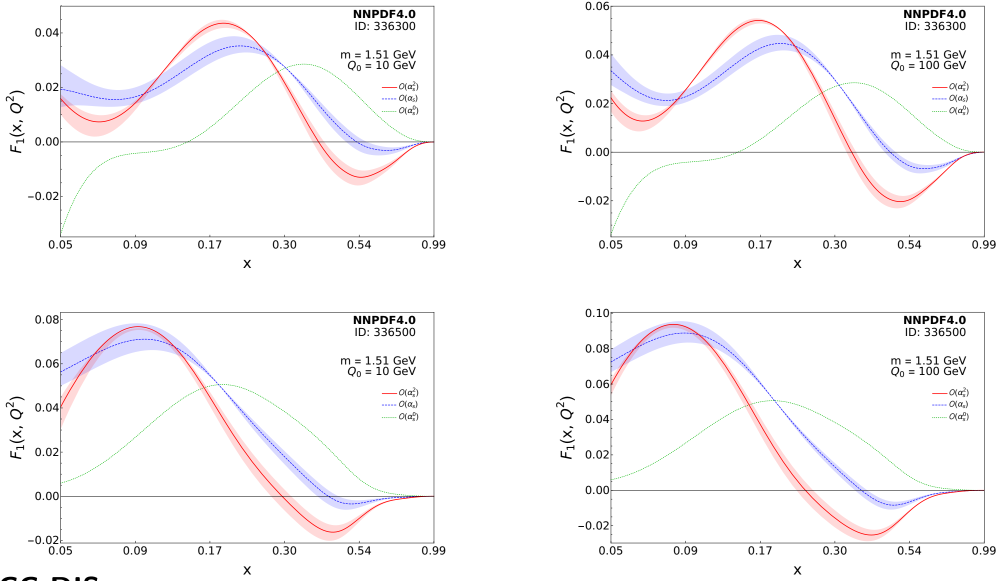
Different cut contributions generate Sudakov logarithms which cancel in physical observables when all the contributions are combined.

Initial state divergences

Initial state collinear divergences are fully regulated by quark mass. No collinear subtraction is required.

The μ_F factorization scale is not generated.

A preliminary numerical study



Summary & Outlook

Summary

- First complete massive NNLO coefficient functions for heavy-quark initiated charged-current DIS
- Analytic results valid for arbitrary

$$m_Q^2/Q^2$$

- Differential equations solved in canonical form
- Boundary constants reconstructed using AMFlow + PSLQ
- Independent checks:
  - massless limits,
  - agreement with known vertex functions,
  - consistency between two γ_5 schemes
- Auxiliary Mathematica files provided for phenomenological applications

Physics impact

- Reduces theoretical uncertainty in intrinsic-charm studies
- Essential ingredient for future GM-VFNS implementations
- Enables consistent NNLO analyses of intrinsic charm in global PDF fits

Outlook

- Full GM-VFNS matching
- Extension to remaining NNLO DIS channels

Based on arXiv:2601.02916 (accepted for publication in JHEP)

An invitation for experts

Extension to remaining NNLO DIS channels

Other channels are involved due to elliptic functions.

1. Regular basis [Lee, Onishchenko, 2019]
2. Methods developed by ϵ-collaboration [Bree+, 2025]
3. Many other methods

Thank you!

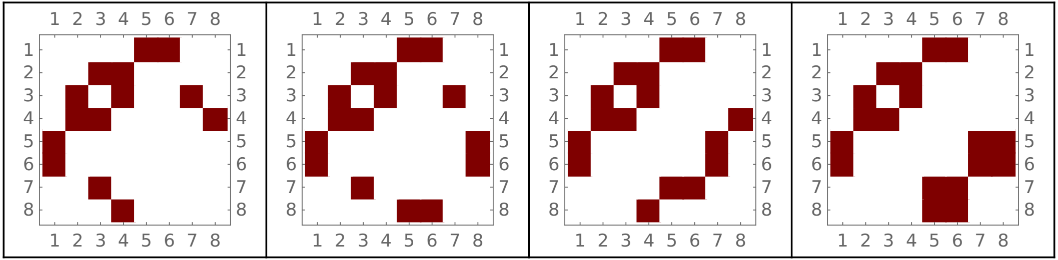
Backup

Adjacency Matrix

An example of an adjacency matrix

$$\begin{pmatrix}
0 & 0 & 0 & 0 & W_M(p_1) & q_1(p_2) & 0 & 0 \\
0 & 0 & W_M(p_1) & q_1(p_2) & 0 & 0 & 0 & 0 \\
0 & W_M(p_1) & 0 & q_1(L_1) & 0 & 0 & q_2(L_1+p_1) & 0 \\
0 & q_1(p_2) & q_1(L_1) & 0 & 0 & 0 & 0 & g(-L_1+p_2) \\
W_M(p_1) & 0 & 0 & 0 & 0 & 0 & q_2(L_2) & q_1(L_2-p_1) \\
q_1(p_2) & 0 & 0 & 0 & 0 & 0 & g(-L_1+L_2-p_1) & \begin{matrix} q_1(-L_1+L_2 \\ -p_1+p_2) \end{matrix} \\
0 & 0 & q_2(L_1+p_1) & 0 & q_2(L_2) & g(-L_1+L_2-p_1) & 0 & 0 \\
0 & 0 & 0 & g(-L_1+p_2) & q_1(L_2-p_1) & \begin{matrix} q_1(-L_1+L_2 \\ -p_1+p_2) \end{matrix} & 0 & 0
\end{pmatrix}$$

[Pak, 2012], [Grigo, Hoff, 2014]



Integral Families

Family	Propagators $D_1, \ldots, D_7$
fam1	$(l_2+p_1)^2-m^2,\ l_2^2,\ (l_1-l_2-p_1+p_2)^2-m^2,$ $(l_1-l_2-p_1)^2,\ l_1^2-m^2,\ (l_1+p_2)^2,\ (l_1-p_1)^2$
fam2	$(l_1+l_2)^2,\ l_1^2-m^2,\ (l_1-p_1)^2,\ (l_1+p_2)^2,$ $l_2^2-m^2,\ (l_2+p_1)^2,\ (l_2-p_2)^2$
fam3	$(l_1+l_2-p_1+p_2)^2,\ l_1^2-m^2,\ (l_1-p_1)^2,$ $(l_1+p_2)^2,\ l_2^2-m^2,\ (l_2+p_2)^2,\ (l_2-p_1)^2$
fam4	$(l_2+p_1)^2-m^2,\ l_2^2,\ (l_1-l_2-p_1-p_2)^2-m^2,$ $(l_1-l_2-p_1)^2,\ l_1^2-m^2,\ (l_1-p_2)^2,\ (l_1-p_1)^2$
fam5	$(l_1+l_2-p_1-p_2)^2,\ l_1^2-m^2,\ (l_1-p_1)^2,$ $(l_1-p_2)^2,\ l_2^2-m^2,\ (l_2-p_2)^2,\ (l_2-p_1)^2$
fam6	$l_1^2-m^2,\ l_2^2,\ (l_2-p_1)^2,\ (l_1-p_2)^2,$ $(l_1-l_2-p_2)^2,\ (l_1-l_2+p_1-p_2)^2,\ (l_1-l_2)^2$