An Analytic Regression Path to Precision QCD

Xiaoyuan Zhang
MIT

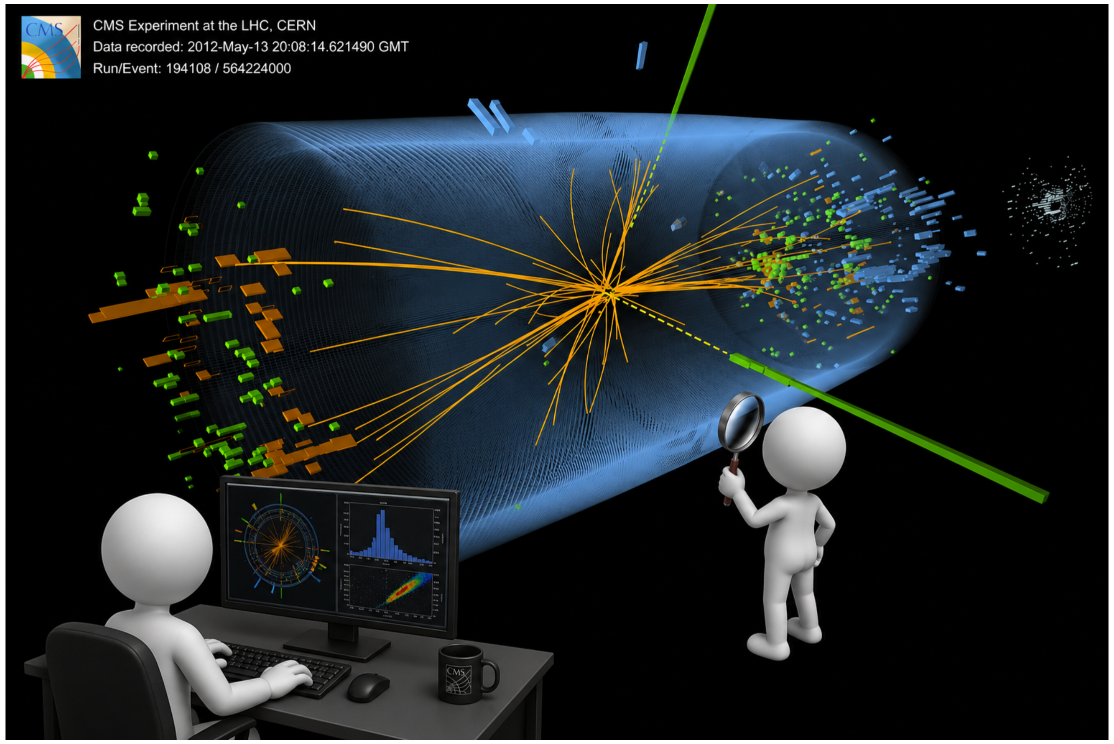
LoopFest XXIV

Based on:

- [2507.17815]: with Oscar Barrera, Aurélien Dersy, Rabia Husain, Matthew Schwartz [Harvard]
- [2509.22782]: with Jianyu Gong, Kai Yan [Shanghai JiaoTong], Andrzej Pokraka [Amsterdam]

Precision QCD

- One of the main tasks in collider physics:



Precision measurements

- Test if a theory is correct
- Determine fundamental parameters
- Search for new physics discovery

[Collins, Soper, Sterman, 1985]

- Calculations are based on the non-perturbative factorization theorem

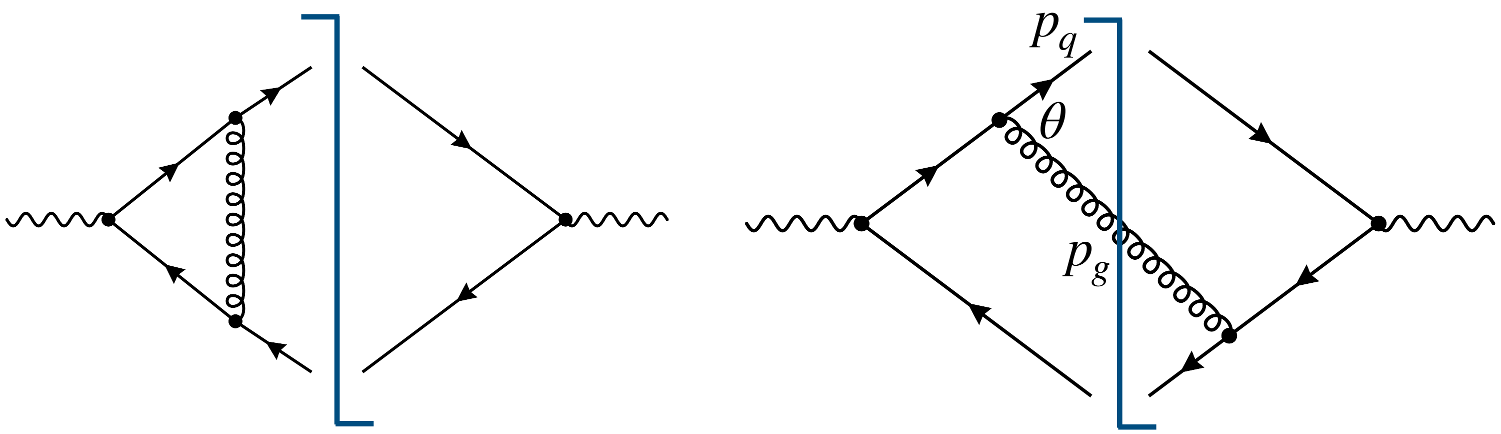
$$d\sigma = \underbrace{f_{P\to a}(x_1) \times f_{P\to b}(x_2)}_{\text{Parton distribution function}} \otimes \underbrace{d\hat{\sigma}}_{\text{Perturbative}} \otimes \underbrace{F}_{\text{Fragmentation}}$$

Parton distribution, Fragmentation: non-perturbative, extracted from data

$d\hat{\sigma}$:Feynman diagrams

Feynman integrals are hard

1. Textbook: NLO total cross-section for $e^+e^- \to \mathrm{hadrons}$



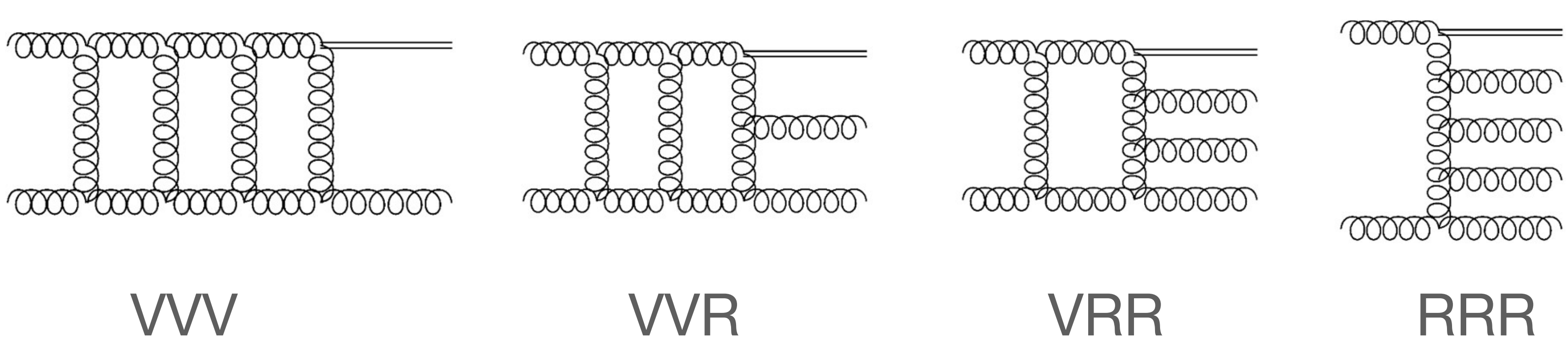
loop integral ⇒ UV divergence

Phase space ⇒ IR divergence

$$d\sigma \sim \alpha_s \frac{d\theta}{\theta}\frac{dE}{E}$$

Although IR divergence cancelled at the observable level, calculations needs to be done with regulator ($d = 4 - 2\epsilon$)

2. One example: Higgs + one jet production at three-loop



[Xiang, Guan, Mistlberger, 2504.06490]

[Badger et al., 2412.06519]

$\mathcal{O}(10^4)$ Feynman diagrams, $\mathcal{O}(10^7)$ integrals, complicated color structure

Integration-by-part (IBP) approach: $\mathcal{O}(10^6)$ CPU-hours, $\mathcal{O}(10^2 - 10^3)$ GB memory

Master integrals: $\mathcal{O}(10^3)$ master integrals that need to compute explicitly

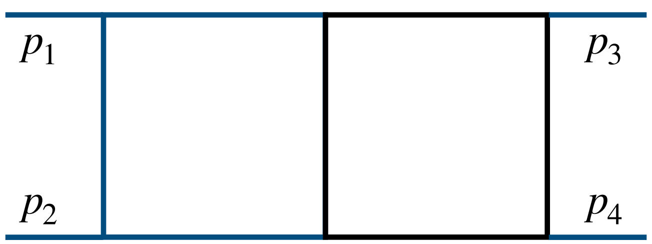
Feynman integrals are also simple

- Total cross-section are in terms of ζ_n value

Transcendental numbers

$$\hat{\sigma}_{e^+e^-} = \sigma_0 \left[1 + \frac{\alpha_s}{4\pi}(3C_F) + \left(\frac{\alpha_s}{4\pi}\right)^2 \left(C_F C_A(132/2 - 44\zeta_3) + C_F T_F n_f(-22 + 16\zeta_3) - C_F^2(3/2)\right) + \cdots\right]$$

- More generally, gives rise to polylogarithmic functions in $\epsilon \to 0$ expansion



$x = t/s$

$$\sim -\frac{4}{\epsilon^4} + \frac{5\log x}{\epsilon^3} - \left(2\log^2 x - \frac{5}{2}\pi^2\right)\frac{1}{\epsilon^2} - \left(\frac{2}{3}\log^3 x + \frac{11}{2}\pi^2\log x - \frac{65}{3}\zeta_3 - 4\mathrm{Li}_3(-x) - 4\log x\,\mathrm{Li}_2(-x) - 2(\log^2 x + \pi^2)\log(1+x)\right)\frac{1}{\epsilon} + \mathcal{O}(\epsilon^0)$$

Classical Polylogarithms $\quad \log(z), \quad \mathrm{Li}_2(z) = -\int_0^z \frac{\log(1-u)}{u} d\,u$

Generalized Polylogarithms (GPLs) $\quad G(a_1, \cdots a_n; z) = \int_0^z \frac{dt}{t - a_1} G(a_2, \cdots a_n; t)$

$$G(a; z) \equiv \log(z - a), \quad G(\vec{0}; z) \equiv \frac{1}{n!}\log^n(z)$$

Elliptic Polylogarithms (eMPLs) + more complicated structure

Transcendental functions

Transcendental functions and Symbol

- Transcendental weight is defined such that differentiation lowers the weight by one
  - weight-0: rational numbers or functions
  - weight-1: $\pi, \log(z), \cdots$
  - weight-2: $\zeta_2, \mathrm{Li}_2(z), \cdots$
  - weight-n: $\zeta_n, \mathrm{Li}_n(z), G(a_1, \cdots, a_n; z), \cdots$

$$\partial_z \mathrm{Li}_n(z) = \frac{1}{z} \mathrm{Li}_{n-1}(z)$$

- Symbol is introduced to condense the information of iterated integrals

$$I = \int_a^b d\log R_1 \circ \cdots \circ d\log R_n \Rightarrow \mathcal{S}[I] = R_1 \otimes \cdots \otimes R_n$$

[Goncharov, Spradlin, Vergu, Volovich, 1006.5703]

e.g. $\mathcal{S}[\log(x)\log(y)] = x \otimes y + y \otimes x, \quad \mathcal{S}[\mathrm{Li}_2(x)] = -(1-x) \otimes x$

$$\mathcal{S}[G(x,0,y)] = -x \otimes x + (x-y) \otimes x - y \otimes x + y \otimes (x-y)$$

- For scattering amplitudes/observables, alphabet $\{R_i\}$ and thus Symbol are constrained by physics information [DNA]

Two types of bootstrap

1. Scattering amplitude bootstrap:

- e.g. Three-point form factor $gg \to Hg$ in planar $\mathcal{N} = 4$ SYM to eight-loop

Kinematic variables & alphabet → Possible symbols × unknown rational constants → Conditions: integrability, symmetry etc. → OPE data → Integrate symbols to functions

Alphabet: $\mathcal{L}_u = \{u, v, w, 1-u, 1-v, 1-w\}$

Constraints ⇒ Linear systems

Ansatz: $\varepsilon^{(L)} = \sum_j c_j F_j^{(2L)}$ c_j: rational constants $F_j^{(2L)} \sim \mathcal{L}_{i_1} \otimes \mathcal{L}_{i_2} \cdots \otimes \mathcal{L}_{i_{2L}}, \quad \mathcal{L}_i \in \mathcal{L}_u$

[Dixon, McLeod, Wilhelm, 2012.12286] [Dixon, Gürdoğan, McLeod, Wilhelm, 2204.11901]

2. Landau bootstrap:

$$\mathcal{I}(p_i, m_e) = \int \prod_{j=1}^{L} \mathrm{d}^D k_j \frac{N(p_i, k_j)}{\prod_{e=1}^{E}(q_e^2 - m_e^2 + i\varepsilon)},$$

- Identify potential singularities using Landau equations

Alphabet ⇒ All possible symbols

- Study local expansion, leading singularities, sequential discontinuities etc.

Solve the unknown coefficients

[Hannesdottir, McLeod, Schwartz, Vergu, 2410.02424]

Numerical approach

Question: Can we use numerical data to exactly reconstruct coefficients?

- High-precision Feynman loop integrals:

Integration-by-Parts (IBP) [Chetyrkin, Tkachov, 1981]

Rational functions — Master integrals (basis)

$$\mathcal{I}(p_i, m_e) = \int \prod_{j=1}^{L} \mathrm{d}^D k_j \frac{N(p_i, k_j)}{\prod_{e=1}^{E} (q_e^2 - m_e^2 + i\varepsilon)} \Rightarrow \mathcal{I} = \sum_a R_a(p_i \cdot p_j, m_e)\, \mathrm{MI}_a$$

Auxiliary mass flow method (AMFlow) [Liu, Ma, 1801.10523; 2107.01864; 2201.11669; 2201.11637]

1. Introduce an auxiliary mass η to some of the propagators
2. Set up closed differential equations w.r.t η using IBPs
3. Solve the differential equations numerically with boundary conditions $\eta \to \infty$

Many higher-loop Feynman integrals are accessible numerically through AMFlow

Analytic regression

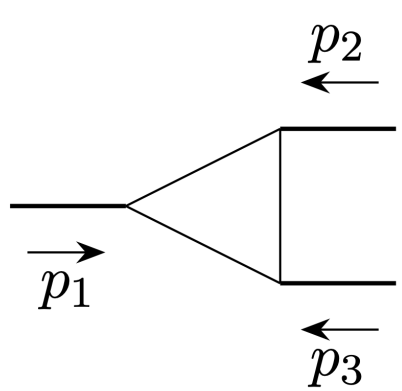
- We can formulate the bootstrap as an analytic regression problem:

Full function, e.g. Feynman integrals

$$f(\mathbf{x}_j) = \sum_{i=1}^{n} c_i \, \mathcal{B}_i(\mathbf{x}_j), \qquad c_i \in \mathbb{Q}$$

Rational coefficients to determine

Basis functions, like polylogarithms integrated from symbols



$$= p_1^2 T_1(p_1^2, p_2^2, p_3^2) = 2\frac{\mathrm{Li}_2(z)}{z-\bar{z}} - 2\frac{\mathrm{Li}_2(\bar{z})}{z-\bar{z}} + \frac{\ln(z\bar{z})\ln\left(\frac{1-z}{1-\bar{z}}\right)}{z-\bar{z}}$$

- Numerical basis functions:
    - FiberSymbol in PolyLogTools + GPL evaluation [Duhr, Dulat, 1904.07279]
    - Can always integrate the symbols numerically along a path [analytically in the first and last entries]

$$\tilde{B} = \sum_{i=1}^{|\tilde{A}|} c_{i_1,i_2,i_3,i_4} \int_0^1 d\log L_{i_4}(\lambda_4) \int_0^{\lambda_4} d\log L_{i_3}(\lambda_3) \int_0^{\lambda_3} d\log L_{i_2}(\lambda_2) \int_0^{\lambda_2} d\log L_{i_1}(\lambda_1)$$

Analytic regression

$$f(\mathbf{x}_j) = \sum_{i=1}^{n} c_i \mathcal{B}_i(\mathbf{x}_j), \quad c_i \in \mathbb{Q}$$

- Can we use matrix inversion to solve the unknown coefficients?

$$M_{ij} = \mathcal{B}_i(\mathbf{x}_j) \quad \Rightarrow \quad c_i = (M^{-1})_{ij} \cdot f(x_j)$$

1. Hard to impose coefficients are rational.

2. Matrix inversion loses precision easily:

$$\frac{||\Delta c||}{||c||} \approx \kappa(M) \left(\frac{||\Delta f||}{||f||} + \frac{||\Delta M||}{||M||} \right) \qquad \kappa(M) = ||M|| ||M^{-1}||$$

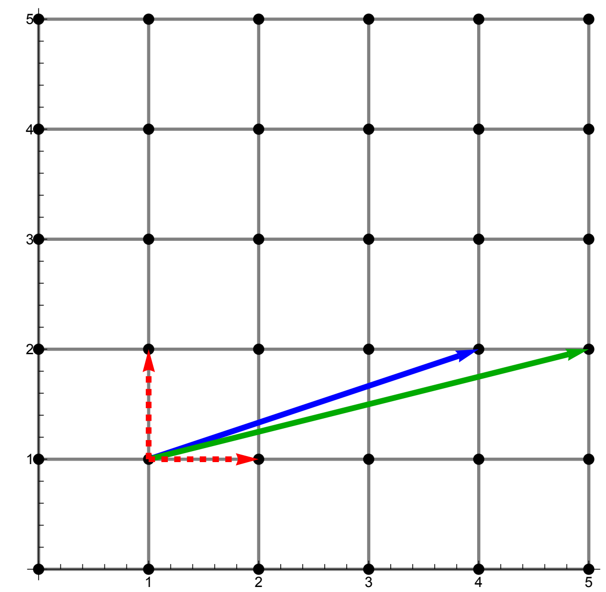
- Precision lost amplified by the matrix condition number κ
- When all entries are the same order of magnitude and normally distributed, $\kappa \sim n, n \gg 1$
  When not uniform, κ is much larger

- [In progress]: with Jesse Thaler

Lattice reduction

- Shortest vector problem: to reduce the initial basis of a lattice into an equivalent basis but with shortest vectors

$$\mathcal{L} = \left\{ \sum_{i=1}^{m} z_i \mathbf{b}_i \middle| z_i \in \mathbb{Z}, \mathbf{b}_i \in B \right\}$$



- Lenstra–Lenstra–Lovás (LLL) algorithm, 1982

  Mathematica implementation LatticeReduce[], C++ implementation FPLLL

- Question: How to use lattice reduction in analytic regression?

Lattice reduction

True function: $f(x) = G(0,1;x) - G(1,-1;x)$ $\quad x_1 = 4/10, x_2 = 9/10$

Basis: $\mathcal{B}(x) = \{G(1,0;x), G(0,1;x), G(0,-1;x), G(1,-1;x)\}$

$$M = \text{round}\ 10^s \left(\begin{array}{c|c} f(\mathbf{x_1}) \cdots f(\mathbf{x_p}) & \\ \mathcal{B}_1(\mathbf{x_1}) \cdots \mathcal{B}_1(\mathbf{x_p}) & 10^{-s}\, \mathbb{I}_{n+1} \\ \vdots & \\ \mathcal{B}_n(\mathbf{x_1}) \cdots \mathcal{B}_n(\mathbf{x_p}) & \end{array} \right) = \begin{pmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \mathbf{v}_3 \\ \mathbf{v}_4 \\ \mathbf{v}_5 \end{pmatrix} = \begin{pmatrix} -35 & -24 & 1 & 0 & 0 & 0 & 0 \\ 92 & 154 & 0 & 1 & 0 & 0 & 0 \\ -45 & -129 & 0 & 0 & 1 & 0 & 0 \\ 36 & 75 & 0 & 0 & 0 & 1 & 0 \\ -10 & -106 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\tilde{M} = \begin{pmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \mathbf{u}_3 \\ \mathbf{u}_4 \\ \mathbf{u}_5 \end{pmatrix} = \begin{pmatrix} 0 & -1 & 1 & 0 & -1 & 0 & 1 \\ 4 & -2 & 2 & 1 & 2 & 2 & 0 \\ 4 & 4 & -2 & -1 & -4 & -4 & 1 \\ 0 & -1 & 4 & 6 & 4 & -7 & -2 \\ 1 & -10 & -5 & 1 & 2 & -6 & -4 \end{pmatrix}$$

$\vec{u}$ and $\vec{v}$ are equivalent basis for the lattice (linear dependent)

$$\mathbf{u}_1 = \mathbf{v}_1 - \mathbf{v}_3 + \mathbf{v}_5$$

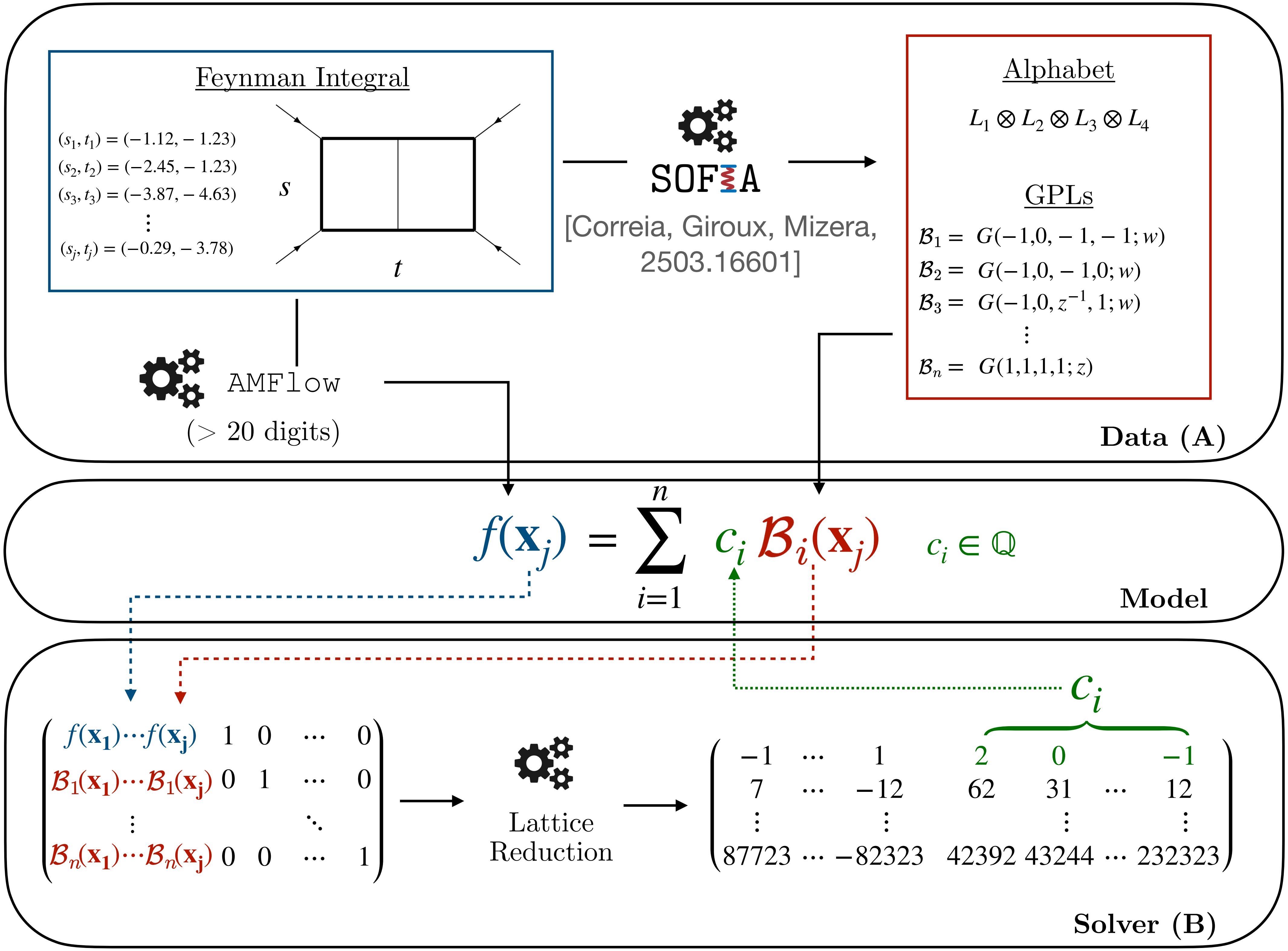
$$0 = f(x_1) - G(0,1;x_1) + G(1,-1;x_1)$$

$$-1 = 10^2 \left[f(x_2) - G(0,1;x_2) + G(1,-1;x_2) \right]$$

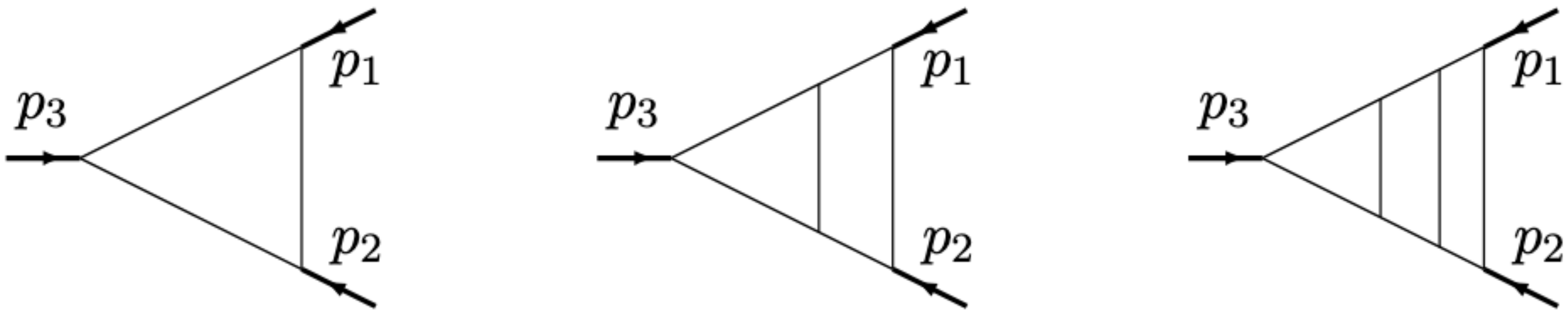
$-1/10^2$ reflects the finite precisions

In this example we use $d = 3$ digits, $s = 2$. We have 4 basis and 2 numerical points.

Analytic regression



Example I: Triangle ladder diagrams



$$z\bar{z} = p_2^2/p_1^2,$$
$$(1-z)(1-\bar{z}) = p_3^2/p_1^2$$

Landau analysis from SOFIA

$$A = \left\{ z\bar{z},\ (1-z)(1-\bar{z}),\ z-\bar{z},\ \frac{\bar{z}}{z},\ \frac{1-z}{1-\bar{z}} \right\}$$

[Correia, Giroux, Mizera, 2503.16601]

We write down all possible symbols up to weight-$2L$ at L-loop and apply integrability conditions for any pair $\{i_{p-1}, i_p\}$:

$$S = \sum_{i_1, i_2, \cdots i_N} c_{i_1, i_2, \cdots i_N} a_{i_1} \otimes a_{i_2} \cdots \otimes a_{i_N}$$

$$\sum_{i_1, i_2, \cdots i_N} c_{i_1, i_2, \cdots i_N} a_{i_1} \otimes a_{i_2} \cdots \otimes a_{i_{p-1}} \otimes a_{i_p} \cdots \otimes a_{i_N}\, d\log a_{i_{p-1}} \wedge d\log a_{i_p} = 0$$

We obtain linear combinations of symbols that are integrated to polylogarithmic functions

Example I: Triangle ladder diagrams

- Form the complete function space: add the product of ζ_n values and $G(\cdots)$

Blue terms are kept in case choosing the wrong branch

$$
\begin{aligned}
&\text{Weight-0:} \quad 1 \\
&\text{Weight-1:} \quad G(a_1,x),\ \pi \\
&\text{Weight-2:} \quad G(a_1,a_2,x),\ \pi\times G(a_1,x),\ \zeta_2 \\
&\text{Weight-3:} \quad G(a_1,a_2,a_3,x),\ \pi\times G(a_1,a_2,x),\ \zeta_2\times G(a_1,x),\ \zeta_3 \\
&\text{Weight-4:} \quad G(a_1,a_2,a_3,a_4,x),\ \pi\times G(a_1,a_2,a_3,x),\ \zeta_2\times G(a_1,a_2,x),\ \pi^3\times G(a_1,x), \\
&\qquad \zeta_3\times G(a_1,x),\ \zeta_4 \\
&\text{Weight-5:} \quad G(a_1,a_2,a_3,a_4,a_5,x),\ \pi\times G(a_1,a_2,a_3,a_4,x),\ \zeta_2\times G(a_1,a_2,a_3,x), \\
&\qquad \pi^3\times G(a_1,a_2,x),\ \zeta_3\times G(a_1,a_2,x),\ \zeta_4\times G(a_1,x),\ \zeta_5,\ \zeta_2\times\zeta_3 \\
&\text{Weight-6:} \quad G(a_1,a_2,a_3,a_4,a_5,a_6,x),\ \pi\times G(a_1,a_2,a_3,a_4,a_5,x),\ \zeta_2\times G(a_1,a_2,a_3,a_4,x), \\
&\qquad \pi^3\times G(a_1,a_2,a_3,x),\ \zeta_3\times G(a_1,a_2,a_3,x),\ \zeta_4\times G(a_1,a_2,x),\ \zeta_5\times G(a_1,x), \\
&\qquad \zeta_2\zeta_3\times G(a_1,x),\ \pi^5\times G(a_1,x),\ \zeta_6,\ \zeta_3^2 \\
&\ldots
\end{aligned}
$$

- We randomly pick points in the Euclidean region $0 < z < \bar{z} < 1$: all GPL are real.
- Obtain the overall rational factor from maximal cut
- Perform lattice reduction

Example I: Triangle ladder diagrams

Diagram	AMFlow point time	Transcendental weights	# points sampled	Basis size	Reduction time
One-loop	15.6 CPU-min	≤ 2	5	full(32) simplified(26) uniform(25)	<1s <1s <1s
Two-loop	1.1 CPU-h	≤ 4	100 100 60	full(488) simplified(393) uniform(366)	9.6 min 10.7 min 3.5 min
Three-loop	5.7 CPU-h	≤ 6	- - 200	full(1373) simplified(972) uniform(806)	- - 1.1 h

3 digits

20 digits

20 digits

Note: we only run AMFlow; in practice, should combine with IBP+ numeric DE

Analytic results

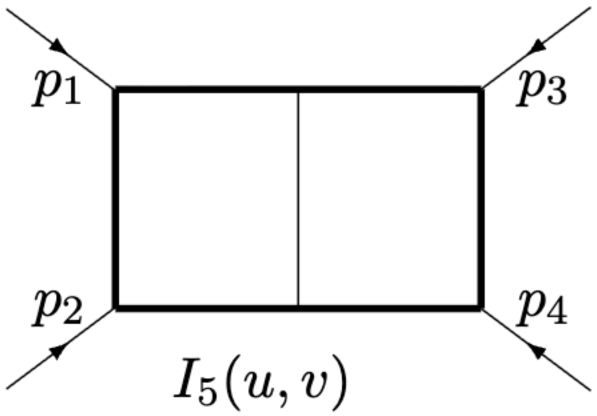
$$T_1(z) = \frac{1}{z-\bar{z}}\left[2\mathrm{Li}_2(z) - \mathrm{Li}_2(\bar{z}) + \log(z\bar{z})\log\left(\frac{1-z}{1-\bar{z}}\right)\right],$$

$$T_2(z) = \frac{1}{(1-z)(1-\bar{z})(z-\bar{z})}\Big[6\mathrm{Li}_4(z) - 6\mathrm{Li}_4(\bar{z}) - 3\log(z\bar{z})\left(\mathrm{Li}_3(z) - \mathrm{Li}_3(\bar{z})\right) + \frac{1}{2}\log^2(z\bar{z})\left(\mathrm{Li}_2(z) - \mathrm{Li}_2(\bar{z})\right)\Big],$$

$$T_3(z) = \frac{1}{(1-z)^2(1-\bar{z})^2(z-\bar{z})}\Big[20\mathrm{Li}_6(z) - 20\mathrm{Li}_6(\bar{z}) - 10\log(z\bar{z})\left(\mathrm{Li}_5(z) - \mathrm{Li}_5(\bar{z})\right) + \log^2(z\bar{z})\left(\mathrm{Li}_4(z) - \mathrm{Li}_4(\bar{z})\right) - \frac{1}{6}\log^3(z\bar{z})\left(\mathrm{Li}_3(z) - \mathrm{Li}_3(\bar{z})\right)\Big]$$

Example II: The two-loop outer-mass double box

[Hannesdottir, McLeod, Schwartz, Vergu, 2410.02424]



$$u = -\frac{4m^2}{s},\ v = -\frac{4m^2}{t},\ \beta_u = \sqrt{1+u},\ \beta_v = \sqrt{1+v},\ \beta_{uv} = \sqrt{1+u+v}$$

$$\tilde{A} = \Big\{u, v, 1+u, 1+v, u+v, 1+u+v, \frac{\beta_u-1}{\beta_u+1}, \frac{\beta_v-1}{\beta_v+1}, \frac{\beta_{uv}-1}{\beta_{uv}+1}, \frac{\beta_{uv}-\beta_u}{\beta_{uv}+\beta_u}, \frac{\beta_{uv}-\beta_v}{\beta_{uv}+\beta_v}, \frac{\beta_{uv}-\beta_u\beta_v}{\beta_{uv}+\beta_u\beta_v}\Big\}$$

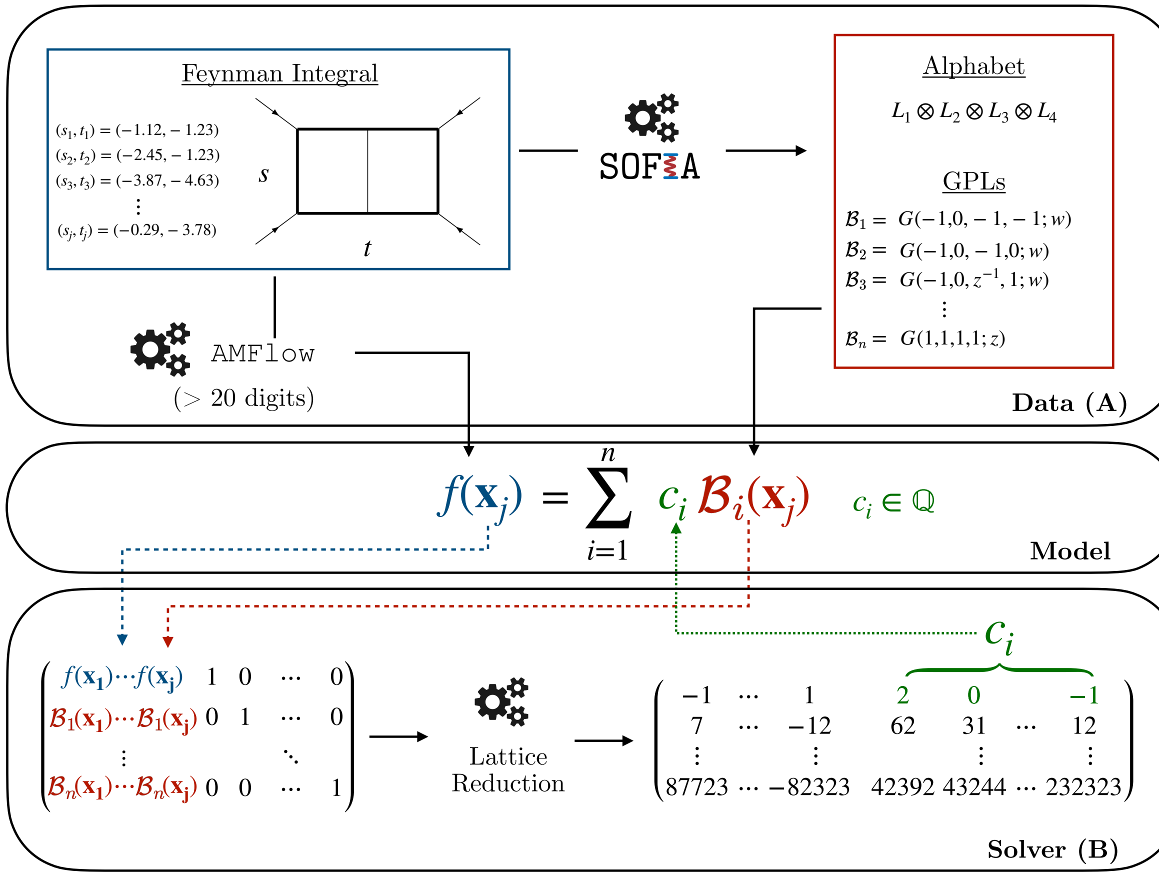
Combining Landau bootstrap and analytic regressions

$$\frac{1}{s^2 t \beta_u \beta_{uv}} \times \Big[-a_6 \otimes \frac{a_1}{a_3} \otimes a_6 \otimes a_9 - a_6 \otimes \frac{a_1}{a_3} \otimes a_9 \otimes a_6$$
$$+ a_6 \otimes a_6 \otimes \frac{a_1 a_2}{a_3 a_5} \otimes a_9 + a_6 \otimes a_9 \otimes \frac{a_2}{a_5} \otimes a_6$$
$$+ a_6 \otimes a_6 \otimes a_8 \otimes a_6 + a_6 \otimes a_9 \otimes a_8 \otimes a_9$$
$$+ a_7 \otimes a_{10} \otimes \frac{a_2}{a_5} \otimes a_6 + a_7 \otimes a_{10} \otimes a_8 \otimes a_9$$
$$+ a_7 \otimes a_7 \otimes \frac{a_1}{a_5} \otimes a_9 + a_7 \otimes a_7 \otimes a_8 \otimes a_6 \Big]$$

Bootstrap Techniques (Easy → Hard)			Numerical Fit (Hard ← Easy): Time | Digits
SOFIA	All Symbols	20736	
Linear systems $c_{ij} J_{ij} = 0$	Integrability	6993	??|??
Even integral under $\sqrt{\cdot} \to -\sqrt{\cdot}$	Galois symmetry	861	$t > 1h \,\|\, d > 25$
Impose α-positivity on $\log(\cdot)$ branch cut	Physical branch cuts	161	$t \sim 30s \,\|\, d = 17$
Discontinuity analysis $\Delta_\lambda I = c \int_{h_\lambda} dI$	Genealogical constraints	28	$t < 1s \,\|\, d = 8$
Landau equations for $s = 4m^2$	α-positive thresholds	6	$t < 1s \,\|\, d = 6$

Toward QCD bootstrap

- So far I introduce analytic regression method with lattice reduction
- It can be combined with Landau bootstrap at any step
- However, we haven't studied Feynman integrals with multiple rational factors



- Similar problem for QCD amplitude bootstrap:
  - Unlike $\mathcal{N} = 4$ SYM, QCD amplitude is not uniform transcendental and one needs to bootstrap rational functions instead of numbers
  - There are less available properties and data for QCD

- A baby step: a bootstrap framework for three-point energy correlators

Problem setup

- We consider the LO 3-point energy correlators, with arbitrary positive power of energies (A finite subset of N-point correlators)

$$F(\{x_{jl}\},\{a_k\}) \equiv \sum_m \sum_{1\le i_1<\cdots<i_n\le m} \int d\sigma_m \times \prod_{1\le k\le n} \frac{E_{i_k}^{a_k}}{Q} \prod_{1\le j<l\le n} \delta\left(x_{jl} - \frac{1-\cos\theta_{i_j i_l}}{2}\right)$$

- Collinear limit:

$$n = 3, m = 4$$

$$\frac{d^3\sigma_{a,b,c}}{dx_{12}dx_{13}dx_{23}} = \left(\frac{\mu^2 e^{\gamma_E}}{4\pi}\right)^{2\epsilon} \frac{2^{4-a-b-c} g^4 \Theta(-\tilde{\Delta}_3)}{(4\pi)^{5-2\epsilon}\Gamma(1-2\epsilon)} (-\tilde{\Delta}_3)^{-\frac{1}{2}-\epsilon}$$

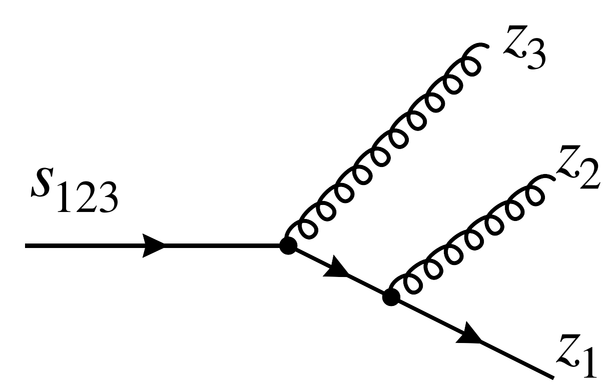
QCD splitting functions

$$G_0(z) \equiv \times \int dz_1 dz_2 dz_3\, \delta(1-z_1-z_2-z_3) z_1^{a+1-2\epsilon} z_2^{b+1-2\epsilon} z_3^{c+1-2\epsilon} \times \frac{P_{1\to 3}(z_1,z_2,z_3)}{s_{123}^2}$$

$$+ \text{perms of } (x_{12},\, x_{13},\, x_{23})$$

$G(z) \equiv$ "Symmetrization": 6 permutations of angles

$$z\bar{z} = x_{13}/x_{12},\ (1-z)(1-\bar{z}) = x_{23}/x_{12}$$



- Finite at $d=4$: as long as $a, b, c \ge 1$
- Not uniform transcendental, involves complicated rational functions

[Chen, Luo, Moult, Yang, XYZ, Zhu, 1912.11050]

Workflow

- A hybrid framework equipped with lattice reduction

Integrands:
$G_0(z) \sim \int z_1^{a+1} z_2^{b+1} z_3^{c+1} \times \frac{P_{1\to3}(z_1, z_2, z_3)}{s_{123}^2}$

↓ Spherical contour

Weight-2 symbols ⇒ functions $G_0^{(2)}(z)$

↓ Single-valued; Spurious pole conditions

Lower weight ansatz $G_0^{(1)}(z) + G_0^{(0)}(z)$ with unknown coefficients partially fixed

↓ Permutations

$G(z) = G_0(z) + G_0(1-z) + \cdots$

↑ Power correction data in squeezed (OPE) limit

Or

↑ Numerical integration + lattice reduction

Spherical contour: [Landau bootstrap]

a method to obtain leading transcendentality of iterated integrals with quadratic singularities

[Arkani-Hamed, Yuan, 1712.09991]

Physical constraints: [amplitude bootstrap]

consistency conditions in different singular limits

$$z\bar{z} = x_{13}/x_{12}, \ (1-z)(1-\bar{z}) = x_{23}/x_{12}$$

Numerical data: [analytic regression]

One example integral

- An example integrand:

$$\int dz_1 dz_2 dz_3\, \delta(1-z_1-z_2-z_3) z_1^{a+1} z_2^{b+1} z_3^{c+1} \times \frac{1}{s_{123}s_{12}z_3}\left(\frac{1}{z_1}+\frac{1}{z_2+z_3}\right)$$

- Leading transcendentality: from spherical contour:

$$\Rightarrow \mathcal{S}(F_1) = -\frac{(1-z)(1-\bar{z})}{2(z-\bar{z})^5}(-z^2-\bar{z}^2-4z\bar{z}+3z^2\bar{z}+3z\bar{z}^2)\left[\frac{1}{(1-z)^2(1-\bar{z})^2}\otimes\frac{\bar{z}}{z}+\frac{1}{z^2\bar{z}^2}\otimes\frac{1-z}{1-\bar{z}}\right]$$

Integrate to Bloch-Wigner function $-8iD_2^-(z) = -4\left(\mathrm{Li}_2(z)-\mathrm{Li}_2(\bar{z})\right) - 2\ln(z\bar{z})\ln\frac{1-z}{1-\bar{z}}$

- Lower transcendentality: all possible poles + single-valued functions

$$\frac{1}{(z-\bar{z})^{n_1}(z\bar{z})^{n_2}((1-z)(1-\bar{z}))^{n_3}} \times \left[\sum_{i,j=0}^{i,j\leq n} a_{i,j} z^i \bar{z}^j + \sum_{i,j=0}^{i,j\leq n} b_{i,j} z^i \bar{z}^j \log(z\bar{z}) + \sum_{i,j=0}^{i,j\leq n} c_{i,j} z^i \bar{z}^j \log((1-z)(1-\bar{z})) + \sum_{i,j=0}^{i,j\leq n} d_{i,j} z^i \bar{z}^j \pi^2\right]$$

- Requiring spurious pole cancellation determines most of the coefficients
- The rest can be fixed by lattice reduction

Bootstrap summary

- Number of parameters: all colors channels in QCD

$$E_i^{a_i} E_j^{a_j} E_k^{a_k}$$

Energy weights	Ansatz	$z - \bar{z}$	Squeezed	Sym.	Squeezed	Indep.
{1,2,1}	6984	1218	395	374	347	171
{1,1,2}	6835	1129	362	343	318	169
{2,2,1}	9045	1478	577	566	470	214
{2,1,2}	8871	1374	555	545	445	220
{1,2,2}	8995	1428	563	553	471	234
{3,1,1}	9114	1477	595	583	485	235
{1,3,1}	9357	1580	611	598	516	248
{1,1,3}	8841	1344	525	516	430	221
{2,2,2}	11311	1703	759	748	602	281

Can be determined by Lattice reduction with 40-50 numerical points

Towards a QCD bootstrap

- Question: Can we also bypass the spherical contour method and some physical constraints in the energy correlator bootstrap?

- If no information from spherical contour, we need a more robust form of rational functions

$$\frac{1}{(z-\bar{z})^{n_1}(z\bar{z})^{n_2}((1-z)(1-\bar{z}))^{n_3}} \times \sum_{i,j=0}^{i,j\leq n} a_{i,j} z^i \bar{z}^j$$

And assign such an ansatz for each transcendental function

- This significantly increases the number of unknown parameters.

- Lattice reduction is powerful, but not enough.

Limitation of lattice reduction

- First intuition:

# of basis function

digits of precision

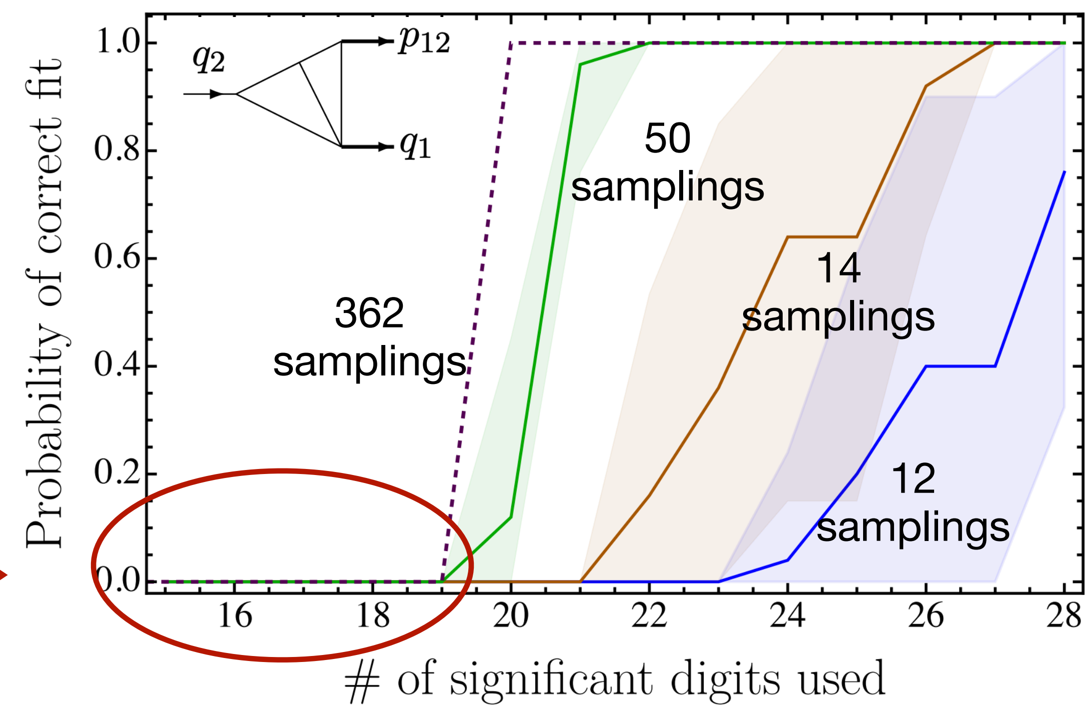
size of integers

$$d \lesssim \frac{nR}{p}$$

# of points

- However, given a Feynman integral, there exists a minimal digit

Always fail below some digit threshold

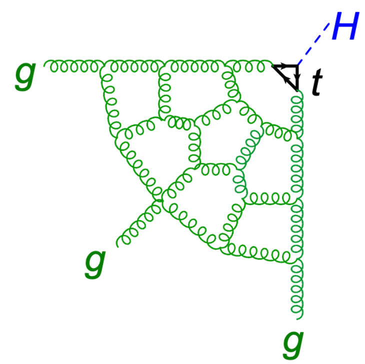


- Barriers towards QCD bootstrap through analytic regressions:
  - Understand the samplings and reduce the minimal digit value
  - Reduce the time cost

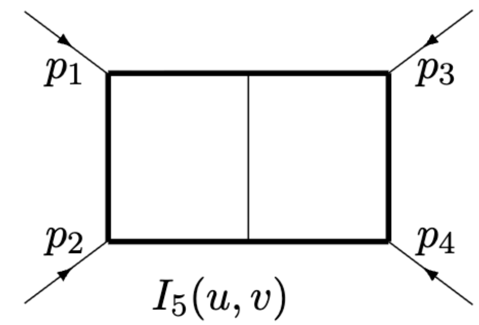
Summary

- Precision calculations are very crucial in studying QCD and collider physics.
- There are mainly two perturbative bootstrap programs:

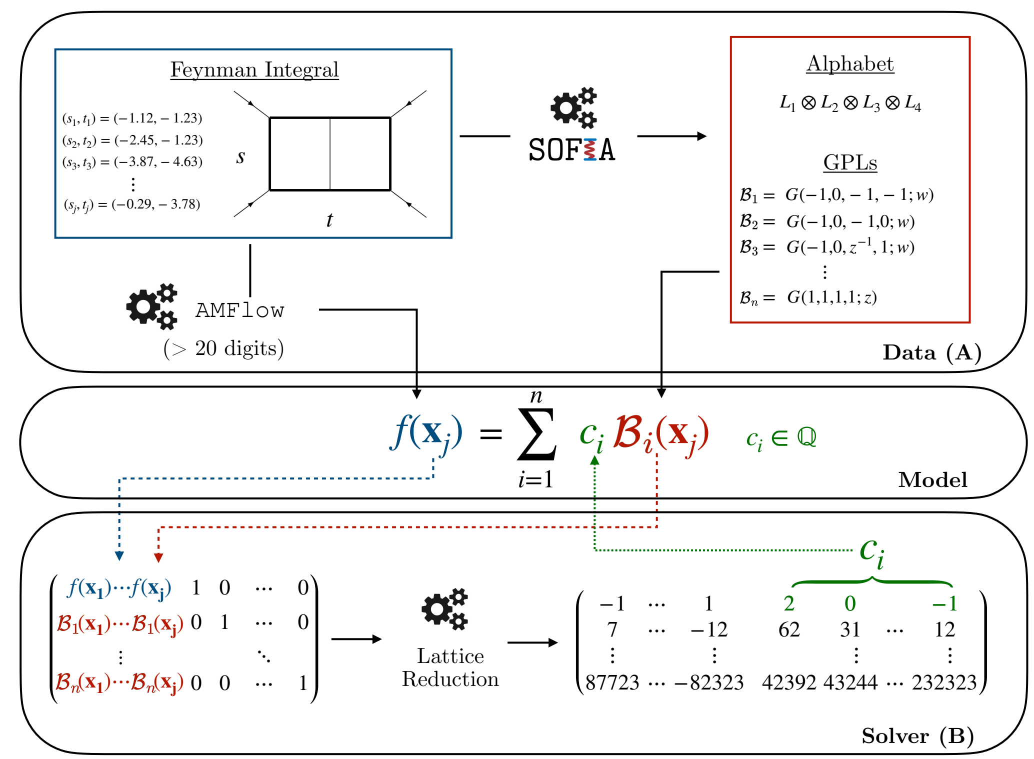
Amplitude bootstrap



Landau bootstrap



- We equip bootstrap with analytic regression method, lattice reduction.
- We develop a hybrid bootstrap program for energy correlators in QCD and obtain analytic results for three-point correlators with arbitrary positive energy powers.
- Can we extend it to a full bootstrap framework, such that it works for higher-points and higher-loops? —Improve lattice reduction



Move towards a QCD bootstrap that connects theory to precision experiments

Backup on energy correlators

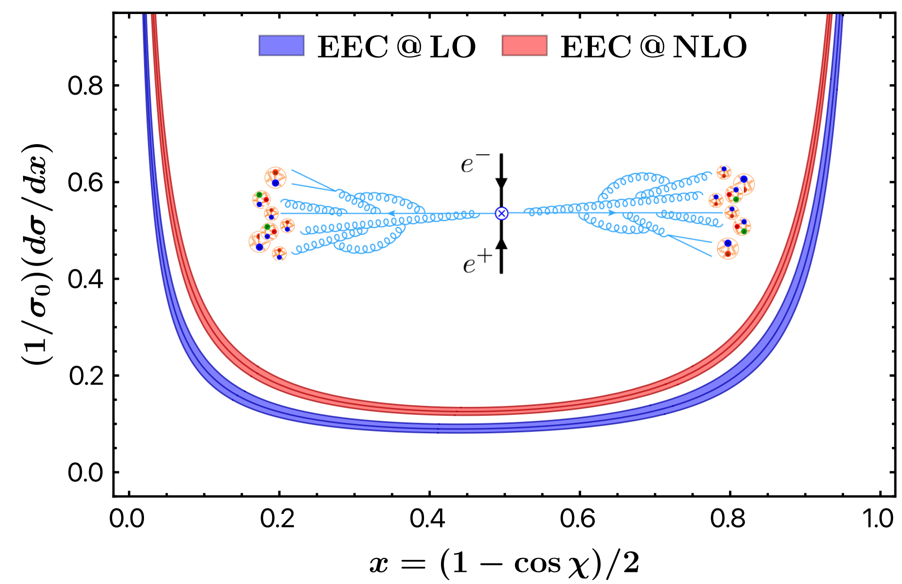
Energy correlators

- Energy-weighted cross section as a function of angles among any two detectors

Energy-energy correlation (2-point), EEC

[Basham, Brown, Ellis, Love, 1978]

$$\mathrm{EEC}(\chi) = \sum_{i,j} \int d\sigma \frac{E_i E_j}{Q^2} \delta(\vec{n}_i \cdot \vec{n}_j - \cos\chi)$$



N-point energy correlator

[Chen, Luo, Moult, Yang, XYZ, Zhu, 1912.11050]

$$\frac{d\sigma}{dx_{12}\cdots dx_{(n-1)n}} = \sum_m \sum_{1\le i_1,\cdots i_n \le m} \int d\sigma_m \times \prod_{1\le k\le n} \frac{E_{i_k}}{Q} \prod_{1\le j<l\le n} \delta\left(x_{jl} - \frac{1-\cos\theta_{i_j i_l}}{2}\right)$$

- Projecting the N-point energy correlator into a one-dimensional distribution: projected N-point correlator (ENC) [Chen, Moult, XYZ, Zhu, 2004.11381]

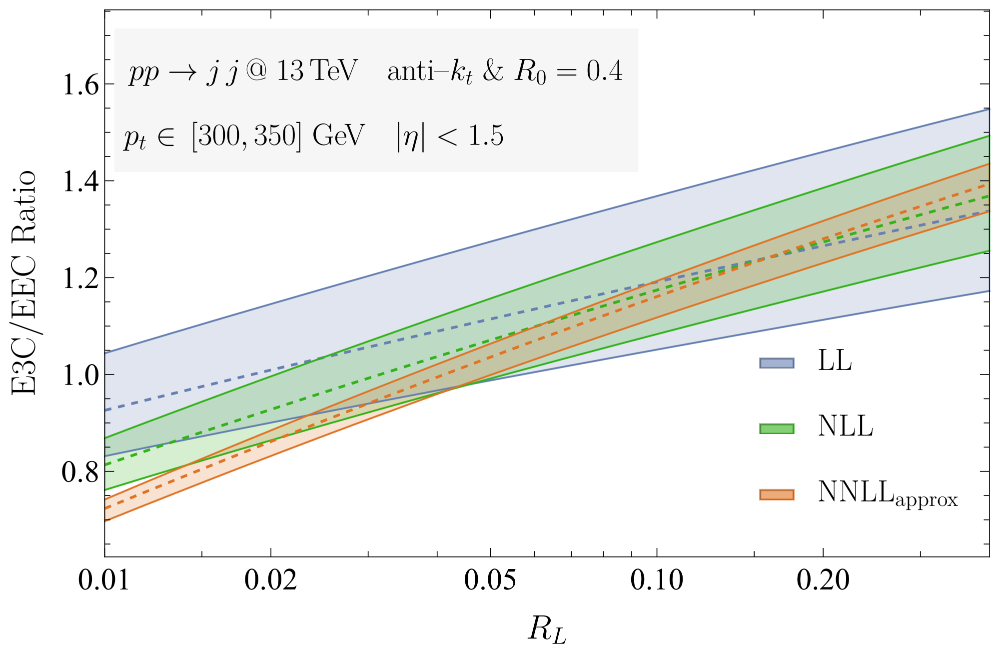
$$\frac{d\sigma^{[N]}}{dx_L} = \sum_m \sum_{1\le i_1,\cdots i_N \le m} \int d\sigma_m \left(\prod_{1\le k\le N} \frac{E_i}{Q}\right) \delta\left(x_L - \max\{x_{i_1,i_2}, x_{i_1,i_3}, \cdots x_{i_{N-1},i_N}\}\right)$$

m is the number of final-state particles

$x_L = (1-\cos\theta_L)/2$ being the largest angular distance

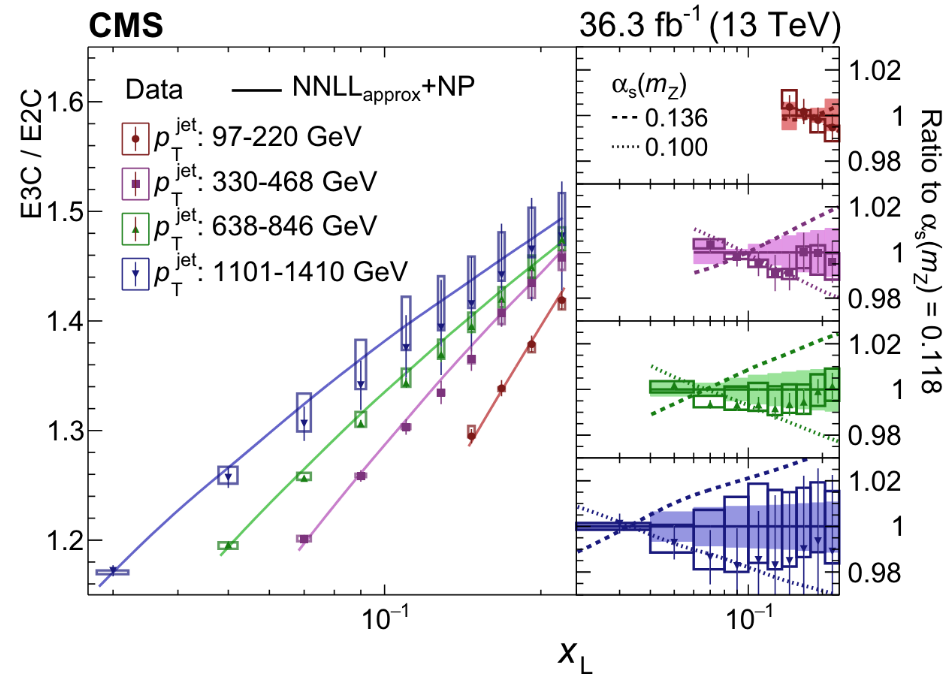
Energy correlators

- CMS measurement of projected 3-point correlator [CMS collaboration, 2402.13864]



NLL: [Lee, Meçaj, Moult, 2205.03414]

aNNLL: [Chen, Gao, Li, Xu, XYZ, Zhu, 2307.07510]

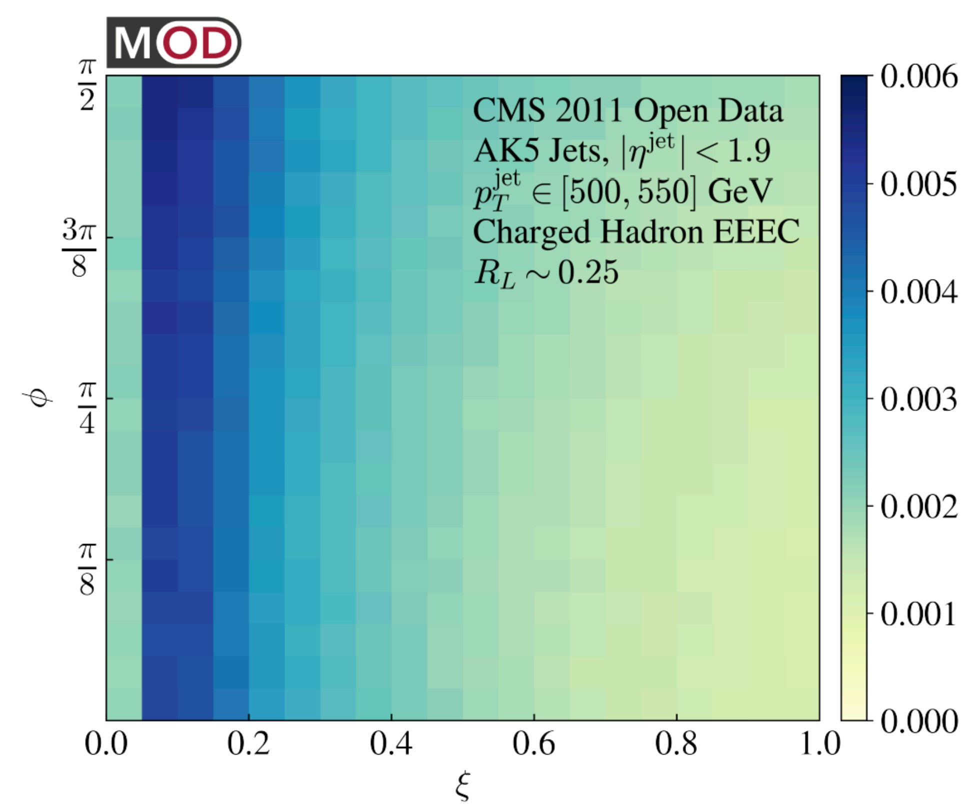


$$\alpha_s(m_Z) = 0.1229^{+0.0014}_{-0.0012}(\text{stat})^{+0.0030}_{-0.0033}(\text{theo})^{+0.0023}_{-0.0036}(\text{exp})$$

- CMS Open Data: 3-point correlator

[Komiske, Moult, Thaler, Zhu, 2201.07800]

[Chen, Moult, Thaler, Zhu, 2205.02857]



- More analytic results for energy correlators are needed

Backup on lattice reduction

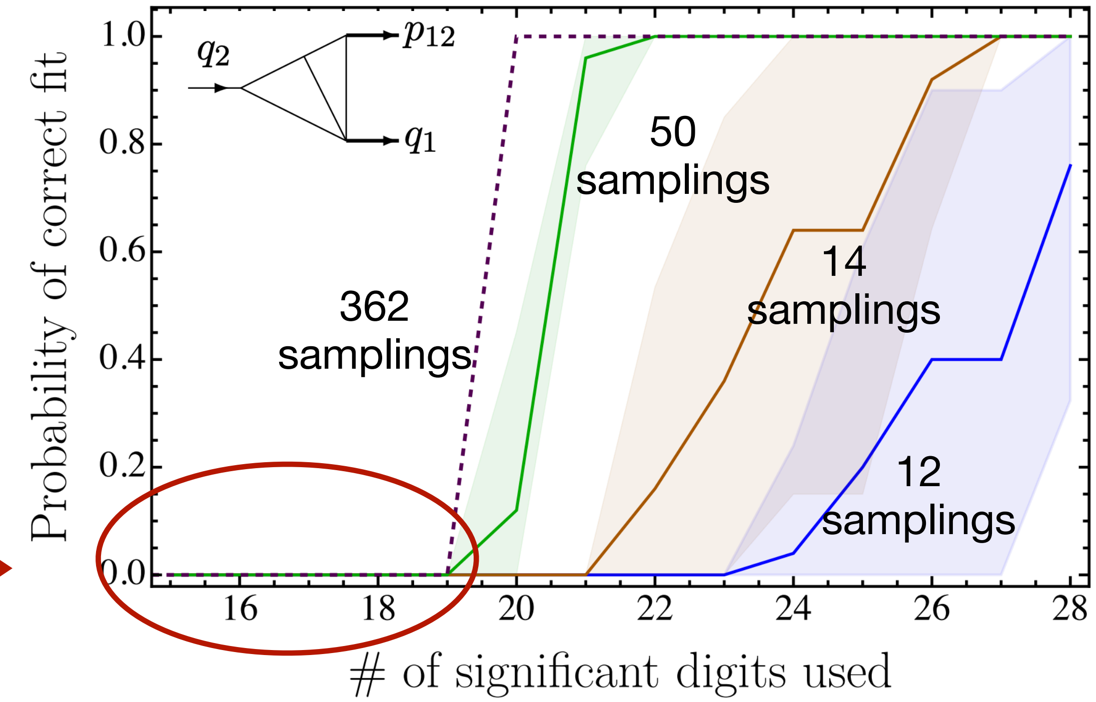
Limitation of lattice reduction

- First intuition:

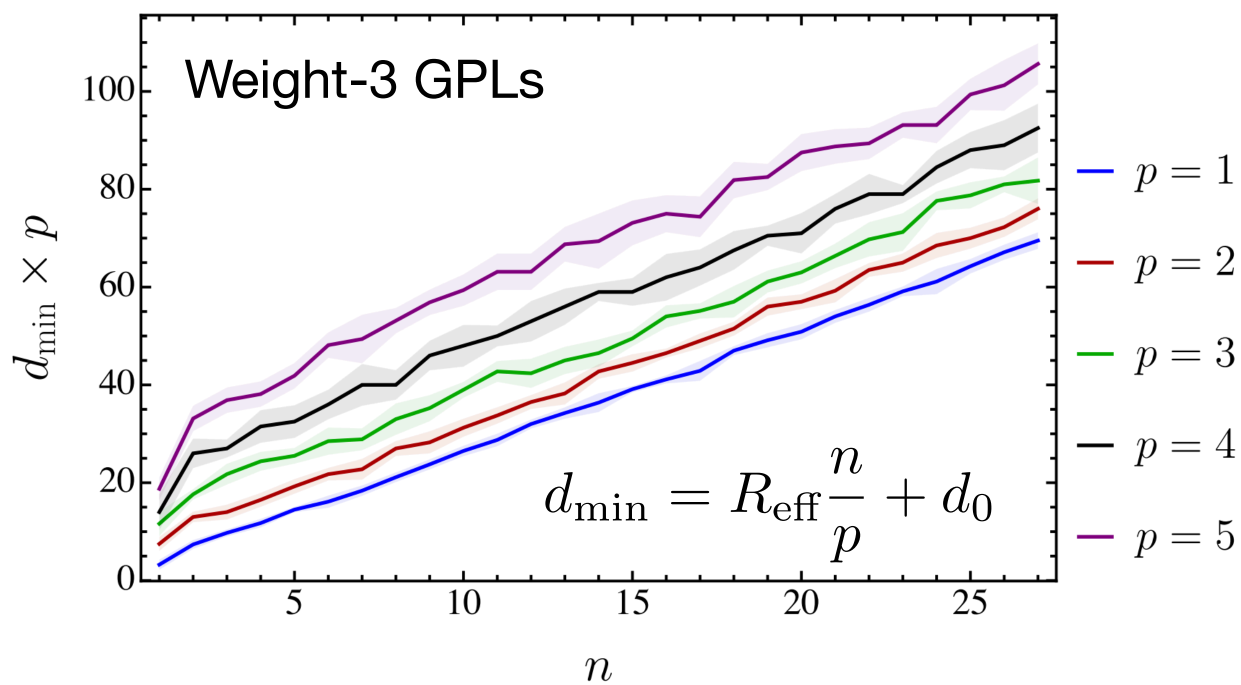
$$d \lesssim \frac{nR}{p}$$

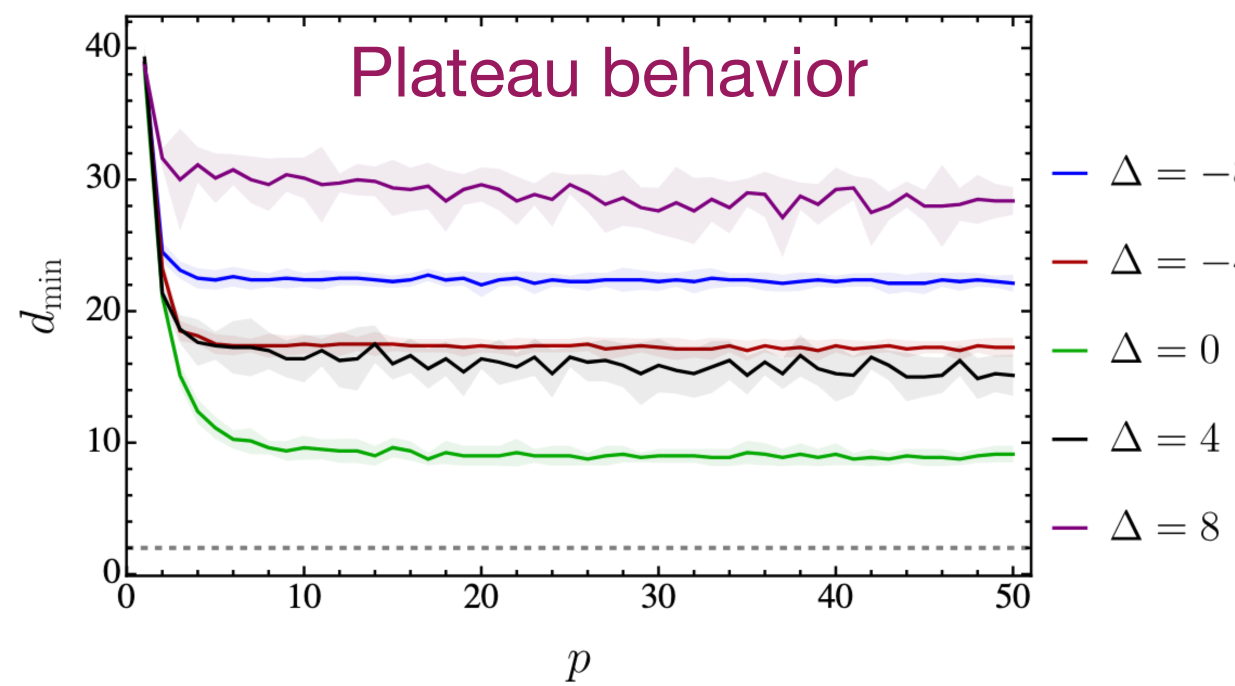
digits of precision → d; # of basis function → n; size of integers → R; # of points → p

Always fail below some digit threshold



- Experiments:



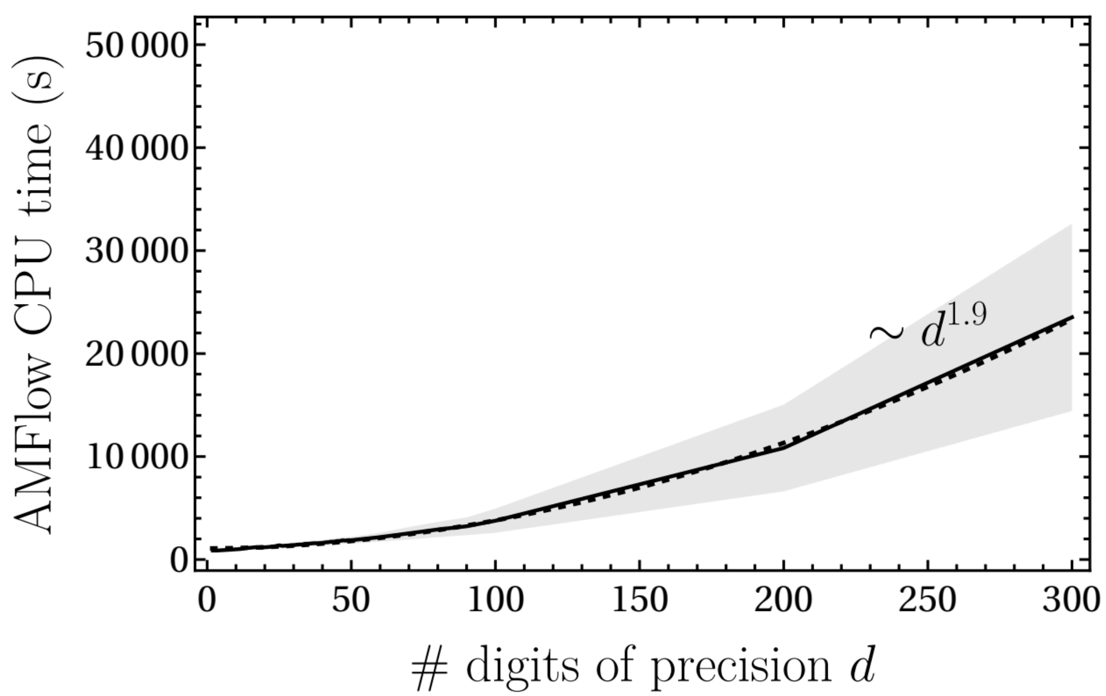


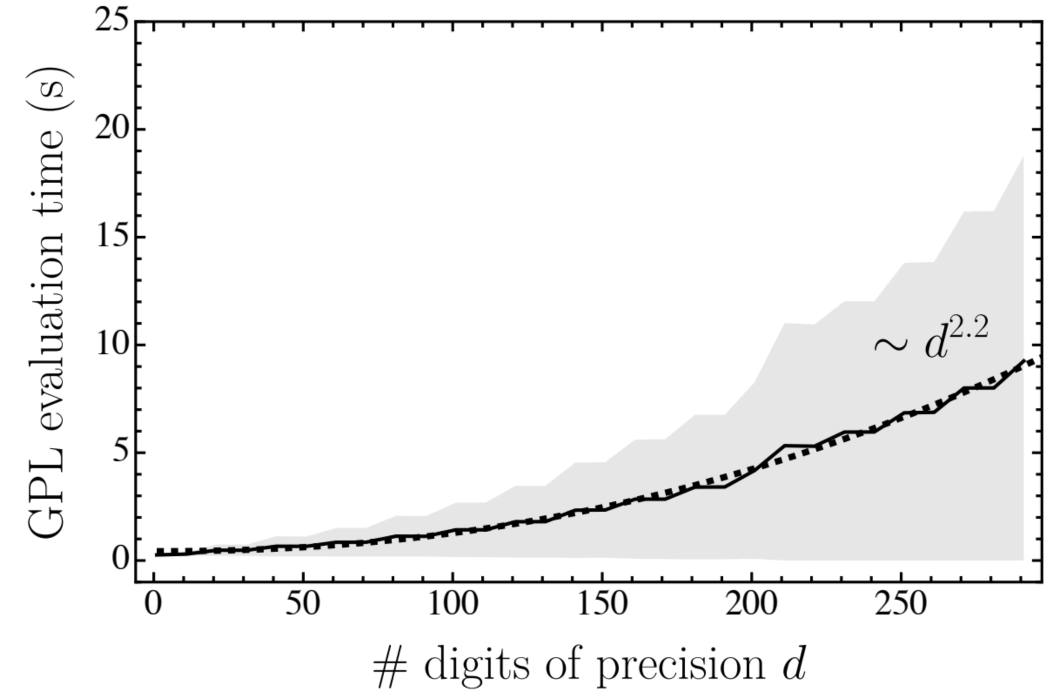
$\mathcal{B}(x) = \{1235.1, 2.3,\ 2.5\}$ Lose information if only 1-2 digit

$n = 15$ single-value GPL(x), with randomly sampling in $1/2 - 10^{\Delta} \leq x \leq 1/2 + 10^{\Delta}$

Limitation of lattice reduction

- Time cost:
  - Lattice reduction: $t(p,n) \sim p \times n^4$, no parallel implementation





Very approximately: $$t(p,n)/\mathrm{ns} \approx \underbrace{10^9 \cdot p \cdot d_{\min}^2(n)}_{\textsc{AMFlow}} + \underbrace{10^4 \cdot n \cdot p \cdot d_{\min}^2(n)}_{\textsc{Ginsh}} + \underbrace{p \cdot n^4}_{\text{fitting}}$$

- Barriers towards QCD bootstrap through analytic regression:
  - Understand the samplings and reduce the digit plateau value
  - Improve the time cost: the n^4 scaling

Backup on
ENC bootstrap

Step1: leading transcendentality

- We start the $G_0(z)$ integrand:

Example terms in splitting functions

$$\int dz_1 dz_2 dz_3\, \delta(1-z_1-z_2-z_3) z_1^{a+1} z_2^{b+1} z_3^{c+1} \times \frac{1}{s_{123}s_{12}z_3}\left(\frac{1}{z_1}+\frac{1}{z_2+z_3}\right)$$

- Feynman parameterization:

$$F_1 = 3\int_0^\infty \frac{dz_1 dz_2 dz_3 dz_4}{\mathrm{GL}(1)} \times \frac{z_2 z_3 z_4^2}{[z_1 z_2 + (z\bar{z}) z_2 z_3 + (1-z)(1-\bar{z}) z_1 z_3 + (z_1+z_2+z_3) z_4]^4} \Rightarrow (XQX)^4$$

$$Q_1 = \frac{1}{2}\begin{pmatrix} 0 & 1 & (1-z)(1-\bar{z}) & 1 \\ 1 & 0 & z\bar{z} & 1 \\ (1-z)(1-\bar{z}) & z\bar{z} & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$

[Arkani-Hamed, Yuan, 1712.09991]

- Read off the symbol entries and calculate the spherical contour residues

$$\mathcal{S}(F_1) = \mathrm{Disc}_{24}\mathrm{Disc}_{13}F_1 \times \left[\frac{1}{(1-z)^2(1-\bar{z})^2} \otimes \frac{\bar{z}}{z}\right] + \mathrm{Disc}_{14}\mathrm{Disc}_{23}F_1 \times \left[\frac{1}{z^2\bar{z}^2} \otimes \frac{1-z}{1-\bar{z}}\right]$$

$$\Rightarrow \mathcal{S}(F_1) = -\frac{(1-z)(1-\bar{z})}{2(z-\bar{z})^5}(-z^2-\bar{z}^2-4z\bar{z}+3z^2\bar{z}+3z\bar{z}^2)\left[\frac{1}{(1-z)^2(1-\bar{z})^2}\otimes\frac{\bar{z}}{z}+\frac{1}{z^2\bar{z}^2}\otimes\frac{1-z}{1-\bar{z}}\right]$$

- Integrate to Bloch-Wigner function

$$-8iD_2^-(z) = -4\left(\mathrm{Li}_2(z)-\mathrm{Li}_2(\bar{z})\right) - 2\ln(z\bar{z})\ln\frac{1-z}{1-\bar{z}}$$

Step1: leading transcendentality

- Example result: $C_F T_F n_f$ color structure in the quark jet

$$
\begin{aligned}
G_0^{(2)}(z) \supset K = & \frac{1}{48}\pi^2 \left(-z^3 - z^2(\bar{z}-2) - z\left(\bar{z}^2 - 2\bar{z} + 2\right) - \bar{z}^3 + 2\bar{z}^2 - 2\bar{z} + 1\right) \\
& + \frac{1}{8}\left(z^3 + z^2(\bar{z}-2) + z\left(\bar{z}^2 - 2\bar{z} + 2\right) + \bar{z}^3 - 2\bar{z}^2 + 2\bar{z} - 1\right) \\
& \times \left[\mathrm{Li}_2\left(1 - \frac{z\bar{z}}{(1-z)(1-\bar{z})}\right) + \frac{1}{2}\log\left(\frac{1}{(1-z)(1-\bar{z})}\right)\log\left(\frac{z\bar{z}}{(1-z)(1-\bar{z})}\right)\right] \\
& + \frac{iD_2^-(z)}{4(z-\bar{z})^{11}}\Big[z^{14} - 2z^{13}(5\bar{z}+1) + z^{12}\left(45\bar{z}^2 + 20\bar{z} + 2\right) - z^{11}\left(120\bar{z}^3 + 90\bar{z}^2 + 20\bar{z} + 1\right) \\
& + 11z^{10}\bar{z}\left(19\bar{z}^3 + 22\bar{z}^2 + 8\bar{z} + 1\right) - 11z^9\bar{z}^2\left(22\bar{z}^3 + 40\bar{z}^2 + 20\bar{z} + 5\right) \\
& + z^8\bar{z}\left(-675\bar{z}^5 + 3114\bar{z}^4 - 2495\bar{z}^3 + 1615\bar{z}^2 - 330\bar{z} + 25\right) \\
& + z^7\bar{z}\left(-1440\bar{z}^6 + 7740\bar{z}^5 - 16838\bar{z}^4 + 14605\bar{z}^3 - 6500\bar{z}^2 + 1255\bar{z} - 76\right) \\
& + z^6\bar{z}\left(-675\bar{z}^7 + 7740\bar{z}^6 - 26050\bar{z}^5 + 41281\bar{z}^4 - 31230\bar{z}^3 + 11695\bar{z}^2 - 1889\bar{z} + 85\right) \\
& + z^5\bar{z}\left(-242\bar{z}^8 + 3114\bar{z}^7 - 16838\bar{z}^6 + 41281\bar{z}^5 - 52400\bar{z}^4 + 33897\bar{z}^3 - 10674\bar{z}^2 + 1348\bar{z} - 36\right) \\
& + z^4\bar{z}^2\left(209\bar{z}^8 - 440\bar{z}^7 - 2495\bar{z}^6 + 14605\bar{z}^5 - 31230\bar{z}^4 + 33897\bar{z}^3 - 18570\bar{z}^2 + 4615\bar{z} - 360\right) \\
& + z^3\bar{z}^3\left(-120\bar{z}^8 + 242\bar{z}^7 - 220\bar{z}^6 + 1615\bar{z}^5 - 6500\bar{z}^4 + 11695\bar{z}^3 - 10674\bar{z}^2 + 4615\bar{z} - 720\right) \\
& + z^2\bar{z}^4\left(45\bar{z}^8 - 90\bar{z}^7 + 88\bar{z}^6 - 55\bar{z}^5 - 330\bar{z}^4 + 1255\bar{z}^3 - 1889\bar{z}^2 + 1348\bar{z} - 360\right) \\
& + z\bar{z}^5\left(-10\bar{z}^8 + 20\bar{z}^7 - 20\bar{z}^6 + 11\bar{z}^5 + 25\bar{z}^3 - 76\bar{z}^2 + 85\bar{z} - 36\right) + \bar{z}^{11}\left(\bar{z}^3 - 2\bar{z}^2 + 2\bar{z} - 1\right)\Big]
\end{aligned}
$$

Very singular behavior:

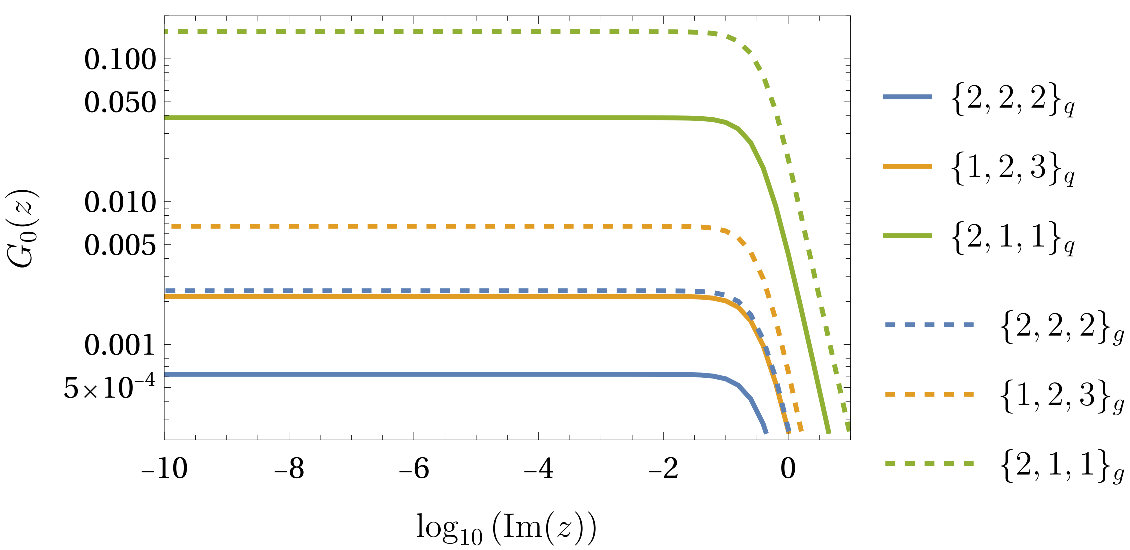
$$
K \overset{z-\bar{z}\to 0}{\approx} \frac{189}{(z-\bar{z})^{10}}(\bar{z}-1)^5\bar{z}^5\left(2\bar{z}^2 - 2\bar{z} + 1\right)\left(\bar{z}\log(1-\bar{z}) - \log(1-\bar{z}) - \bar{z}\log(\bar{z})\right) + \mathcal{O}\left[(z-\bar{z})^{-9}\right]
$$

$$
K \overset{z\to 0}{\approx} \frac{9t^6\left(t^8 + 10t^6 + 20t^4 + 10t^2 + 1\right)(\log(r) - 1)}{2r^4\left(t^2-1\right)^{10}} + \mathcal{O}(r^{-3}), \quad z \equiv r \times t,\ \bar{z} \equiv r/t
$$

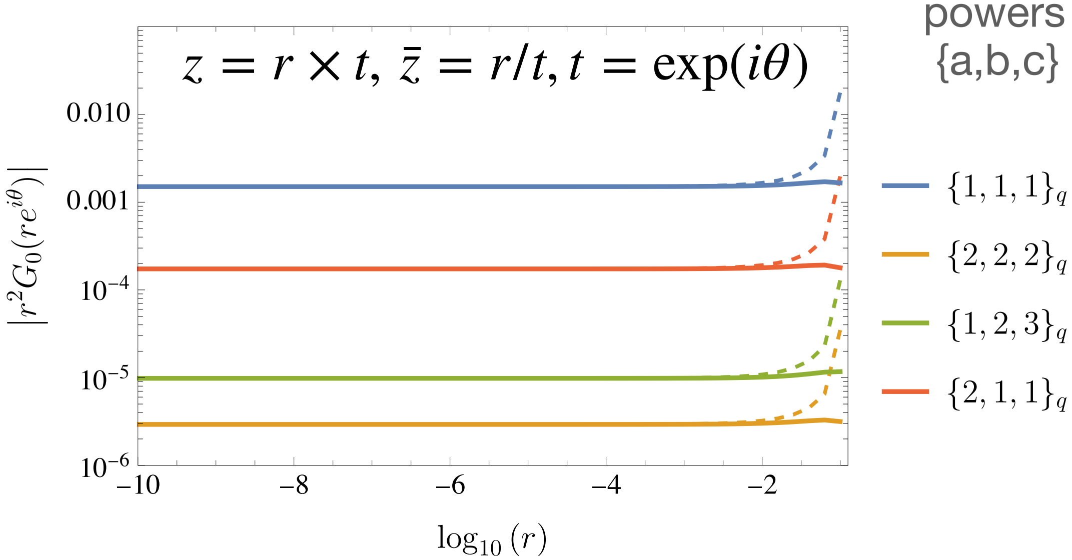
Step2: lower-weight function

- Different kinematic limits

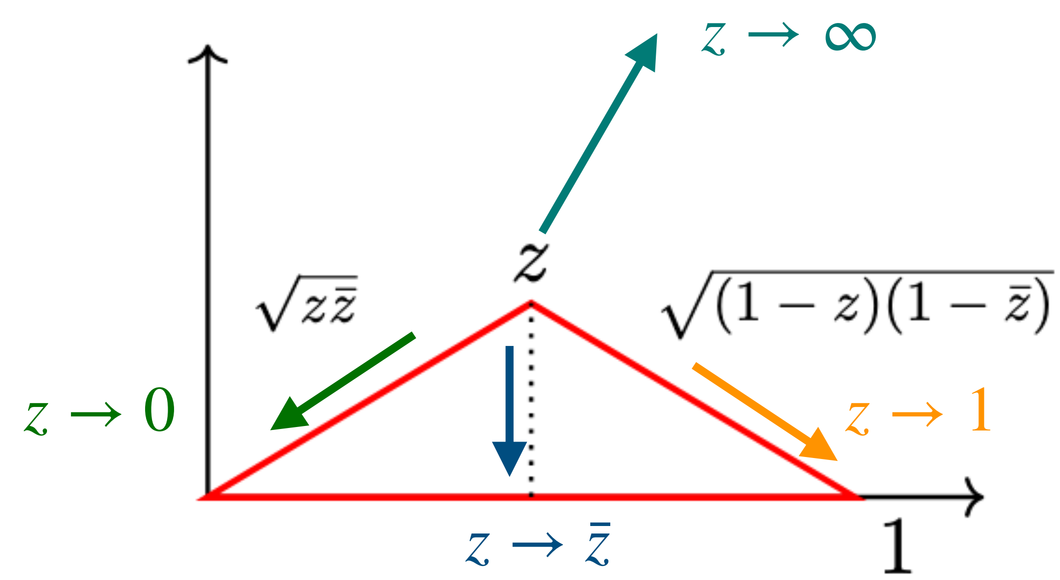
(1). Collapsed limit $z \to \bar{z}$: non-singular



(2). Squeezed limit $z \to 0$: $1/(z\bar{z})$



Nontrivial cancellation between weight-2 and lower-weights



Ansatz Single-valued functions, $\log(z\bar{z})$ etc

$$\frac{1}{(z-\bar{z})^{n_1}(z\bar{z})^{n_2}((1-z)(1-\bar{z}))^{n_3}} \times \left[\sum_{i,j=0}^{i,j\le n} a_{i,j} z^i \bar{z}^j + \sum_{i,j=0}^{i,j\le n} b_{i,j} z^i \bar{z}^j \log(z\bar{z}) + \sum_{i,j=0}^{i,j\le n} c_{i,j} z^i \bar{z}^j \log((1-z)(1-\bar{z})) + \sum_{i,j=0}^{i,j\le n} d_{i,j} z^i \bar{z}^j \pi^2\right]$$

$n_{1,2,3}$ match to weight-2 denominators

Step2: lower-weight function

1. Any power divergences of $z - \bar{z}$ have to vanish
2. When $z \to 0$, any power divergences or logarithmic divergences that are faster than $1/(z\bar{z})$ (or $1/r^2$) need to vanish.
3. Similar for $z \to 1$, $z \to \infty$

Channel	Ansatz	$z-\bar{z}$	$z=1$	$z=0$	$z=\infty$	Symmetrization	r^{-2}	r^0	r^2	r^4
$\mathcal{N}=4$	70	27	26	17	16	12	2	0	0	0
$P_{qq'\bar{q}'}$	664	99	98	97	28	24	13	0	0	0
$P_{qq\bar{q}}$	931	216	135	134	85	81	59	31	0	0
$P_{qgg}^{(C_F)}$	442	27	26	6	6	6	2	0	0	0
$P_{qgg}^{(C_A)}$	664	99	98	97	28	24	13	0	0	0
$P_{gq\bar{q}}^{(C_F)}$	511	46	45	21	20	16	5	0	0	0
$P_{gq\bar{q}}^{(C_A)}$	748	133	132	33	32	28	14	0	0	0
P_{ggg}	1134	319	318	197	116	107	86	51	7	0
QCD Sum	5094	939	852	585	315	288	192	82	7	0

[Chen, Moult, Zhu, 2011.02492, 2104.00009, 2202.04085]

Determine remaining parameters: power correction data in $z \to 0$ limit, need calculations

Step3: lattice reduction

- Question: can we bypass the tedious calculations for squeezed limit data?

- Consider C_F^2 after applying the physical constraints

$$G(z) = D(z) + a_{0,11}\, g_1(z) + a_{0,12}\, g_2(z) + a_{1,12}\, g_3(z) + a_{2,12}\, g_4(z) + c_{1,12}\, g_5(z) + d_{2,12}\, g_6(z)$$

$$\begin{aligned} g_3(z) = \frac{1}{(z-1)^2 z^2 (\bar{z}-1)^2 \bar{z}^2} \Big[& z^4\bar{z}^2 - z^4\bar{z} + 8z^3\bar{z}^3 - 14z^3\bar{z}^2 + 8z^3\bar{z} - z^3 + z^2\bar{z}^4 - 14z^2\bar{z}^3 \\ & + 24z^2\bar{z}^2 - 14z^2\bar{z} + z^2 + \big(2z^4\bar{z}^3 - 3z^4\bar{z}^2 + z^4\bar{z} + 2z^3\bar{z}^4 - 8z^3\bar{z}^3 + 9z^3\bar{z}^2 - 4z^3\bar{z} + z^3 \\ & - 3z^2\bar{z}^4 + 9z^2\bar{z}^3 - 12z^2\bar{z}^2 + 9z^2\bar{z} - 3z^2 + z\bar{z}^4 - 4z\bar{z}^3 + 9z\bar{z}^2 - 8z\bar{z} + 2z + \bar{z}^3 - 3\bar{z}^2 \\ & + 2\bar{z}\big) \log((z-1)(\bar{z}-1)) + \big(-2z^4\bar{z}^3 + 3z^4\bar{z}^2 - z^4\bar{z} - 2z^3\bar{z}^4 + 8z^3\bar{z}^3 - 9z^3\bar{z}^2 + 2z^3\bar{z} \\ & + 3z^2\bar{z}^4 - 9z^2\bar{z}^3 + 6z^2\bar{z}^2 - z\bar{z}^4 + 2z\bar{z}^3\big) \log(z\bar{z}) - z\bar{z}^4 + 8z\bar{z}^3 - 14z\bar{z}^2 + 8z\bar{z} - \bar{z}^3 + \bar{z}^2 \Big] \end{aligned}$$

- Treat $\{g_i\}$ as transcendental basis and $a_{i,j}, b_{i,j}, c_{i,j}, d_{i,j}$ as coefficients to determine

- (1). Run lattice reduction among $\{g_i(z)\}$ to get a linear-independent basis $\{\tilde{g}_i(z)\}$
- (2). Run lattice reduction among both $G(z)$ and $\{\tilde{g}_i(z)\}$

In this case, 2 numerical points and 13 digits fix all six parameters

Step3: lattice reduction

- $\mathcal{N} = 4$ and all QCD channels with reasonable computation cost

Channel	# of basis	# of linearly-independent basis	# of points	Digits
$\mathcal{N} = 4$	12	6	2	5
$P_{qq'\bar{q}'}$	24	12	10	16
$P_{qq\bar{q}}$	81	39	30	31
$P_{qgg}^{(C_F)}$	6	6	2	13
$P_{qgg}^{(C_A)}$	24	12	10	14
$P_{gq\bar{q}}^{(C_F)}$	16	8	3	11
$P_{gq\bar{q}}^{(C_A)}$	28	14	10	14
P_{ggg}	107	54	50	22

- Higher energy weights: Number of parameters

Lattice reduction

40-50 points

Energy weights	Ansatz	$z - \bar{z}$	Squeezed	Sym.	Squeezed	Indep.
{1,2,1}	6984	1218	395	374	347	171
{1,1,2}	6835	1129	362	343	318	169
{2,2,1}	9045	1478	577	566	470	214
{2,1,2}	8871	1374	555	545	445	220
{1,2,2}	8995	1428	563	553	471	234
{3,1,1}	9114	1477	595	583	485	235
{1,3,1}	9357	1580	611	598	516	248
{1,1,3}	8841	1344	525	516	430	221
{2,2,2}	11311	1703	759	748	602	281

Step4: results

- Validation:

(1) Increasing the number of points and digits in lattice reduction

(2) Compare obtained expression with numerical data using different points

(3) Expected behavior: