

Topical Review

Chapter 5: Analytic bootstraps for scattering amplitudes and beyond

Georgios Papathanasiou* 

Deutsches Elektronen-Synchrotron DESY, Notkestr. 85, 22607 Hamburg, Germany

E-mail: georgios.papathanasiou@desy.de

Received 29 April 2022, revised 14 June 2022

Accepted for publication 5 July 2022

Published 30 November 2022

**Abstract**

One of the main challenges in obtaining predictions for collider experiments from perturbative quantum field theory, is the direct evaluation of the Feynman integrals it gives rise to. In this chapter, we review an alternative bootstrap method that instead efficiently constructs physical quantities by exploiting their analytic structure. We present in detail the setting where this method has been originally developed, six- and seven-particle amplitudes in the large-color limit of $\mathcal{N} = 4$ super Yang–Mills theory. We discuss the class of functions these amplitudes belong to, and the strong clues mathematical objects known as cluster algebras provide for rendering this function space both finite and of relatively small dimension at each loop order. We then describe how to construct this function space, as well as how to locate the amplitude inside of it with the help of kinematic limits, and apply the general procedure to a concrete example: the determination of the two-loop correction to the first nontrivial six-particle amplitude. We also provide an overview of other areas where the realm of the bootstrap paradigm is expanding, including other scattering amplitudes, form factors and Feynman integrals, and point out the analytic properties of potentially wider applicability that it has revealed.

Keywords: scattering amplitudes, gauge theory, cluster algebras

(Some figures may appear in colour only in the online journal)

* Author to whom any correspondence should be addressed.

Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](https://creativecommons.org/licenses/by/4.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

Contents

1. Introduction	2
2. Planar $\mathcal{N} = 4$ SYM Amplitudes in a Nutshell	3
2.1. Planar limit, color-ordering and discrete symmetries	3
2.2. Helicity dependence	4
2.3. The kinematic space: dual conformal invariance and momentum twistors	4
2.4. Amplitude normalizations	6
3. Cluster polylogarithmic functions	8
3.1. Multiple polylogarithms and symbols	8
3.2. Cluster algebras and amplitude singularities	11
3.2.1. Basics of cluster algebras	11
3.2.2. Symbol letters from cluster variables	15
3.3. Cluster adjacency and extended Steinmann relations	17
4. The Steinmann Cluster Bootstrap for $\mathcal{N} = 4$ SYM Amplitudes	21
4.1. Restrictions on the first and last symbol entries	21
4.2. Constructing the function space containing the n-particle amplitude	23
4.3. Singling out the amplitude: special kinematic limits	24
4.4. A simple example: bootstrapping the two-loop six-particle MHV amplitude	25
5. New frontiers	32
5.1. $n > 7$ singularities from tropical Grassmannians	32
5.1.1. The positive tropical Grassmannian	33
5.1.2. A tropical sieve for rational letters	34
5.1.3. Square-root letters from infinite mutation sequences	36
5.2. Bootstrapping Feynman integrals	38
5.2.1. Survey of explicit results	38
5.2.2. The role of cluster algebras	40
5.3. Bootstrapping a three-particle form factor	42
Acknowledgments	44
Data availability statement	44
References	44

1. Introduction

The idea that analytic properties could be exploited so as to circumvent difficulties in obtaining predictions from the traditional perturbative quantum field theory approach is not new [1]. Yet it is only over the last decade that it has been applied so successfully in order to compute physical quantities for general values of their kinematic parameters, albeit mostly in the large-color limit of the simplest interacting four-dimensional gauge theory, known as $\mathcal{N} = 4$ super Yang–Mills (SYM).

In its modern reincarnation, this perturbative analytic bootstrap has been initiated by Dixon, Drummond and Henn in order to determine the three-loop correction to the essentially first nontrivial six-particle scattering amplitude of the theory [2], following a remarkable earlier

conjecture on the all-loop structure of amplitudes of any multiplicity [3], as well as the exploration of its consequences at strong coupling [4–7] via the gauge/string duality [8]. The cornerstone of the method is the construction of a finite space of functions expected to contain the physical quantity in question, from which the latter may then be uniquely determined from the knowledge of its behavior in special kinematic limits. Its upgrade to higher loops [2, 9–15] and to multiplicity seven [16–19] has been achieved in parallel with the discovery of new analytic properties [20, 21] that prune the initial space of functions, thereby making the method more efficient.

Perhaps more importantly, the discovered analytic properties have potential for applicability to more general gauge theories [22, 23]. And as a proof of concept that the bootstrap may also work for phenomenologically relevant physical quantities, it has been used to rederive the three-loop soft anomalous dimension in quantum chromodynamics (QCD), up to an overall numerical factor [24]. Further areas expanding the realm of the bootstrap paradigm include amplitudes in simpler kinematics [25–27] and three-particle form factors in $\mathcal{N} = 4$ SYM [28, 29], as well as individual Feynman integrals [9, 20, 30–34], or basis sets thereof [35, 36].

Given that there already exists a six-particle bootstrap review [37], as well as a more recent amplitude bootstrap review [38] based on a talk by the author, our goal here will be two-fold: on the one hand, to cover some of the key concepts behind the amplitude bootstrap in more detail, and to show how it works in a concrete example; and on the other hand, to give an overview of the new frontiers of its application. In section 2 we briefly introduce the $\mathcal{N} = 4$ SYM amplitudes of interest, and in section 3 we discuss the class of polylogarithmic functions they belong to. We also describe how mathematical objects known as cluster algebras provide strong clues for the singularities of these amplitudes at multiplicities $n = 6, 7$, as well as for how these singularities can appear consecutively. This is the information that renders the function space finite and of relatively small size, and we explain how to construct it and how to locate the amplitude inside it in section 4, also working out all the steps explicitly for the two-loop six-particle case in the simplest helicity configuration. We close with an overview of the new frontiers in section 5.

2. Planar $\mathcal{N} = 4$ SYM amplitudes in a nutshell

2.1. Planar limit, color-ordering and discrete symmetries

For the most part we will be focusing on $\mathcal{N} = 4$ SYM theory [39, 40], the gauge theory containing the maximal amount of supersymmetry, and whose particle content apart from gluons also contains their scalar and fermionic partners, all of which are in the adjoint representation. We will also be restricting to the origin in the moduli space of the theory, where the scalar fields have zero vacuum expectation value, and all particles are massless. Many of the properties of the theory and of its amplitudes are discussed in chapter 1 [41] of the SAGEX review [42], so here we will briefly recall some of their main features that we will need later on. Unless otherwise noted we will also be considering 't Hooft's planar limit [43], where the number of colors $N \rightarrow \infty$ while its product $g_{\text{YM}}^2 N$ with the gauge theory coupling is held fixed. The latter is the only parameter of the theory that survives in the limit, and in our normalization conventions we will be denoting the perturbative expansion of any quantity F with respect to it as

$$F = \sum_{L=0}^{\infty} g^{2L} F^{(L)}, \quad g^2 = \frac{g_{\text{YM}}^2 N}{16\pi^2}. \quad (1)$$

A major simplification that occurs in the planar limit, is that only a single color structure $\text{Tr}(T^{a_1} \dots T^{a_n})$ of the gauge group generators T^{a_i} is the leading term in the $1/N$ expansion of the n -particle amplitude, as is reviewed in [41] or [44]. We may thus restrict our attention to the coefficient of this leading color structure, the *color-ordered amplitude* A_n . Due to its relation to the trace, it is evident that it is invariant under cyclic shifts $i \rightarrow i + 1$ of the particle labels. What is less obvious, is that A_n is also invariant (up to an overall sign for odd n) under reflections $i \rightarrow n + 1 - i$, where the equivalence $n + i \sim i$ is understood. At tree level, this holds both for gluons and quarks in any gauge theory, and is a consequence of the (anti-)symmetry of the (three-) four-point vertex of the color-ordered Feynman rules, see e.g. [45]. As a consequence of supersymmetry, in $\mathcal{N} = 4$ SYM this property also persists at loop level and for all types of external states, after grouping them in a single superfield [46], and similarly combining all component amplitudes in a single superamplitude [47]. Together cyclic permutations and reflections form the discrete *dihedral group*, which is thus a symmetry group of A_n .

2.2. Helicity dependence

Having factored out the color degrees of freedom from the amplitude in the planar limit, the remaining physical quantities it will depend on are the momenta and helicities of the external particles. We recall that *helicity*, namely the projection of a spin in the direction of momentum is a good quantum number for massless particles, as it cannot be altered by Lorentz transformations. With the help of spinor-helicity variables, which as the name suggests are friendly to the aforementioned quantum number, and which by now have also made their way into quantum field theory textbooks [45, 48], it is straightforward to show that all gluon amplitudes where all or all but one helicities are positive (for example in conventions where all momenta are outgoing) vanish at tree level. While this vanishing is lifted at loop level in generic gauge theories (see [49] for a relatively recent example in QCD), it does persist in supersymmetric theories such as $\mathcal{N} = 4$ SYM, since it is a consequence of supersymmetric Ward identities [47, 50].

Hence the first nontrivial helicity configuration at any loop order corresponds to the so-called *maximally helicity violating* (MHV) amplitudes, with all but two gluons having positive helicity, and similarly amplitudes with all but $(k + 2)$ gluons of positive helicity are denoted as $N^k\text{MHV}$. As reviewed in e.g. [51], different distributions of the negative helicity states are also simultaneously accounted for in the superamplitude, which has a natural grading with respect to total helicity, following from the fact that the entire aforementioned superfield it depends on has well-defined helicity. Also note that helicity degrees k and $\bar{k} = n - 4 - k$ are related by a parity or spatial reflection transformation, which in spinor-helicity variables simply corresponds to complex conjugation of spinors. As a result, we may always restrict $k \leq \lfloor \frac{n-4}{2} \rfloor$, where $\lfloor x \rfloor$ denotes the integer part of x . Summarizing what we have discussed so far, the nontrivial physical quantity encoding scattering in planar $\mathcal{N} = 4$ SYM is the color-ordered n -particle, helicity degree- k superamplitude

$$A_{n,k}(p_1, \dots, p_n), \quad n \geq 4, \quad 0 \leq k \leq \lfloor \frac{n-4}{2} \rfloor, \quad (2)$$

where p_i denotes the momentum of the i th particle. In other words only $A_{n,0}$ or the MHV superamplitude is needed for $n = 4, 5$, additionally $A_{n,1}$ or the NMHV superamplitude is needed for $n = 6, 7$, and so on. Evidently, amplitudes obeying $k = (n - 4)/2$, such as $A_{4,0}$ or $A_{6,1}$, will be invariant under parity.

2.3. The kinematic space: dual conformal invariance and momentum twistors

With not only color but also helicity dependence specified as above, we now move on to describe what is the space of kinematics of the amplitude. Remarkably, planar $\mathcal{N} = 4$ SYM possesses *dual conformal symmetry* [4, 52–54], which acts on dual position variables x_i related to the usual momenta by

$$p_i \equiv x_{i+1} - x_i. \quad (3)$$

This symmetry, which is also reviewed in chapter 1 of this review [41], and combines with the usual conformal symmetry of the theory so as to form an infinite-dimensional Yangian symmetry [55], implies that the amplitude (in the appropriate normalization, which we will describe at the end of this section) depends on $3n - 15$ instead of $3n - 10$ kinematic variables¹ said differently, instead of an algebraically independent subset of the $n(n - 3)/2$ distinct Mandelstam invariants

$$s_{i,\dots,j-1} \equiv (p_i + p_{i+1} + \dots + p_{j-1})^2 = (x_i - x_j)^2 \equiv x_{ij}^2, \quad j \geq i + 1 \pmod n, \quad (4)$$

one must instead pick an algebraically independent subset of their distinct $n(n - 5)/2$ *conformal cross ratios*

$$u_{ij} \equiv \frac{x_{ij+1}^2 x_{i+1j}^2}{x_{ij}^2 x_{i+1j+1}^2}, \quad j \geq i + 2 \pmod n. \quad (5)$$

The fact that dual conformal invariance reduces the number of independent kinematic variables of the amplitude to $3n - 15$ has the following important implication: *Only normalized n -particle amplitudes with $n \geq 6$ have nontrivial kinematic dependence*, and will therefore be the focus of this article. For the simplest cases with $n = 6, 7$ the abbreviated notation

$$u_i \equiv u_{i+8-n, i+11-n} \quad (6)$$

for the cross ratios (5) is also used.

The algebraic relations among different Mandelstam invariants (cross ratios) are known as (conformal) Gram determinant constraints, and simply encode the fact that the number of independent vectors is bounded in a given spacetime dimension. In the presence of conformal symmetry, these constraints were worked out in [56], and while their solution yields an independent subset of cross ratios that can be used to parametrize the kinematics, in practice this parametrization turns out to be quite complicated.

Instead, massless, planar, dual conformal invariant kinematics may be most conveniently described in terms of *momentum twistors* [57], which are also very nicely reviewed in e.g. [58, 59]. Very briefly, one way for obtaining these variables is by representing $x^\mu \in \mathbb{R}^{1,3}$ as a projective null vector $X^M \in \mathbb{R}^{2,4}$, $X^2 = 0$, $X \sim \lambda X$. This $SO(2, 4)$ vector X^M is also equivalent to an antisymmetric representation X^{IJ} of $SU(2, 2)$, since the two algebras are isomorphic (in practice one representation can be converted into the other by six-dimensional analogues of the Pauli matrices, which may similarly be used in order to transform Lorentz vectors to 2×2 antisymmetric matrices). The antisymmetric representation can in turn be built out of two copies of the fundamental representation Z^I of $SU(2, 2)$, or, after complexifying, $SL(4, \mathbb{C})$.

¹ This counting corresponds to the number of independent components of the lightlike-separated points x_i , minus the dimension of the 4D conformal group or Poincaré group, for planar $\mathcal{N} = 4$ SYM or a generic massless gauge theory, respectively.

Momentum twistors precisely correspond to these Z 's, and we see that our original point $x^\mu \in \mathbb{R}^{1,3}$ is mapped to a pair of points, i.e. a line in momentum twistor space. As is the case for the vector X they originate from, momentum twistors are also defined up to rescalings $Z \sim tZ$, thus they may be equivalently viewed as homogeneous coordinates on complex projective space \mathbb{P}^3 . It is then possible to show that the usual Mandelstam invariants (4) can be expressed in terms of momentum twistors as

$$x_{ij}^2 \propto \langle i - 1ij - 1j \rangle, \tag{7}$$

up to proportionality factors that drop out from conformally invariant quantities, where

$$\langle ijkl \rangle \equiv \langle Z_i Z_j Z_k Z_l \rangle = \det(Z_i Z_j Z_k Z_l) \tag{8}$$

is a *four-bracket* of momentum twistors.

The advantage of momentum twistor variables is that they automatically satisfy both momentum conservation and the constraint on external lightlike momenta, $p_i^2 = x_{i+1}^2 = 0$: indeed, from equation (3) it is evident that the former already holds for the dual space coordinates giving rise to them, and the latter follows from equations (7) and (8). Furthermore, conformal transformations of the dual space coordinates x map to $SO(2, 4)$ rotations of X , and in turn to $SL(4, \mathbb{C})$ transformations of the momentum twistors. Therefore the space of dual conformal invariant kinematics can be written as a $4 \times n$ matrix, whose columns are the cyclically ordered momentum twistors/homogeneous \mathbb{CP}^3 coordinates defined up to rescalings, and modulo $SL(4, \mathbb{C})$ transformations. Fixing this gauge redundancy, so as to obtain explicit parametrizations of the kinematics in terms of $3n - 15$ independent variables is then very straightforward; one such example are the web variables, that may be algorithmically constructed for any n [60], see also [61, 62] for a simplified reformulation.

Let us see how these types of kinematic parametrizations may be used in practice in the $n = 6$ case. The matrix of momentum twistors in terms of web variables is

$$(Z_1, \dots, Z_6) = \begin{pmatrix} 1 & 0 & 0 & 0 & -1 & -1 - x_1 - x_1 x_2 - x_1 x_2 x_3 \\ 0 & 1 & 0 & 0 & 1 & 1 + x_1 + x_1 x_2 \\ 0 & 0 & 1 & 0 & -1 & -1 - x_1 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix}, \tag{9}$$

from which we can compute any four-bracket by choosing the corresponding 4×4 minor according to equation (8), such that e.g.

$$\langle 1346 \rangle = 1 + x_1 + x_1 x_2, \tag{10}$$

and similarly evaluate any other kinematic variable that depends on them. For example due to equation (7), the cross ratios (6) become

$$u_1 = \frac{x_2 x_3}{(1 + x_1 + x_1 x_2)(1 + x_2 + x_2 x_3)}, \quad u_2 = \frac{x_1 x_2}{1 + x_1 + x_1 x_2}, \quad u_3 = \frac{1}{1 + x_2 + x_2 x_3}.$$

It is worth noting that the space of kinematics in momentum twistor variables is also very closely related to the *Graßmannian* $\text{Gr}(m, n)$, defined as the space of m -dimensional planes going through the origin in n -dimensional space: from this definition, it follows that $\text{Gr}(m, n)$ may also be realized as an $m \times n$ matrix, this time modulo $GL(m)$ transformations. As was first noted in [63], based on the previously discovered relevance of the Graßmannian for the amplitude integrand [64], the comparison of their matrix realizations reveals that the space of external momentum kinematics is in fact equivalent to the quotient $\text{Gr}(4, n)/(\mathbb{C}^*)^{n-1}$.

2.4. Amplitude normalizations

Let us finally come to address the question of the normalization of the amplitude. In any massless gauge theory, loop amplitudes have infrared divergences, arising from integration regions where the loop momenta become soft or collinear. It can be shown that these divergences exponentiate quite universally [65], and that particularly in planar $\mathcal{N} = 4$ SYM they are captured to all loops by the Bern–Dixon–Smirnov (BDS) ansatz [3]. The latter is essentially the exponential of the one-loop amplitude times the cusp anomalous dimension Γ_{cusp} ,

$$\frac{1}{4}\Gamma_{\text{cusp}}(g^2) = g^2 - 2\zeta_2 g^4 + 22\zeta_4 g^6 - [219\zeta_6 + 8(\zeta_3)^2]g^8 + \dots, \quad (11)$$

also known to all loops [66] thanks to the integrability of the theory, as reviewed for example in [67]². So the BDS ansatz by construction satisfies the exponentiation property of infrared divergences, but also includes additional one-loop contributions.

From the above discussion, it therefore follows that it is possible to obtain an infrared-finite normalized amplitude by dividing out $A_{n,k}$ by the BDS ansatz. It should also make apparent, however, that this normalization is not unique: similarly to the difference between renormalization schemes in any gauge theory, there is still freedom in the finite, dual conformal invariant terms that the infrared-divergent factor may be chosen to absorb. While this choice ultimately leads to just equivalent representations of the amplitude, it proves advantageous to *tune* it such that the normalized amplitude inherits certain important physical properties of $A_{n,k}$, and hence becomes simpler to compute. Indeed, when n is not a multiple of 4, choosing to factor out the closely related BDS-like ansatz, which naturally appears in the strong-coupling description of the amplitude [6], ensures that the normalized amplitude respects the Steinmann relations, whose significance will be discussed in subsection 3.3.

In what follows, whenever possible we will thus focus on BDS-like normalized amplitude, denoted as

$$E_{n,k} \equiv \frac{A_{n,k}}{A_{n,0}^{(0)} A_n^{\text{BDS-like}}}, \quad n \bmod 4 \neq 0, \quad (12)$$

where for convenience we have also additionally divided by the tree-level MHV superamplitude, $A_{n,0}^{(0)}$. Since the precise form of the BDS and BDS-like ansätze will not be important for our purposes, we will refrain from quoting them here, and refer the interested reader to the original references, or e.g. [69, 70]. The ratio of the two ansätze is however closely related to the one-loop correction to $E_{n,0}$,

$$\frac{A_n^{\text{BDS}}}{A_n^{\text{BDS-like}}} = \exp\left[\frac{\Gamma_{\text{cusp}}}{4} E_{n,0}^{(1)}\right], \quad (13)$$

where for $n = 6, 7$ we explicitly have,

$$E_{6,0}^{(1)} = \sum_{i=1}^3 \text{Li}_2\left(1 - \frac{1}{u_i}\right), \quad (14)$$

$$E_{7,0}^{(1)} = \sum_{i=1}^7 \left[\text{Li}_2\left(1 - \frac{1}{u_i}\right) + \frac{1}{2} \log\left(\frac{u_{i+2}u_{i-2}}{u_{i+3}u_{i-3}}\right) \log u_i \right]. \quad (15)$$

²The integrability of a deformation of $\mathcal{N} = 4$ SYM theory is also described in chapter 9 [68] of the SAGEX review [42].

Converting between BDS and BDS-like normalizations also follows immediately from equation (13). For example, the BDS-normalized MHV amplitude, which in the original literature was expressed in terms of an exponentiated *remainder function* \mathcal{R}_n , is related to $\mathcal{E}_{n,0}$ by

$$e^{\mathcal{R}_n} = e^{-\frac{\Gamma_{\text{cusp}}}{4} \mathcal{E}_{n,0}^{(1)}} \mathcal{E}_{n,0}. \tag{16}$$

To recapitulate the main lesson of this section, infrared-normalized, color-ordered superamplitudes in the planar limit of $\mathcal{N} = 4$ SYM only depend on the particle number n , the helicity degree k , $3n - 15$ variables in the space of dual conformally invariant kinematics, and the order L of loops or perturbative corrections.

3. Cluster polylogarithmic functions

Having reviewed the parameters that normalized amplitudes $\mathcal{E}_{n,k}^{(L)}$ in planar $\mathcal{N} = 4$ SYM theory depend on, here we will continue to describe the type of functions they evaluate to. As we will recall in subsection 3.1, the latter fall in the general class of multiple polylogarithms, which are also relevant for a wide range of standard model processes at the forefront of precision phenomenology, especially when mediated by internal particles that can be considered as massless, see for example [22, 49, 71–73]. Within this class, however, there still exists an infinite number of these functions at each loop order, depending on where they are allowed to have singularities. In subsection 3.2 we will then see that beautiful mathematical objects known as cluster algebras appear to correctly predict these singularities at multiplicity $n = 6, 7$, thereby making the function spaces expected to contain $\mathcal{E}_{n,k}^{(L)}$ finite. Furthermore, in subsection 3.3 we will see that cluster algebras also dictate how these singularities are allowed to appear consecutively, and how these additional restrictions can be physically interpreted as the (extended) Steinmann relations of axiomatic quantum field theory. The finite function spaces that are further reduced by the latter restrictions will be the starting point for bootstrapping the corresponding amplitudes in the most efficient manner, to be discussed in the next section.

3.1. Multiple polylogarithms and symbols

All explicit calculations to date, as well as an analysis at the level of the integrand [74] (note however the subtleties pointed out in [75]), suggest that at least for $k = 0, 1$, $\mathcal{E}_{n,k}^{(L)}$ can be expressed in terms of generalized or Goncharov or *multiple polylogarithms* (MPLs) [76–78] of weight $m = 2L$. These functions are also mentioned in chapters 3 [79] and 4 [80] of the SAGEX review [42], and in addition they are discussed in great detail in the recent textbook [81]. Let us briefly collect here the definitions and properties that will be useful for our purposes.

A function F_m is defined as an MPL of weight m if its total differential obeys

$$dF_m = \sum_{\phi_\beta \in \Phi} F_{m-1}^{\phi_\beta} d \log \phi_\beta, \tag{17}$$

such that $F_{m-1}^{\phi_\alpha}$ is an MPL of weight $m - 1$,

$$dF_{m-1}^{\phi_\beta} = \sum_{\phi_\alpha \in \Phi} F_{m-2}^{\phi_\alpha, \phi_\beta} d \log \phi_\alpha, \tag{18}$$

and so on, with this recursive definition (17) terminating at $m = 1$ with the usual logarithms on the left-hand side, and rational numbers as coefficients of the total differentials on the right-hand side. The arguments of the dlogs ϕ_{α_i} are algebraic functions of the independent variables of F_m known as the (symbol) *letters*, and similarly their collection Φ from all steps of the recursion is called the (symbol) *alphabet*. Evidently, it encodes the positions of the possible branch point singularities of F_m , which may appear when $\phi_{\alpha_i} = 0, \infty$.

The iterative structure we have described forms part of the *coaction* operation Δ [82–85] (also loosely referred to as the *coproduct*), which ‘decomposes’ an MPL of weight m into linear combinations of pairs of MPLs with weight $\{m - m_1, m_1\}$ for $m_1 = 0, 1, \dots, m$. Concretely, the total differential (17) is essentially equivalent to the $\{m - 1, 1\}$ component of Δ ,

$$\Delta_{m-1,1}F_m = \sum_{\phi_\beta \in \Phi} F_{m-1}^{\phi_\beta} \otimes [\log \phi_\beta \bmod (i\pi)]. \tag{19}$$

The coaction may be repeatedly applied to either the first or the second factor of the $\{m - m_1, m_1\}$ pair when $m_1 > 1$, yielding a decomposition of an MPL of weight m into subspaces of MPLs with weight $\{m_1, \dots, m_r\}$, $\sum_{i=1}^r m_i = m$, that is unique thanks to the coassociativity property of the coaction. Denoting the projection of the coaction on each of these subspaces by Δ_{m_1, \dots, m_r} , for example the equations (17) and (18) combine to yield the $\{m - 2, 1, 1\}$ coproduct,

$$\Delta_{m-2,1,1}F_m = \sum_{\phi_\alpha, \phi_\beta \in \Phi} F_{m-2}^{\phi_\alpha, \phi_\beta} \otimes \log \phi_\alpha \otimes \log \phi_\beta, \tag{20}$$

where from equation (20) onwards, identification of $\log \phi$ factors up to $i\pi$ will be implied. Furthermore, the maximal iteration of the coaction defines the *symbol* [86, 87],

$$S[F_m] = \underbrace{\Delta_{1, \dots, 1}}_{m \text{ times}} F_m = \sum_{\phi_{\alpha_1}, \dots, \phi_{\alpha_m}} F_0^{\phi_{\alpha_1}, \dots, \phi_{\alpha_m}} [\log \phi_{\alpha_1} \otimes \dots \otimes \log \phi_{\alpha_m}], \tag{21}$$

where it is also customary to adopt a more compact notation by replacing $\log \phi_{\alpha_i} \rightarrow \phi_{\alpha_i}$. To make the above general definitions more tangible, we will also apply them to a concrete example towards the end of this subsection.

Comparing (17) and (19), we see that derivatives only act on the rightmost factor of the coaction, and the same carries over to the symbol. Similarly, the discontinuities of MPLs can be shown to be encoded in the leftmost factor of their coaction. For example, at the level of the symbol the discontinuity of F_m when going around a potential branch point $\phi_\beta = 0$ with no other letter vanishing simultaneously is given by

$$S[\text{Disc}_{\phi_\beta}(F_m)] = 2\pi i \sum_{\phi_{\alpha_1}, \dots, \phi_{\alpha_m}} F_0^{\phi_{\alpha_1}, \dots, \phi_{\alpha_m}} \delta_{\alpha_1\beta} [\phi_{\alpha_2} \otimes \dots \otimes \phi_{\alpha_m}], \tag{22}$$

in other words it is equivalent to clipping off the first entry.

We now move on to present some further definitions that we will only need when we perform the sample bootstrap computation of $\mathcal{E}_{6,0}^{(2)}$ in subsection 4.4. The reader only interested in conceptual aspects may thus choose to skip to the next subsection. Alternatively to the differential definition of MPLs presented above, one may also reverse the direction and define them as iterated integrals. Choosing the integration contour in the simplest possible manner, leads to the (*G*-function) definition

$$G(a_1, \dots, a_m; z) = \int_0^z \frac{dt}{t - a_1} G(a_2, \dots, a_m; t), \tag{23}$$

where the recursion starts with $G(; z) = 1$, and for the special case where all the a_i are zero, we define

$$G(\underbrace{0, \dots, 0}_{m \text{ times}}; z) = \frac{1}{m!} \log^m z. \tag{24}$$

Indeed, by differentiating equation (23), applying the identity

$$\frac{\partial}{\partial a} \frac{1}{t - a} = - \frac{\partial}{\partial t} \frac{1}{t - a} \tag{25}$$

and partial fractioning, it can be shown that dG takes the general form of equation (17).

In the integral definition (23), we see that the weight m corresponds to the number of iterated integrations. The single-variable case with $a_i \in \{-1, 0, 1\}$ has also independently appeared in the physics literature under the name of *harmonic polylogarithms* (HPLs) [88], up to the different sign convention,

$$H(a_1, \dots, a_m; z) = (-1)^p G(a_1, \dots, a_m; z), \quad a_i \in \{-1, 0, 1\}, \tag{26}$$

where p counts how many a_i are equal to $+1$. Since HPLs only depend on the outermost integration bound z in equation (23), their differentiation is trivial, and it is easy to show that the definitions (17)–(21) specialize to

$$\Delta_{m-1,1} H(a_1, \dots, a_m; z) = (-1)^{\text{sgn}(a_1)} H(a_2, \dots, a_m; z) \otimes (z - a_1), \tag{27}$$

$$\Delta_{m-2,1,1} H(a_1, \dots, a_m; z) = (-1)^{\text{sgn}(a_1) + \text{sgn}(a_2)} H(a_3, \dots, a_m; z) \otimes (z - a_2) \otimes (z - a_1),$$

⋮

$$S[H(a_1, \dots, a_m; z)] = (-1)^p [(z - a_m) \otimes \dots \otimes (z - a_1)]. \tag{28}$$

For later convenience let us also note that a more compact notation for all a_i arguments can be adopted, whereby a string of subsequent zeros is replaced by

$$\underbrace{0, 0, \dots, 0}_{m-1 \text{ times}}, \pm 1 \rightarrow \pm m, \tag{29}$$

and the resulting, shorter string of arguments is placed as indices of the function. For example, classical logarithms, that are contained in HPLs, in this notation correspond to

$$\text{Li}_m(z) = H_m(z) = H(\underbrace{0, \dots, 0}_{m-1 \text{ times}}, 1; z) \tag{30}$$

and again for completeness their symbol will be (recalling that $d \log(z - 1) = d \log(1 - z)$)

$$S[\text{Li}_m(z)] = - \left[(1 - z) \otimes \underbrace{z \otimes \dots \otimes z}_{m-1 \text{ times}} \right]. \tag{31}$$

Finally, let us mention that products of G -functions with the same rightmost argument, and thus also HPLs, may be reexpressed as linear combinations thereof. Namely they form a

(shuffle) algebra, as can be simply inferred from the definition (23), by appropriately splitting the integration range so that all dummy variables have a specific order.

In practice, by now there exist a variety of software tools that allow the evaluation of the functions presented in this subsection, as well as the application of their transformation properties, either in free computer algebra systems such as GiNaC [89], or in proprietary ones such as Mathematica and Maple, by SAGEX industry partners Wolfram Research and Maplesoft, respectively. These for example include the Mathematica packages HPL [90] and PolyLogTools [91], as well as the native functionality of GiNaC (Maple [92]) with respect to the numerical evaluation (and symbolic manipulation) of these functions.

3.2. Cluster algebras and amplitude singularities

Once we have identified MPLs as the general class of functions that contain the amplitude, the next step is to clarify what the corresponding symbol alphabet is. For $n = 6$, this came as a result of an explicit Feynman diagram computation [93, 94] at two loops, and was further supported by the analysis of closely related integrals [95, 96]. For general n , strong motivation is provided by the cluster algebra structure [63] of the space of kinematics.

More precisely, in section 2 we mentioned that the space of kinematics in terms of momentum twistors can be realized as the quotient $\text{Gr}(4, n)/(\mathbb{C}^*)^{n-1}$ of a Grassmannian. It is the latter space that is naturally endowed with a cluster algebra structure [97], thus making it sensible to explore any implications this may have on the symbol alphabet. Before we spell that out, we will begin with a brief introduction on cluster algebras [98–101], which have become a very active research area in contemporary mathematics since their inception in early 2000s. It is also worth noting that they have already found applications in other areas of mathematical physics in the past, such as the proof of periodicity of Y -systems and associated thermodynamic Bethe ansätze of certain integrable models [102], or the determination of BPS state spectra in supersymmetric field theories [103–105]. In the realm of scattering amplitudes, their role was first appreciated at the level of the integrand, which also exhibits Grassmannian structure [64]. For more recent work relating cluster algebras and tree-level amplitudes or loop integrands, see also [106, 107].

3.2.1. Basics of cluster algebras. With many excellent introductory articles on cluster algebras available in the literature, as well as articles with detailed review sections on their connection to scattering amplitudes [108, 109], here will simply aim to highlight some of their features while mostly following one concrete example that will be relevant later on.

The building blocks of cluster algebras are certain variables a_i , known as (Fomin–Zelevinsky) cluster \mathcal{A} -coordinates, that are grouped into overlapping subsets $\{a_1, \dots, a_d\}$ of rank d , the *clusters*. Starting from an initial cluster, cluster algebras are constructively defined by a *mutation* operation on the \mathcal{A} -coordinates. They can also be generalized so as to contain *frozen variables* or *coefficients* $\{a_{d+1}, \dots, a_{d+m}\}$, whose main difference from the \mathcal{A} -coordinates is that they do not mutate.

In the simplest case, which is also sufficient for $\mathcal{N} = 4$ SYM amplitudes, cluster algebras can be described by directed graphs or quivers. On the left-hand side of figure 1, the initial cluster of the $\text{Gr}(4, 6)$ cluster algebra, relevant for six-particle scattering, is shown. Unboxed and boxed vertices of the quiver denote the \mathcal{A} -coordinates and frozen variables, respectively, and we observe that they all correspond to the four-brackets (or equivalently Plücker coordinates) defined in subsection 2.3, namely 4×4 minors of the $4 \times n$ matrix realization of $\text{Gr}(4, n)$, here for $n = 6$. The observant reader may however notice two differences between frozen and

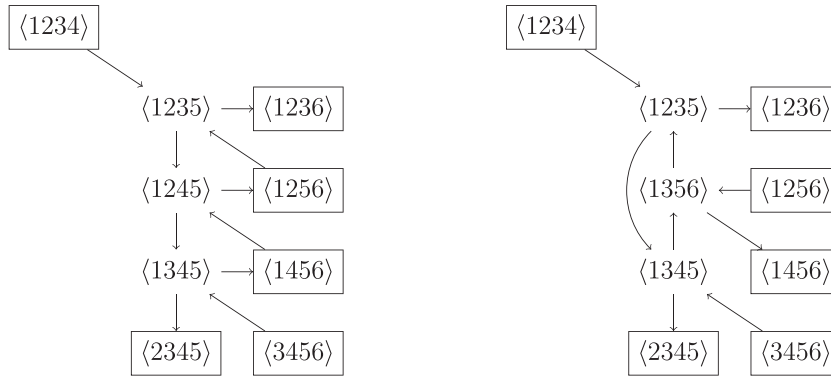


Figure 1. Left: the quiver diagram for the $\text{Gr}(4,6)$ initial cluster. Right: the quiver that arises by mutating $\langle 1245 \rangle$ of the initial cluster, where the effect of the mutation is described in equations (32) and (33) for the variable, and below equation (34) for the arrows of the quiver. Reproduced from [38]. © Copyright owned by the author(s) under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0).

cluster variables, which turn out to hold more generally: first, that the former always have consecutive indices $\langle i, i + 1, i + 2, i + 3 \rangle$, modulo $n + i \sim i$ identifications, whereas this is not the case for the latter. And second, that the former are not allowed to have arrows between them.

The arrows of the quiver encode how its \mathcal{A} -coordinates will transform under mutation. Concretely, if a_k is a cluster \mathcal{A} -coordinate, mutating it replaces it by

$$a_k \rightarrow a'_k = \frac{1}{a_k} \left(\prod_{\text{arrows } i \rightarrow k} a_i + \prod_{\text{arrows } k \rightarrow j} a_j \right). \tag{32}$$

Let us see this in action in the $\text{Gr}(4,6)$ initial cluster, but for reasons that will become apparent very shortly, let us first switch to shorthand notation where each four-bracket is expressed in terms of the complement of the twistor labels it contains, for example $\langle 1235 \rangle = (46)$, $\langle 1345 \rangle = (26)$, with the sign convention chosen such that increasing order of labels on the left is mapped to increasing order of labels on the right (this is an instance of the more general $\text{Gr}(k, n) \simeq \text{Gr}(n - k, n)$ duality exchanging k - and $(n - k)$ -planes). Then, the mutation of $\langle 1245 \rangle$ on the left-hand side of figure 1 yields

$$\langle 1245 \rangle = (36) \rightarrow \frac{(46)(23) + (26)(34)}{(36)} = (24), \tag{33}$$

where the shorthand notation allowed us to arrive at the last equality by using the familiar also in other contexts three-term Plücker relation or $SL(2, \mathbb{C})$ Schouten identity,

$$(ik)(jl) = (ij)(kl) + (il)(jk), \tag{34}$$

for $i = 2, j = 3, k = 4$ and $l = 6$.

Except for $a_k \rightarrow a'_k$, in the new quiver produced by this mutation, all the rest of the \mathcal{A} -coordinates and coefficients remain unchanged. However, the arrows of the new quiver will

differ, and may be obtained by those of the quiver before the mutation of a_k by applying the following rules³:

- For each path $i \rightarrow k \rightarrow j$ add an arrow $i \rightarrow j$, except if both i and j are frozen variables.
- Reverse the direction of all arrows pointing to or originating from k .
- Remove any pairs of arrows pointing in opposite directions, \rightleftarrows .

Going back to our example, we see that by virtue of these rules the mutation of $\langle 1245 \rangle$ in the $\text{Gr}(4, 6)$ initial cluster leads to the new cluster shown on the right-hand side of figure 1.

We have thus specified all the rules of the game, and obtaining the entire cluster algebra is a matter of applying them over and over at each vertex of every quiver we encounter. While the graphical representation and rules we described so far are more accessible for a first exposure to cluster algebras, to this end it proves more efficient to exploit the fact that every quiver is in bijection with a skew-symmetric exchange matrix B with elements

$$b_{ij} = (\# \text{ arrows } i \rightarrow j) - (\# \text{ arrows } j \rightarrow i). \tag{35}$$

In this manner, e.g. the exchange matrices B, B' associated to the left- and right-hand side quivers of figure 1, respectively, have nonzero elements with $i < j$ that are equal to

$$\begin{aligned} b_{12} = b_{15} = b_{23} = b_{26} = b_{37} = b_{39} = -b_{14} = -b_{16} = -b_{27} = -b_{38} = 1, \\ b'_{13} = b'_{15} = b'_{27} = b'_{39} = -b'_{12} = -b'_{14} = -b'_{23} = -b'_{26} = -b'_{38} = -b'_{67} = 1, \end{aligned} \tag{36}$$

when ordering our \mathcal{A} -coordinates and frozen variables as $\{\langle 1235 \rangle, \langle 1245 \rangle, \langle 1345 \rangle, \langle 1234 \rangle, \langle 1236 \rangle, \langle 1256 \rangle, \langle 1456 \rangle, \langle 3456 \rangle, \langle 2345 \rangle\}$ for B , and similarly with $\langle 1245 \rangle \rightarrow \langle 1356 \rangle$ for B' . With this rearrangement of information, it can be shown that the \mathcal{A} -coordinate mutation (32) becomes

$$a'_k = a_k^{-1} \left(\prod_{i=1}^{d+m} a_i^{[b_{ik}]_+} + \prod_{i=1}^{d+m} a_i^{[-b_{ik}]_+} \right), \tag{37}$$

with $[x]_+ = \max(0, x)$. Similarly, the rules we discussed below equation (34) for the transformation of the quiver translate into the following mutation rule for the exchange matrix,

$$b'_{ij} = \begin{cases} -b_{ij}, & \text{if } i = k \text{ or } j = k, \\ b_{ij} + [-b_{ik}]_+ b_{kj} + b_{ik} [b_{kj}]_+, & \text{otherwise,} \end{cases} \tag{38}$$

as can be readily verified in the example of equation (36).

Apart from the fact that the alternative definitions (37) and (38) are more amenable to computer implementation, they can be also generalized so as to describe cluster algebras with skew-symmetrizable instead of skew-symmetric exchange matrices (more precisely, their principal part with indices $i, j \leq d$)⁴. In this more general setting, it is possible to prove that finite cluster algebras are classified by Dynkin diagrams, and that a skew-symmetric cluster algebra is finite if and only if one of its clusters takes the form of the associated Dynkin diagram

³ It is very interesting to note that essentially the same quiver (but not cluster variable) mutation rules were independently proposed in the context of $\mathcal{N} = 1$ quiver gauge theories so as to describe their Seiberg duality [110], which generalizes the usual electric-magnetic duality of abelian gauge theory.

⁴ Alternatively, skew-symmetrizable cluster algebras may be defined by generalizing quivers to so-called valued quivers [111].



Figure 2. Left: geometric interpretation of the $\text{Gr}(4, 6) \simeq \text{Gr}(2, 6)$ cluster shown in the left of figure 1, where (1234) corresponds to the edge (56) etc, as a triangulation of a hexagon by non-crossing diagonals. Right: mutations of this cluster algebra may be similarly geometrically interpreted as flips $(ik) \leftrightarrow (jl)$ of the diagonals of any quadrilateral subdiagram, see also equation (34).

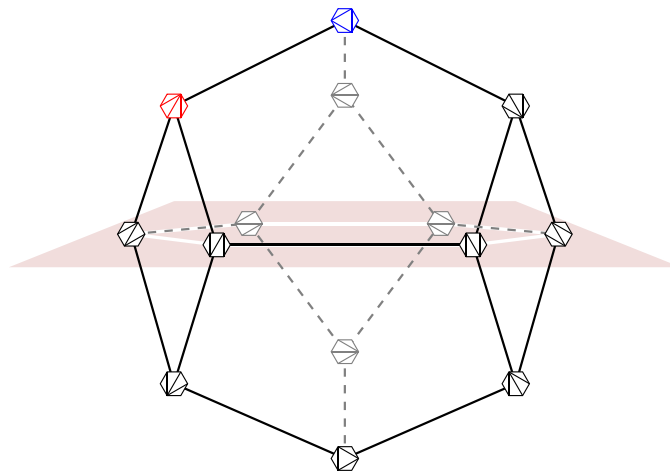


Figure 3. The cluster polytope of the $\text{Gr}(4, 6) \simeq A_3$ cluster algebra, with clusters representing the different triangulations of a hexagon, as discussed in the text and in figure 2, where also the vertex labels may be found. The initial and mutated cluster in the left- and right-hand side of figure 1 are color-coded in red and blue, respectively. The parity invariant plane is also drawn in pink. Reproduced from [109]. CC BY 4.0.

when dropping arrow orientations as well as frozen variables and arrows from/to them [99]. Inspecting figure 1, we can thereby infer that the $\text{Gr}(4, 6)$ cluster algebra is of finite A_3 type.

Despite the relative simplicity of the definitions and properties of cluster algebras, the reader may perhaps be left wondering where they come from, if there is any physical or mathematical intuition behind them. To address this, drawing from [98] let us return to our $\text{Gr}(4, 6)$ example, and note that the labels of the dual two-component brackets (ij) can be interpreted as the vertices of a hexagon, such that the frozen variables correspond to its edges, whereas the \mathcal{A} coordinates of the clusters we have encountered so far to non-crossing diagonals. This observation in fact extends to the entire cluster algebra, since it can be shown that *all* mutations have the form of the three-term identity (34) with $1 \leq i < j < k < l \leq 6$, and hence they are geometrically equivalent to flipping the diagonal of a quadrilateral inside the hexagon, see

figure 2. Therefore the $\text{Gr}(4, 6)$ cluster algebra is in natural bijection with all triangulations of the hexagon with non-crossing diagonals. Note that this bijection also includes the exchange matrix B in equation (35), or in other words the arrows of the quivers as seen e.g. in figure 1, which become arrows between adjacent sides of any triangle of the triangulation (not allowing arrows between two edges of the hexagon), with the orientation chosen in the anti-clockwise direction. This geometric picture of cluster algebras as triangulations is in fact a very profound and universal one, as up to 18 exceptional cases, it has been shown to hold for all cluster algebras with a finite number of exchange matrices, even if they have an infinite number of clusters/variables [112].

Before concluding this introduction on cluster algebras, let us mention one further connection they have with geometry, that will be useful for us in what follows: representing each cluster with a vertex, and each mutation with an edge yields the *exchange graph* of a finite rank- d cluster algebra, which in fact defines a *simple polytope* [102], namely a geometric object with flat faces generalizing the polygon to higher dimensions, whose vertices are in addition adjacent to exactly d edges. As an example, the cluster polytope of the $\text{Gr}(4, 6) \simeq A_3$ cluster algebra is shown in figure 3. The bijection with triangulations discussed in the previous paragraph allows one to easily work this out, and infer that it has a total of nine diagonals/cluster variables spread into 14 vertices/clusters, as well as 21 edges and nine faces. In addition to its topological and combinatorial nature, this polytope also geometrically describes the compactification of the *positive region* of $\text{Gr}(4, n)/(\mathbb{C}^*)^{n-1}$, defined as the region where $\langle ijkl \rangle > 0$ for $i < j < k < l$. In particular, each edge of the polytope can be assigned a (Fock–Goncharov) \mathcal{X} -coordinate, related to the \mathcal{A} -coordinates by [113]

$$x_i \equiv \prod_{l=1}^{d+m} a_l^{b_{li}}, \quad i = 1, \dots, d, \tag{39}$$

such that each cluster provides a local coordinate chart describing this compactification, with the interior of the positive region corresponding to all $\infty > x_i > 0$. The significance of the positive region will also be highlighted in chapter 7 [114] of the SAGEX review [42].

3.2.2. Symbol letters from cluster variables. With this background knowledge on cluster algebras, we can now state their role in $\mathcal{N} = 4$ SYM amplitudes: in [63] the remarkable observation was made, based on the then-known explicit computations, that

$\text{Gr}(4, n)$ cluster \mathcal{A} -coordinates appear as symbol letters of the n -particle amplitude.

That is, for the examples considered, the amplitude could be expressed in terms of polylogarithmic functions as defined in equation (17), with ϕ_{α_i} coinciding with \mathcal{A} -coordinates (more precisely multiplicative combinations thereof, also including the frozen variables, that respect the scale and hence the dual conformal invariance of the theory).

For the six-particle amplitude, as we have seen the associated $\text{Gr}(4, 6)$ cluster algebra consists of nine \mathcal{A} -coordinates, and these precisely coincide with the symbol alphabet of the two-loop correction to the amplitude [87, 93, 94]. The six-particle or *hexagon bootstrap* was initiated in [2] based on the assumption that this alphabet also remains stable at higher loops, and in this case its cluster algebraic structure may be considered as evidence backing this assumption. Explicitly, in the usual four-bracket notation a convenient choice for the six-particle alphabet

reads [21]⁵,

$$a_1 = \frac{\langle 1245 \rangle^2 \langle 3456 \rangle^2 \langle 6123 \rangle^2}{\langle 1234 \rangle \langle 2345 \rangle \cdots \langle 6123 \rangle}, \quad m_1 = \frac{\langle 1356 \rangle \langle 2346 \rangle}{\langle 1236 \rangle \langle 3456 \rangle}, \quad y_1 = \frac{\langle 1345 \rangle \langle 2456 \rangle \langle 1236 \rangle}{\langle 1235 \rangle \langle 1246 \rangle \langle 3456 \rangle}, \tag{40}$$

together with two more cyclic transformations $l_1 \rightarrow l_{1+i}$ with $l \in \{a, m, y\}$ induced by shifting $Z_m \rightarrow Z_{m-2i}$ on the right-hand side. The discrete parity and flip transformations of the letters may be inferred from the cluster polytope of figure 3, where they correspond to up-down and left-right reflection, respectively. For example, parity is equivalent to a $i \rightarrow i + 3$ shift of momentum twistor labels, and thus transforms $y_i \rightarrow 1/y_i$, while leaving a_i, m_i invariant.

The initial quiver for generic $\text{Gr}(4, n)$, from which the next nontrivial $n = 7$ case may be studied with the same set of rules we spelled out, is depicted in figure 4. While for the latter case the initial quiver does not have the topology of a Dynkin diagram, mutating $\langle 1256 \rangle, \langle 1456 \rangle$ and $\langle 1345 \rangle$ does lead to an E_6 -shaped cluster, and hence $\text{Gr}(4, 7)$ is also a finite cluster algebra. In particular, one finds 42 different \mathcal{A} -coordinates distributed in 833 distinct clusters (the order of the variables in each cluster does not matter). Again, these \mathcal{A} -coordinates exactly match the symbol alphabet of the two-loop correction to the seven-particle amplitude [115], and this observation was of central importance for generalizing the bootstrap program to higher multiplicity $n = 7$ in [16]. The choice for the corresponding symbol alphabet adopted in the latter reference is

$$\begin{aligned} a_{11} &= \frac{\langle 1234 \rangle \langle 1567 \rangle \langle 2367 \rangle}{\langle 1237 \rangle \langle 1267 \rangle \langle 3456 \rangle}, & a_{41} &= \frac{\langle 2457 \rangle \langle 3456 \rangle}{\langle 2345 \rangle \langle 4567 \rangle}, \\ a_{21} &= \frac{\langle 1234 \rangle \langle 2567 \rangle}{\langle 1267 \rangle \langle 2345 \rangle}, & a_{51} &= \frac{\langle 1(23)(45)(67) \rangle}{\langle 1234 \rangle \langle 1567 \rangle}, \\ a_{31} &= \frac{\langle 1567 \rangle \langle 2347 \rangle}{\langle 1237 \rangle \langle 4567 \rangle}, & a_{61} &= \frac{\langle 1(34)(56)(72) \rangle}{\langle 1234 \rangle \langle 1567 \rangle}, \end{aligned} \tag{41}$$

where we have defined

$$\langle a(bc)(de)(fg) \rangle \equiv \langle abde \rangle \langle acfg \rangle - \langle abfg \rangle \langle acde \rangle, \tag{42}$$

together with the letters a_{ij} obtained from a_{i1} by cyclically relabeling the momentum twistors $Z_m \rightarrow Z_{m+j-1}$. It is interesting to note that for $n = 7$, and more generally when n is odd, any individual cluster \mathcal{A} -coordinate can be rendered invariant under rescalings $Z_i \rightarrow tZ_i$ by suitable products of powers of frozen variables. Here it is slightly more nontrivial to show that parity transformations map $a_{2i} \leftrightarrow a_{3,i-1}$ and $a_{4i} \leftrightarrow a_{5i}$.

The appearance of not only Plücker variables but also homogeneous polynomials thereof (42) as \mathcal{A} -coordinates is a qualitatively new feature that persists for $\text{Gr}(4, n)$ cluster algebras with $n \geq 8$. However the relation of these cluster algebras with n -particle alphabets is a significantly more subtle issue which we will address in subsection 5.1.

While not of direct relevance for this article, before closing let us briefly mention two more connections that have been established between cluster algebras and scattering amplitudes. The

⁵ Note that any set of equal size, consisting of multiplicatively independent combinations of these letters, would make an equally valid choice. Indeed, in the original literature [2, 9–12] the letters u_i and $1 - u_i$ were used. The relation with the presently used alphabet is $a_i = u_i/(u_{i-1}u_{i+1})$ and $m_i = 1 - 1/u_i$.

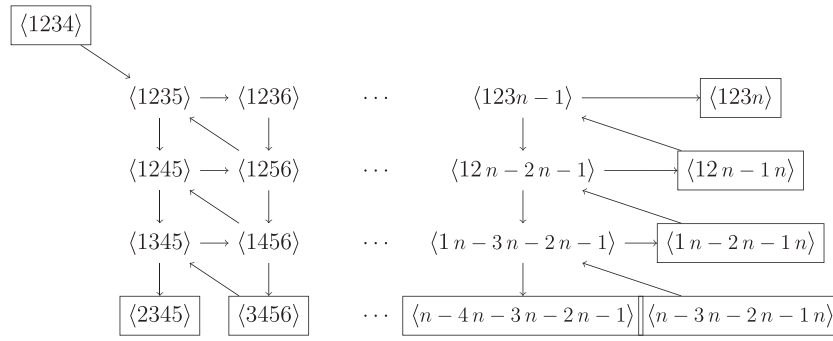


Figure 4. Quiver diagram for the initial $\text{Gr}(4, n)$ cluster. Reproduced with permission from [38]. © Copyright owned by the author(s) under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0).

first one also pertains to planar n -particle amplitudes in $\mathcal{N} = 4$ SYM, and focuses on identifying appropriate cluster \mathcal{X} -coordinates as arguments of the MPLs needed to describe them [116–119], based on their Poisson structure [120]. This approach has been used very successfully in promoting symbols of two-loop MHV amplitudes [115] to functions [121–123], yet there is evidence that this is no longer possible at different MHV degree [124], and furthermore it is unclear how to generalize to higher loops. Finally, cluster algebras may also be used in order to define natural generalizations of string amplitudes [125–127].

3.3. Cluster adjacency and extended Steinmann relations

So far we have seen that cluster variables dictate the singularities of $\mathcal{N} = 4$ SYM amplitudes. However cluster algebras have more structure than just the variables, for instance the clusters. It is thus natural to ask, do they also play a role in this context? Very interestingly, there is evidence that they do, in the form of [20]

Cluster adjacency : in a symbol whose alphabet contains $\text{Gr}(4, n)$ cluster \mathcal{A} -coordinates, two of them can appear consecutively only if there exists a cluster where they both appear.

Let us distill the implications of cluster adjacency⁶ in our familiar $\text{Gr}(4, 6)$ cluster algebra example. From our discussion of the geometric interpretation of the clusters as triangulations of a hexagon with non-crossing diagonals, it is evident that pairs of \mathcal{A} -coordinates not found in a cluster together will in turn correspond to crossing diagonals, and are thus forbidden from appearing next to each other in the symbol of the six-particle amplitude, as shown in figure 5.

The same information may be alternatively described by a *neighbor set* [109] of an \mathcal{A} -coordinate, that is the union of all clusters containing it, which in other words contains all the other variables (including the frozen ones), that can appear next to it in the symbol. For the $n = 6$ case, reverting to the usual four-bracket notation these are

$$\langle 1245 \rangle = \{ \langle 1245 \rangle, \langle 2456 \rangle, \langle 1345 \rangle, \langle 1246 \rangle, \langle 1235 \rangle, \text{ \& frozen variables} \}, \tag{43}$$

⁶ Note that in the mathematics literature there exists a related notion of simultaneous inclusion in a cluster, which is known as ‘compatibility’ [102].

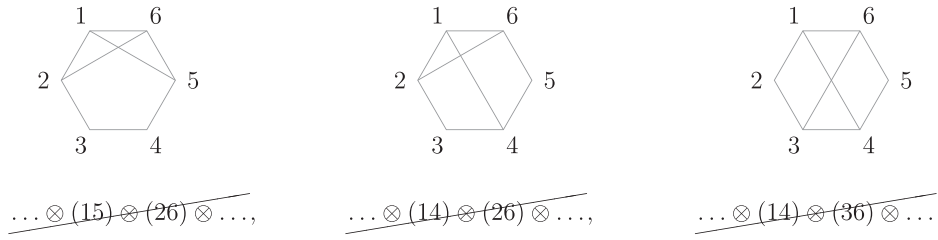


Figure 5. Pairs of $\text{Gr}(4, 6)$ \mathcal{A} -coordinates not found in a cluster together, and hence not allowed to appear consecutively in the symbol of the six-particle amplitude, up to cyclic permutations and order reversal.

$$\langle 1235 \rangle = \{ \langle 1235 \rangle, \langle 2456 \rangle, \langle 2356 \rangle, \langle 1356 \rangle, \langle 1345 \rangle, \langle 1245 \rangle, \text{ \& frozen variables} \}, \quad (44)$$

as well as their cyclic permutations, two for the first line and five for the second. Then, as a constraint on a polylogarithmic function F , and in the notation of equation (20), cluster adjacency may be formulated as

$$\overline{F^{\phi_\beta, \phi_\beta}} = 0, \quad (45)$$

where $\overline{\phi_\beta}$ does not belong to the neighbor set of ϕ_β , with the same condition also holding recursively for all left factors in the coproduct of F .

As we have discussed, actual symbol letters are dual conformal invariant, namely products of the cluster and frozen variables that are invariant under rescalings of the twistors, i.e. homogeneous. It is thus more convenient to take this information into account on the right-hand side of equations (43) and (44), by defining the corresponding *homogeneous neighbor sets*,

$$\text{hns}[\langle 1245 \rangle] = \text{hns}[a_1] = \{ a_1, m_2, m_3, y_1, y_2 y_3 \}, \quad (46)$$

$$\text{hns}[\langle 1235 \rangle] = \left\{ a_1, a_2, \frac{m_1}{y_2}, \frac{m_2}{y_2 y_3}, m_3, y_1 y_2 y_3 \right\}, \quad (47)$$

again plus cyclic permutations. In the first line we could also promote the left-hand side to the conformally invariant letter a_1 , since $\langle 1245 \rangle$ is the only cluster variable it depends on. This is not possible for the second line, relevant for the remaining letters m_i, y_i .

Moving on to the case $n = 7$, all letters (42) depend on a single \mathcal{A} -coordinate, so we can directly focus on the homogeneous neighbor sets. These are generated by

$$\text{hns}[a_{11}] = \{ a_{11}, a_{14}, a_{15}, a_{21}, a_{22}, a_{24}, a_{25}, a_{26}, a_{31}, a_{33}, a_{34}, a_{35}, a_{37}, a_{41}, a_{43}, a_{46}, a_{51}, a_{53}, a_{56}, a_{62}, a_{67} \} \quad (48)$$

$$\text{hns}[a_{21}] = \{ a_{11}, a_{13}, a_{14}, a_{15}, a_{17}, a_{21}, a_{23}, a_{24}, a_{25}, a_{26}, a_{31}, a_{33}, a_{34}, a_{36}, a_{37}, a_{41}, a_{43}, a_{45}, a_{46}, a_{52}, a_{53}, a_{55}, a_{57}, a_{62}, a_{64}, a_{66} \} \quad (49)$$

$$\text{hns}[a_{41}] = \{ a_{11}, a_{13}, a_{16}, a_{21}, a_{23}, a_{24}, a_{26}, a_{31}, a_{33}, a_{35}, a_{36}, a_{41}, a_{43}, a_{46}, a_{51}, a_{62}, a_{67} \} \quad (50)$$

$$\text{hns}[a_{61}] = \{ a_{12}, a_{17}, a_{23}, a_{25}, a_{27}, a_{32}, a_{34}, a_{36}, a_{42}, a_{47}, a_{52}, a_{57}, a_{61} \}, \quad (51)$$

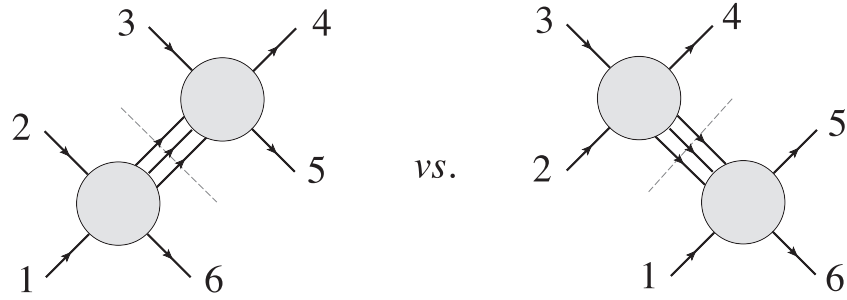


Figure 6. Channels $s_{345} \propto \langle 2356 \rangle = (14)$ and $s_{234} \propto \langle 1245 \rangle = (36)$ for $3 \rightarrow 3$ kinematics. The Steinmann relations state that the discontinuity in one channel should not know about the discontinuity in the other channel, and this yields the same constraint as the non-cluster adjacent pair on the right of figure 5. Reprinted (figure) with permission from [14], Copyright (2016) by the American Physical Society.

together with images under parity transformations and cyclic permutations. We will also comment on higher-multiplicity generalizations towards the end of this subsection.

What about the physical interpretation of cluster adjacency? It turns out that some of its restrictions can be understood as the extended Steinmann relations [21, 31], as we will now explain. These are generalizations of the usual Steinmann relations [128–130], which demand that the double discontinuities of any Feynman diagram (and thus of the amplitude they contribute to) vanish when taken in overlapping channels. In section 3.1 we have seen that a discontinuity may be labeled by a Mandelstam invariant $s_{i,\dots,j-1}$ which is analytically continued around its branch point. At the same time, by virtue of the Cutkosky rules [131] this discontinuity may be obtained by placing on-shell the internal particles whose total energy equals $s_{i,\dots,j-1}$, that is by replacing their propagators with delta functions. This is the notion of a cut, which splits the Feynman diagram into two parts, as seen in figure 6.

By the same logic, overlapping channels correspond to cut lines that intersect, in other words they divide the external particles of the Feynman diagram into four non-empty sets. In the example of the figure, these sets are $\{2\}$, $\{3, 4\}$, $\{5\}$, and $\{6, 1\}$. Focusing on three-particle Mandelstam invariants, but allowing the number of external particles n to be arbitrary, the Steinmann relations at the level of the amplitude then imply

$$\text{Disc}_{s_{j,j+1,j+2}}(\text{Disc}_{s_{i,i+1,i+2}}(A_{n,k})) = 0, \quad \text{for } j = i \pm 1, i \pm 2, \tag{52}$$

with an obvious generalization to higher-particle Mandelstam invariants. Note that we refrain from considering two-particle invariants, since it is necessary for an invariant to be independent for the sake of analytic continuation. That is, no other Mandelstam invariant is allowed to change sign but the one we analytically continue, and this is generically not the case with two-particle invariants.

As equation (22) reveals, at the level of the symbol a discontinuity around $\phi_\beta = 0$ amounts to clipping off this particular letter from its first entry, with iterated discontinuities obtained by applying this procedure repeatedly. That is, if no other letter vanishes simultaneously, which can be ensured for rational alphabets such as the $n = 6, 7$ ones, due to the multiplicative independence of the letters. In this case, the Steinmann relations are therefore statements about which letters can appear next to each other in the first two entries of the symbol, and in the past they have been exploited in the amplitude bootstrap so as to simplify the construction of the function spaces containing the amplitude [14, 17]. Particularly for $n = 6$, we then notice

that the unique, up to cyclic permutations, double discontinuity shown in figure 6 precisely coincides with the non-cluster adjacent pair restriction shown on the right of figure 5.

Then, the analysis of a wealth of data obtained by the amplitude bootstrap, independently revealed that these restrictions on consecutive pairs of symbol letters apply not only in the first two slots, but to *all* depths in the symbol, and they were thereby coined the *extended Steinmann relations*⁷ [21, 31]. Given that in the $n = 6, 7$ alphabets (40) and (41) only a_i and a_{1i} are proportional to three-particle invariants, letting F denote the appropriately normalized amplitudes, or any of their coproducts, or any finite integral within the space of MPLs with these alphabets, in this case they may thus concretely be expressed as⁸

extended steinmann relations:
$$\begin{cases} F^{a_i, a_{i+1}} = 0, & 1 \leq i \leq 3, & \text{for } n = 6 \\ F^{a_{1i}, a_{1i+\delta}} = 0, & \delta = 1, 2, & 1 \leq i \leq 7, \text{ for } n = 7. \end{cases} \quad (53)$$

As physical quantities in perturbative quantum field theory are multivalued functions, whose different branches represent different kinematic regions, it is plausible that the extended Steinmann relations follow from the validity of the usual Steinmann relations in all branches. That is, since moving from one branch to another involves shifting functions by their discontinuities, and since this operation at the level of the symbol amounts to removing first entries, a condition between any pair of adjacent entries could be converted to the same one between the first two entries. In this manner, the bootstrap may point, for the first time, to a more general property of quantum field theory, at least in the planar limit. Indeed, the extended Steinmann relations have been confirmed to hold also for all planar two-loop five-point master integrals [22], though not for a family of non-planar integrals with the same external kinematics [73].

Also in the form (53), it is clear that the extended Steinmann relations are contained in the cluster adjacency constraints (45), (46) and (48). What is less obvious but also true, is that for well-defined functions with physical branch cuts built out of the $n = 6, 7$ alphabets (40) and (41), the extended Steinmann relations also automatically imply all remaining cluster adjacency restrictions. In other words there exists a quite nontrivial equivalence between the former and the latter with respect to the transcendental part of amplitudes. On the other hand, cluster adjacency also has important implications for the rational parts of amplitudes [109, 133–136], as well as how these correlate with the transcendental parts, and this additional information has been very useful for bootstrapping $A_{7,1}^{(4)}$ [18]. Note that the relation between symbol letters and rational parts of amplitudes is not only confined to $\mathcal{N} = 4$ SYM, as is has also been observed in five-gluon amplitudes in QCD [137].

Apart from a mathematical curiosity or a formal physical property, it is also reasonable to ask what the practical significance of cluster adjacency/extended Steinmann relations is. The answer to this is that they drastically reduce the size of the function space containing the amplitude, thus making it far simpler to construct the former, and uniquely identify the latter. By now they have thus been incorporated in the construction of this function space, which we will detail in the next section, but just to illustrate their power, in tables 1 and 2, we provide the function space dimension as the weight increases, in comparison with the dimensions of the more redundant spaces used at earlier stages of the bootstrap.

⁷ These results for $n = 6$ were initially reported at *Amplitudes 2017*, in a talk by the author [132].

⁸ Imposing these equations in combination with the integrability conditions, to be discussed in the next section, automatically implies that they hold also with the order of letters reversed.

Table 1. Dimensions of spaces of integrable symbols containing the six-particle amplitude, refined as one moves from top to bottom by successively imposing the analytic properties indicated on the left-hand side. Reproduced with permission from [38]. © Copyright owned by the author(s) under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0).

Weight n	0	1	2	3	4	5	6	7	8	9	10	11	12	13
First entry	1	3	9	26	75	218	643	1929	5897	?	?	?	?	?
Steinmann	1	3	6	13	29	63	134	277	562	1117	2192	4263	8240	?
Ext. Stein.	1	3	6	13	26	51	98	184	340	613	1085	1887	3224	5431

Table 2. Dimensions of spaces of integrable symbols containing the seven-particle amplitude, refined as one moves from top to bottom by successively imposing the analytic properties indicated on the left-hand side. Reproduced with permission from [38]. © Copyright owned by the author(s) under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0).

Weight n	0	1	2	3	4	5	6	7
First entry	1	7	42	237	1288	6763	?	?
Steinmann	1	7	28	97	322	1030	3192	9570
Ext. Stein.	1	7	28	97	308	911	2555	6826

Finally, let us comment on the status of cluster adjacency and extended Steinmann relations at multiplicity $n \geq 8$. While cluster adjacency has been confirmed for all MHV amplitudes at $L = 1, 2$ with the help of the Sklyanin bracket [138], as we will see in subsection 5.1 in general the symbol letters do not quite coincide with the cluster variables for $n \geq 8$, and this creates several subtleties. For example a naive application of cluster adjacency appears to be violated in certain integrals [139]. As understood also in the latter reference, on this front it is the extended Steinmann relations that are on a firmer footing when expressed as a statement about multiple discontinuities generalizing equation (52), however in this case this does not simply translate into a statement about adjacent pairs in the symbol, analogous to equation (53).

4. The Steinmann cluster bootstrap for $\mathcal{N} = 4$ SYM amplitudes

In the previous section, we described the essential characteristics of the cluster polylogarithmic function spaces containing six- and seven-particle amplitudes in $\mathcal{N} = 4$ SYM theory. Here, we will start subsection 4.1 by presenting certain additional analytic properties these spaces are endowed with, and in subsection 4.2 we will explain how to construct them iteratively in the weight. Subsection 4.3 will then focus on how the amplitude may be singled out from this space with the help of independent information on its behavior in certain kinematic limits. Finally in subsection 4.4 we will apply all this knowledge in order to bootstrap the two-loop six-particle MHV amplitude $\mathcal{E}_{6,0}^{(2)}$. While the bootstrap has currently been applied to multiplicity $n = 6, 7$, whenever possible we will keep n general in the presentation.

4.1. Restrictions on the first and last symbol entries

In local perturbative quantum field theories, amplitudes can only have singularities when intermediate particles go on shell, corresponding to vanishing propagator denominators in contributing Feynman diagrams (vertices can either be constant or polynomial in the momenta). In massless theories, this is possible only when a Mandelstam invariant vanishes, and in the planar limit this is further restricted to the subset of Mandelstam invariants of cyclically adjacent momenta $s_{i,\dots,j-1} = x_{ij}^2$ defined in equation (4). Given that the singularities of

multiple polylogarithms are encoded in the first entry of their symbols, this implies that only the subset of letters formed exclusively out of products and ratios of the aforementioned Mandelstam variables is allowed to appear in the latter [140]. In $\mathcal{N} = 4$ SYM this subset precisely corresponds to the $n(n - 5)/2$ conformal cross ratios defined in equation (5), so concretely the *first entry condition* becomes

$$\text{first symbol entry of } A_{n,k} \in u_{ij}. \tag{54}$$

As a consequence of equation (7), in momentum twistor language the equivalent statement is that the first symbol entries are restricted to multiplicative dual conformal invariant combinations of the $\langle i - 1j - 1j \rangle$ Plücker variables, and of no other Plücker variable or algebraic function thereof. In this manner, one can for example immediately identify that for the particular choices (40) and (41) of $n = 6, 7$ alphabets we have made, (54) specializes to

$$\text{first symbol entry of } A_{n,k} \in \begin{cases} a_i, & i = 1, \dots, 3, \text{ for } n = 6, \\ a_{1i}, & i = 1, \dots, 7, \text{ for } n = 7. \end{cases} \tag{55}$$

Similarly to the first entry, the last entry of the symbol of $\mathcal{N} = 4$ amplitudes is also constrained, this time by the \bar{Q} -equation [141], which encodes how their Yangian symmetry is broken at loop level by infrared divergences. The precise form of the constraint depends on the helicity k , in line with the fact that beyond the $k = 0$ or MHV case the transcendental functions associated to loop corrections to the amplitude are also multiplied by rational functions of helicities and momenta with tree-level origin (recall that in our normalized amplitude definition (12) we have divided out by the corresponding MHV rational factors). Hence, every linearly independent rational factor in the superamplitude, whose number for the first few helicity configurations has been found to be [47]⁹,

$$\begin{aligned} & \# \text{ linearly independent components of } A_{n,k} \quad \text{for} \\ k = 0: & \quad 1, \\ k = 1: & \quad \binom{n-1}{4}, \\ k = 2: & \quad \frac{(n-5)(n-4)^2(n-3)^2(n-2)^2(n-1)}{4!5!}, \end{aligned} \tag{56}$$

will also come with a nontrivial transcendental function multiplying it. In general, it will be the final entries of these functions times the rational factors that are related to each other by the \bar{Q} -equation.

For simplicity, here we will thus only quote the *MHV final entry condition*,

$$\text{final symbol entry of } A_{n,0} \in \langle ij - 1jj + 1 \rangle, \tag{57}$$

which we can again specialize to

$$\text{final symbol entry of } A_{n,0} \in \begin{cases} m_i, y_i, & i = 1, \dots, 3, \text{ for } n = 6, \\ a_{2i}, a_{3i}, & i = 1, \dots, 7, \text{ for } n = 7. \end{cases} \tag{58}$$

⁹ Just to get a sense of these numbers, the NMHV amplitude has 5, 15, 35 and 70 NMHV components for $n = 6, 7, 8, 9$ and the N²MHV one 105 and 490 components for $n = 8, 9$, respectively. The complete classification of these Yangian invariant rational functions, according to n, k and cyclic class, has been carried out in [64].

in our choices of dual conformal invariant alphabets (40) and (41). The NMHV final entries may be found in [13, 17] for $n = 6, 7$, and in [142] for arbitrary n .

Finally, it is worth noting that the \bar{Q} -equation in fact contains significantly more information than these final entry conditions, as it provides alternative representations for amplitudes, as integrals over a collinear limit of amplitudes with higher multiplicity and MHV degree, and lower loop order. In this manner, not only does it predict similar constraints deeper inside the symbol, such as MHV next-to-final-entry conditions [139], but it also offers an alternative route for the direct computation of amplitudes, as was successfully carried out in [142–144].

4.2. Constructing the function space containing the n -particle amplitude

In the previous subsection we discussed the special restrictions holding at the two endpoints of the symbol, how about its remaining entries? One may be tempted to think that all possible combinations of letters or ‘words’ of a given alphabet give rise to well-defined functions (so that for an alphabet of size $|\Phi|$ at weight k we would obtain $|\Phi|^k$ such functions). However this is not true, since any well-defined function F must satisfy the property that double derivatives with respect to two different independent variables x_i, x_j commute,

$$\frac{\partial^2 F}{\partial x_i \partial x_j} - \frac{\partial^2 F}{\partial x_j \partial x_i} = 0, \quad i \neq j. \tag{59}$$

When F is an MPL, due to equations (17) and (18) this requirement implies the existence of linear relations between its double coproducts $F^{\phi_\alpha, \phi_\beta}$,

$$\sum_{\alpha, \beta=1}^{|\Phi|} D_{i\alpha\beta} F^{\phi_\alpha, \phi_\beta} = 0, \quad i = 1, 2, \dots, l, \tag{60}$$

where D is a tensor with purely numeric entries, and l counts the number of independent equations, which generally depends on the choice of alphabet. These equations are known as the *integrability conditions*, and from equation (20) we see that they are also equivalent to the fact that of all $|\Phi|^2$ combinations of letters, only a subset of weight-two symbols can appear at the last two coproduct slots.

Further focusing on the MPL function spaces containing $A_{n,k}$, in subsection 2.3 we have already mentioned that the variables x_i in equation (59) can be chosen to be an algebraic independent subset of cross ratios, or more conveniently to coincide with the variables of any momentum twistor parametrization of the kinematics, such as the $\text{Gr}(4, n)$ \mathcal{X} -coordinates of a given cluster. For the known six- and seven-particle alphabets, in this manner it can be shown that there exist $l = 26$ equations for the $9^2 = 81$ double coproducts and $l = 729$ equations for the $42^2 = 1764$ double coproducts, respectively [17]. Their explicit form may be found in the ancillary file accompanying the `arXiv` submission of [38], together with the extended Steinmann relations (53), which also assume the general form (60) and thus may be described by an enlarged matrix D .

With the knowledge of the n -particle alphabet as a starting point, the above properties allow us to recursively construct the space of (extended Steinmann/cluster adjacent) n -gon functions containing the L -loop amplitude at weight $m = 2L$, which we shall denote $\mathcal{H}_{n,m}$, as follows: given that a basis of functions on $\mathcal{H}_{n,m-1}$ is known, we consider their $\{m-2, 1\}$ coproduct representation (19), and attach another letter to them to the right in all possible ways. From this $\{m-2, 1, 1\}$ tensor product space that resembles equation (20), we then obtain $\mathcal{H}_{n,m}$ by imposing the integrability and extended Steinmann relations (60) on its elements. The procedure starts at $m = 1$ with the functions dictated by the first entry condition (54), and

terminates at the desired weight $m = 2L$, where also the final entry condition (57) (for $k = 0$) or its generalization (for $k > 0$) may be imposed if one is interested in a particular helicity configuration.

Hence the bootstrap method simplifies amplitude computations by transforming them into linear algebra, and has so far successfully been applied to determine $A_{n,k}$ for $n = 6, 7$ to unprecedented loop orders, that would have been completely out of reach with traditional Feynman diagram methods [2, 9–19, 21]. The construction of $\mathcal{H}_{n,m}$ is computationally the most challenging part of the bootstrap program as m increases, but still the resulting systems of linear equations are many orders of magnitude smaller than, e.g. the integration-by-parts identities needed to determine the basis of master integrals at the same loop order. As first proposed and applied in [16], they can be most efficiently solved by finite field methods that avoid intermediate expression swell, implemented for example in software such as IML [145], SageMath [146] or SpaSM [147] has been used the past. More recently, the SymBuild [148] package and the FiniteFlow framework [149], with dedicated capabilities for constructing integrable symbols, have also been made available.

The procedure we have described works equally well for functions or for their symbols, and ensures that if $\mathcal{H}_{n,m-1}$ has physical branch cuts, so will $\mathcal{H}_{n,m}$, with the following exception for the case of functions: the space of solutions of the integrability conditions (60) also contains functions such as

$$\zeta_{m-1} \log \phi_\alpha, \tag{61}$$

where ζ_{m-1} is the Riemann zeta function and ϕ_α is a letter that is not an allowed first entry (54), which do not have physical branch cuts.

It is therefore necessary to eliminate such functions from our space, and one way to do so is by noting that the branch point at $\phi_\alpha = 0$ also manifests itself as a pole in the derivative of the function. So ensuring that the function is analytic at $\phi_\alpha = 0$ can be achieved by requiring that the corresponding residue of its derivative, or by virtue of equation (17) its left coproduct factor, vanishes as $\phi_\alpha \rightarrow 0$.

In practice, it is simpler to impose such *branch cut conditions* on kinematic limits where more letters vanish simultaneously. For the purposes of this review, it will be sufficient to mention one such limit that has been used in the literature for $n = 6$, the *soft limit*. In the choice (40) for the six-particle alphabet and in the i th orientation it amounts to¹⁰

$$\begin{aligned} a_i \rightarrow \infty, \quad a_{i-1} \rightarrow \frac{1}{a_{i+1}}, \\ \text{soft}_i : \quad m_{i+1} \rightarrow \frac{\sqrt{a_i}}{\sqrt{a_{i+1}}}, \quad m_{i-1} \rightarrow \sqrt{a_i} \sqrt{a_{i+1}}, \quad y_i \rightarrow 1, \end{aligned} \tag{62}$$

$$\text{with } a_{i+1}, y_{i+1}, y_{i-1}, m_i \sqrt{a_i} \text{ fixed, } \quad i = 1, 2, 3,$$

with each of the soft limits corresponding to one square face, that also intersect the parity-even surface, of the $\text{Gr}(4, 6)$ cluster polytope shown in figure 3. In these limits, functions in $H_{6,m}$ should additionally satisfy the following branch cut conditions [10, 13]

$$F^{m_i} \Big|_{\text{soft}_i} = F^{y_{i-1}} \Big|_{\text{soft}_i} = F^{y_{i+1}} \Big|_{\text{soft}_i} = 0, \quad i = 1, 2, 3. \tag{63}$$

Analogous conditions for $n = 7$ have been obtained in [19].

¹⁰This limit may also be expressed in terms of the more conventional cross ratios (6) as $u_i \rightarrow 1, u_{i-1}, u_{i+1} \rightarrow 0$ with $u_j/(1 - u_j)$ held fixed for $j \neq i$.

4.3. Singling out the amplitude: special kinematic limits

Once the space of n -gon functions $H_{n,m}$ containing the n -particle amplitude has been constructed, the final step of the bootstrap method is to uniquely identify the latter from within this space, using information from kinematic limits where the behavior of the amplitude is already known. The simplest of these special kinematic configurations is the limit where the momenta of two consecutive external particles become collinear: indeed, the BDS ansatz correctly captures not just the infrared singularity structure of the amplitude, but also its behavior under collinear factorization. As a consequence, in this limit the BDS-normalized amplitude smoothly reduces to the same amplitude with one leg less, for example for the MHV case this immediately carries over directly for the remainder function (16),

$$\lim_{i+1\parallel i} \mathcal{R}_n^{(L)} = \mathcal{R}_{n-1}^{(L)}, \tag{64}$$

with the cascade terminating at $\mathcal{R}_5^{(L)} = 0$. As an example that will also be useful to us in the next subsection, let us quote here how the symbol alphabet behaves in the case of the $n = 6$ collinear limit,

$$\begin{aligned} & m_i \rightarrow \infty, \quad m_{i-1} \rightarrow \frac{1}{m_{i+1}}, \\ \text{collinear}_i : & \quad a_i \rightarrow \frac{(1 + m_{i+1})^2}{m_i m_{i+1}}, \quad a_{i-1} \rightarrow m_i m_{i+1}, \quad a_{i+1} \rightarrow \frac{m_i}{m_{i+1}}, \\ & y_{i-1} \rightarrow 1, \quad y_{i+1} \rightarrow 1, \quad \text{with } m_{i+1}, y_i \text{ fixed, } i = 1, 2, 3. \end{aligned} \tag{65}$$

The three collinear limits correspond to the three edges of the $\text{Gr}(4, 6)$ cluster polytope shown in figure 3, that lie on the parity-even surface.

Beyond the strict collinear limit, Mellin–Barnes-like integral representations for every term in the series expansion around an $(n - 5)$ -fold collinear limit may be predicted with the help of the integrability-based Wilson loop or pentagon operator product expansion (OPE) [150–160], to all loops. These integral representations can then be systematically evaluated in closed form [15, 161, 162], thus providing direct input for the amplitude bootstrap, and in some cases the entire series expansion or certain well-defined subsector thereof may even be resummed so as to access more general kinematic configurations [163–169].

Last but not least, an excellent source of boundary kinematic data for the bootstrap is offered by the high energy or multi-Regge kinematics (MRKs), a very rich subject in its own right, which will be the focus of chapter 15 [170] of this review [42]. This owes to the development of an effective description of the latter by Balitsky, Fadin, Lipatov and Kuraev originally in QCD, that was later extended also to planar $\mathcal{N} = 4$ SYM [12, 25, 163, 171–177]. Interestingly, the dual conformal invariance of the theory renders it equivalent to the soft limit, so in order for the normalized amplitude to have nontrivial kinematic dependence there, it is necessary to first analytically continue away from the Euclidean region. At multiplicity $n = 6$ we have already seen the soft/multi-Regge limit in equation (62), from where it becomes apparent that is natural to organize the weak coupling expansion of the amplitude also with respect to the order of the divergent logarithm, $\log^{L-p-1} a_1$, denoted as the (next-to) p -leading-logarithmic ($N^p\text{LL}$) approximation. Remarkably, this double expansion can be computed at any loop order and logarithmic approximation, not only for $n = 6$ [178], but also at arbitrary multiplicity [179], thanks to an analytic continuation connecting the multi-Regge with the near-collinear limit mentioned above. For $n = 7$, these all-loop results have been recently checked against all available bootstrap data [180].

4.4. A simple example: bootstrapping the two-loop six-particle MHV amplitude

With all the bootstrap technology in place, let us now see it at work in a concrete example, the computation of the first nontrivial correction at $L = 2$ loops (since the $L = 1$ correction is by construction part of the BDS ansatz) to the six-particle amplitude in the MHV helicity configuration. This subsection is thus intended for the reader who is interested in learning how to perform actual bootstrap computations, and may otherwise be skipped.

In the first instance, we will construct the space of hexagon functions $\mathcal{H}_{6,m}$ containing the amplitude and its derivatives, for $1 \leq m \leq 2L = 4$. At $m = 1$, the first entry condition (55) requires this space to be

$$H_{6,1} : \log a_i, \quad i = 1, 2, 3. \tag{66}$$

To go to higher weight, we first need to determine the integrability conditions in our choice (40) for the six-particle alphabet. To this end, for a generic function $F \in \mathcal{H}_{6,m}$ we compute the commutator of double derivatives with the help of the definitions (17) and (18), which generically takes the form

$$\sum_{\phi_\alpha, \phi_\beta \in \Phi} F^{\phi_\alpha, \phi_\beta} \left[\frac{\partial \log \phi_\alpha}{\partial x^i} \frac{\partial \log \phi_\beta}{\partial x^j} - \frac{\partial \log \phi_\alpha}{\partial x^j} \frac{\partial \log \phi_\beta}{\partial x^i} \right], \tag{67}$$

since terms with both derivatives acting on the letters ϕ automatically commute and cancel out. Next, we express the letters in terms of the three independent variables of any momentum twistor parametrization, such as the $\text{Gr}(4, 6)$ \mathcal{X} -coordinates of the initial cluster, which we quoted in equation (9), from which the term in brackets in the above equation is trivial to compute. This gives three equations for each $1 \leq i < j \leq 3$, which should hold for any value of the x_i . We can therefore convert them into equations for the double coproducts $F^{\phi_\alpha, \phi_\beta}$ with purely numeric coefficients, either by collecting all terms under a common denominator and demanding that they hold separately for each coefficient of the polynomials in the numerator, or by evaluating them for sufficiently many values of the x_i . Explicitly, and in the shorthand notation $F^{[x,y]} = F^{x,y} - F^{y,x}$ the thus derived linear equations read

$$\begin{aligned} F^{[a_1, a_2]} &= F^{[a_1, m_1]} = F^{[a_1, y_1]} = F^{[a_1, y_2]} - F^{[a_1, y_3]} = F^{[m_1, y_2]} - F^{[m_1, y_3]} = 0, \quad i = 1, 2, 3, \\ F^{[m_1, m_2]} + F^{[m_1, m_3]} &= F^{[m_3, a_1]} + F^{[a_2, m_3]} + F^{[m_1, m_3]} + F^{[y_1, y_2]} = 0, \quad i = 1, 2, \\ F^{[a_2, y_1]} + F^{[a_3, y_1]} + F^{[m_1, y_1]} &= F^{[a_1, y_2]} + F^{[a_3, y_1]} + F^{[m_2, y_2]} = 0, \\ F^{[a_1, y_2]} + F^{[a_2, y_1]} + F^{[m_3, y_3]} &= F^{[a_2, y_1]} - F^{[a_3, y_1]} - F^{[m_2, y_1]} + F^{[m_3, y_1]} = 0, \\ F^{[a_1, y_2]} - F^{[a_3, y_1]} - F^{[m_1, y_2]} + F^{[m_3, y_1]} &= F^{[m_2, a_1]} + F^{[a_3, m_2]} + F^{[m_1, m_2]} + F^{[y_1, y_3]} = 0, \\ F^{[m_1, m_2]} - F^{[y_1, y_2]} + F^{[y_1, y_3]} - F^{[y_2, y_3]} &= 0. \end{aligned} \tag{68}$$

With the six-particle integrability conditions (68) at hand, constructing the weight-2 space of functions is now simply a matter of building an ansatz for the most general form of their $\Delta_{1,1}$ coproduct, obtained by adding another hexagon letter to $H_{6,1}$ in all possible ways (in compact symbol notation $\log \phi \rightarrow \phi$),

$$\sum_{i=1}^3 \sum_{j=1}^9 c_{ij} a_i \otimes \phi_j, \tag{69}$$

and imposing the integrability conditions simultaneously with the extended Steinmann relations (53) on this ansatz¹¹. In total these are 29 homogeneous linear equations, but when applied to the above ansatz only 21 of them are linearly independent, and fix an equal number of the 27 unknowns c_{ij} . Thus the coefficients of the remaining six unknowns will span the allowed weight-two space, and explicitly we find these to be

$$H_{6,2} : \begin{cases} -[(a_{i+1}a_{i-1}) \otimes m_i] = -[(1 + m_i)^2 \otimes m_i] \rightarrow -2H_{-2}(m_i), \\ a_i \otimes a_i \rightarrow \frac{1}{2} \log^2 a_i, \end{cases} \quad i = 1, 2, 3. \quad (70)$$

In the first line we chose to express the letters using their subset m_i as the independent variables,

$$a_i = \frac{(1 + m_{i+1})(1 + m_{i-1})}{1 + m_i}. \quad (71)$$

The simple form of the $\Delta_{1,1}$ coproducts also allows us to immediately identify the corresponding functions, see in particular equation (27) for the first line, indicated with arrows in the above equation. More precisely, the $\Delta_{1,1}$ coproduct is equivalent to the total differential of a function and thus specifies it up to a constant, which we are free to choose for example so as to simplify the functional expression, as is done here. This choice is tantamount to the choice of base point for the integration of the total differential, which is usually chosen as a potentially singular point of the functions. So if we wish our basis to be independent of this choice of singular base point, it is natural to also include the transcendental constants the above functions evaluate to at these points, in this case $\zeta_2 = H_2(1)$.

Moving on to weight three, to find a basis of functions we similarly form an ansatz of the six functions of equation (70) tensored with the nine hexagon letters, giving rise to a total of 54 unknowns. We do not need to include $\zeta_2 \log \phi_j$ in our ansatz since these functions have identically vanishing double coproducts, and hence correspond to trivial solutions of the integrability conditions and extended Steinmann relations, which we know beforehand. We then apply the aforementioned constraints on the right two slots of the $\Delta_{1,1,1}$ coproduct of the ansatz, in the first instance obtaining $29 \times 3 = 87$ equations, since these should hold separately for each of the leftmost coproduct slots, which are algebraically independent. However these reduce to 41 linearly independent equations, whose solution space is then found to be¹²,

$$H_{6,3} : \begin{cases} 2H_{-2}(m_i) \otimes m_i \rightarrow 2H_{-3}(m_i), \\ -2H_{-2}(m_i) \otimes (1 + m_i)^2 \rightarrow -4H_{-1,-2}(m_i), \\ \frac{1}{2} \log^2 a_i \otimes a_i \rightarrow \frac{1}{6} \log^3 a_i, \\ -H_{-2}(m_i) \otimes \frac{a_{i-1}}{a_{i+1}} - \frac{1}{2} (\log^2 a_{i-1} - \log^2 a_{i+1}) \otimes m_i \rightarrow -H_{-2}(m_i) \log \frac{a_{i-1}}{a_{i+1}}, \\ \tilde{F} \equiv \sum_{i=1}^3 \left[\frac{1}{2} (\log^2 a_{i-1} + \log^2 a_{i+1}) - 2H_{-2}(m_1) - 2H_{-2}(m_2) - 2H_{-2}(m_3) \right] \otimes y_i, \end{cases} \quad (72)$$

¹¹ Note that cluster adjacency may be exploited so as to reduce the number of initial unknowns and equations, by attaching a letter ϕ_j only to those functions whose final entries belong to the neighbor set (46) and (47) of ϕ_j [18, 109]. While this increases the efficiency of the method, for the sake of simplicity we will refrain from applying it here.

¹² From this point on, we will employ the compact symbol notation only to the right factor of the $\Delta_{m-1,1}$ coproduct, and otherwise retain the complete functional form of the left factor.

with $i = 1, 2, 3$. Notice that unlike all functions encountered thus far, \tilde{F} is parity odd, owing to the appearance of the parity-odd letters y_i for the first time. Of the remaining even functions, only the last one is not immediately identifiable with the help of the usual HPL definitions, and requires expressing the difference of the squares of logarithms as a product thereof, using the m_i as the independent variables, and applying the shuffle algebra relations mentioned in subsection 3.1.

At this stage all functions are defined modulo $\zeta_2 \log \phi_j$, the trivial solution we chose not to include in our ansatz. This ambiguity may be fixed by further imposing the branch cut conditions (63), which were automatically satisfied at lower weight since no analogous ambiguity existed. While it is easy to check that all even $\mathcal{H}_{6,3}$ functions in equation (72) satisfy them, for \tilde{F} we find that in the soft limits (62)

$$\tilde{F}^{y_{i-1}} \Big|_{\text{soft}_i} = \tilde{F}^{y_{i+1}} \Big|_{\text{soft}_i} = -4\zeta_2. \tag{73}$$

The only slightly nontrivial step needed to show this, and more generally to evaluate all kinematic limits we will consider later in this subsection, are HPL $x \rightarrow 1/x$ argument inversion identities, which for example can be obtained with the package HPL [90]. From the above equation, it is clear that in order to ensure that \tilde{F} has good branch cuts, we need to redefine it as

$$\tilde{F} \equiv \sum_{i=1}^3 \left[\frac{1}{2} (\log^2 a_{i-1} + \log^2 a_{i+1}) - 2H_{-2}(m_1) - 2H_{-2}(m_2) - 2H_{-2}(m_3) + 4\zeta_2 \right] \otimes y_i, \tag{74}$$

and now the only ambiguity remaining in its definition is its value at a point. This may then be fixed by picking this point anywhere on the parity-invariant surface, where by definition any parity-odd function vanishes. In our choice of basis, \tilde{F} is in fact equal to twice the transcendental part of the six-dimensional hexagon integral studied in [96], where also an explicit expression of the latter in terms of classical polylogarithms may be found. Finally, in $\mathcal{H}_{6,3}$ we may additionally include the part of the trivial solution which obviously also satisfies the branch cut conditions, $\zeta_2 \log a_i$, as well as ζ_3 , by the same reasoning that led to the inclusion of ζ_2 at one weight lower.

Arriving at weight 4, since our goal here is MHV amplitude, which is parity even and obeys the final entry condition (58), we may simplify the calculation by incorporating these constraints directly in our initial ansatz. Namely we tensor the even $\mathcal{H}_{6,2}$ functions with m_i , and \tilde{F} with y_i . For this subspace of $\mathcal{H}_{6,4}$, which we shall denote $\mathcal{H}_{6,4}^{+,\text{MHV}}$, we find the basis

$$\mathcal{H}_{6,4}^{+,\text{MHV}} : \begin{cases} -2H_{-3}(m_i) \otimes m_i \rightarrow -2H_{-4}(m_i), \\ 4H_{-1,-2}(m_i) \otimes m_i \rightarrow 4H_{-2,-2}(m_i), \\ \hat{\Omega}_i^{(2)}, \\ -2\zeta_2 H_{-2}(m_i), \end{cases} \quad i = 1, 2, 3, \tag{75}$$

where

$$\Delta_{3,1}\hat{\Omega}_i^{(2)} = \hat{\Omega}_i^{(2),m_i} \otimes m_i + \hat{\Omega}_i^{(2),m_{i+1}} \otimes m_{i+1} - \tilde{F} \otimes (y_i y_{i+1}), \quad (76)$$

$$\begin{aligned} \hat{\Omega}_i^{(2),m_i} \equiv & -4H_{-3}(m_i) + 2[H_{-1,-2}(m_i) + H_{-1,-2}(m_{i+1}) - H_{-1,-2}(m_{i-1})] \\ & + \frac{1}{3}\log^3 a_{i+1} - H_{-2}(m_{i-1})\log \frac{a_{i+1}}{a_i} + H_{-2}(m_i)\log \frac{a_{i-1}}{a_{i+1}} \\ & - H_{-2}(m_{i+1})\log \frac{a_i}{a_{i-1}} + 4\zeta_2 \log a_{i+1} \end{aligned} \quad (77)$$

$$\hat{\Omega}_i^{(2),m_{i+1}} = \Omega_i^{(2),m_i} \Big|_{\substack{m_j \leftrightarrow m_{j+1} \\ a_j \leftrightarrow a_{j+1}}} \quad (78)$$

Clearly, $\hat{\Omega}_i^{(2)}$ is symmetric under exchange of letters with indices $i \leftrightarrow (i + 1)$. Recalling that \tilde{F} vanishes in the parity-invariant surface containing (intersecting) the collinear (soft) limit, and using HPL argument inversion identities as before, it is a straightforward exercise to show that the above functions already satisfy the branch cut conditions (63), and thus require no further modification. As with lower weights, we also include the constant ζ_4 in $\mathcal{H}_{6,4}^{+,MHV}$.

Therefore the only task remaining in order to fully specify our $\mathcal{H}_{6,4}^{+,MHV}$ basis is the value of $\hat{\Omega}_i^{(2)}$ at a point. We can do this by noting that its $\Delta_{3,1}$ coproduct vanishes in the $m_{i-1} \rightarrow \infty$ orientation of the collinear limit (65), therefore it reduces to a constant which is natural to also set to zero. This choice in fact renders $\hat{\Omega}_i^{(2)}$ four times the double pentagon integral $\Omega^{(2)}$, which we will encounter again in section 5.2, in its three possible orientations.

Having fully specified $\hat{\Omega}_i^{(2)}$ in this manner, it is not difficult to similarly obtain its other two nontrivial collinear limits, which we will need later on, from the coproduct representation (76). In the $m_i \rightarrow \infty$ orientation in particular we find,

$$\begin{aligned} \Delta_{3,1}\hat{\Omega}_i^{(2)} \rightarrow & 4 \left[(H_{-3}(m_{i+1}) + H_{-2,0}(m_{i+1}) - 2H_{-2,-1}(m_{i+1}) \right. \\ & + \log m_i H_{-2}(m_{i+1})) \otimes m_i + (H_{-2,0}(m_{i+1}) - 2H_{-2,-1}(m_{i+1}) \\ & - H_{-1,-2}(m_{i+1}) + 4H_{-1,-1,-1}(m_{i+1}) - 2H_{-1,-1,0}(m_{i+1}) \\ & - 2 \log m_i (H_{-1,-1}(m_{i+1}) + H_{-1,0}(m_{i+1}) + H_{-2}(m_{i+1})) \\ & \left. + \left(\frac{1}{2}\log^2 m_i + \zeta_2 \right) H_{-1}(m_{i+1}) \right) \otimes m_{i+1} \Big]. \end{aligned} \quad (79)$$

This expression can be trivially integrated to yield $\hat{\Omega}_i^{(2)}$ up to a constant, since it is equivalent to an ordinary differential equation for the function with respect to the only surviving finite variable m_{i+1} . That is, focusing on the coproduct component in question, equation (27) allows us to replace

$$H_{\bar{l}}(m_{i+1}) \otimes m_{i+1} \rightarrow H_{0,\bar{l}}(m_{i+1}) + c. \quad (80)$$

The integration constant is then fixed by the fact that $\hat{\Omega}_i^{(2)}$ should vanish at the $m_{i+1} \rightarrow 0$ endpoint of the $m_i \rightarrow \infty$ collinear line: indeed, this point is an overlap with the $m_{i-1}, m_i \rightarrow \infty, m_{i+1} \rightarrow 0$ soft limit, and the latter in turn also overlaps with the $m_{i-1} \rightarrow \infty$ collinear limit, where as we have seen the function vanishes. In this manner, we finally obtain

$$\begin{aligned}
 \frac{\hat{\Omega}_i^{(2)}}{4} \xrightarrow[m_i \rightarrow \infty]{\text{coll.}} & H_{-3,0}(m_{i+1}) - 2H_{-3,-1}(m_{i+1}) - H_{-2,-2}(m_{i+1}) \\
 & + 4H_{-2,-1,-1}(m_{i+1}) - 2H_{-2,-1,0}(m_{i+1}) \\
 & - 2 \log m_i [H_{-2,-1}(m_{i+1}) + H_{-2,0}(m_{i+1}) + H_{-3}(m_{i+1})] \\
 & + \left(\frac{1}{2} \log^2 m_i + \zeta_2 \right) H_{-2}(m_{i+1}), \tag{81}
 \end{aligned}$$

whereas the third collinear limit orientation follows for free by exploiting the flip symmetry of the function so as to replace $m_i \leftrightarrow m_{i+1}$ in the above formula.

Now that the hard part of constructing the function space containing the amplitude is over, all that is left to determine the latter is to form an ansatz from all the basis functions, and determine the coefficients by comparing it to special kinematic limits where we have independent information on the behavior of the amplitude. Taking into account the dihedral symmetry of the BDS-like normalized six-particle amplitude, our initial ansatz contains merely five unknowns,

$$\begin{aligned}
 \mathcal{E}_{6,0}^{(2)} = & c_1 \sum_{i=1}^3 [-2H_{-4}(m_i)] + c_2 \sum_{i=1}^3 4H_{-2,-2}(m_i) \\
 & + c_3 \sum_{i=1}^3 \hat{\Omega}_i^{(2)} + c_4 \sum_{i=1}^3 [-2\zeta_2 H_{-2}(m_i)] + c_5 \zeta_4, \tag{82}
 \end{aligned}$$

where we remind the reader that the function $\hat{\Omega}_i^{(2)}$ is defined by its $\Delta_{3,1}$ coproduct (76) and the fact that it vanishes in the $w_{i-1} \rightarrow \infty$ collinear limit (65).

In order to fix the coefficients of the ansatz we will also consider the collinear limit, where equations (11), (16) and (64) imply that at this loop order the amplitude has the simple behavior

$$\mathcal{E}_{6,0}^{(2)} \xrightarrow{\text{collinear}} \frac{1}{2} \left(\mathcal{E}_{6,0}^{(1)} \right)^2. \tag{83}$$

More precisely, given that our ansatz has already taken dihedral symmetry into account, we may focus on a single orientation, say $m_3 \rightarrow \infty$, where the one-loop correction (14) reduces to

$$\mathcal{E}_{6,0}^{(1)} \xrightarrow[m_3 \rightarrow \infty]{\text{collinear}} -\frac{1}{2} \log^2 m_1 - \frac{1}{2} \log^2 m_3 - 2\zeta_2. \tag{84}$$

The final step is to also evaluate our ansatz in the limit. For $\hat{\Omega}_i^{(2)}$ this has already been done in equation (81), and for the other functions we proceed similarly. Recalling also the MPL identity (24), after the dust settles the difference of the right-hand sides of equations (83) and (82), evaluated on the $m_3 \rightarrow \infty$ collinear limit will be a sum of functions

$$\log^i(m_3) \zeta_j H_{-l_1, -l_2, \dots, -l_k}(m_1), \tag{85}$$

of total weight four, with coefficients depending on the unknowns c_i . Since each of these functions is algebraically independent, their coefficients should vanish separately, and solving this set of equations yields the unique solution

$$c_1 = -2, \quad c_2 = -\frac{1}{4}, \quad c_3 = \frac{1}{4}, \quad c_4 = 1, \quad c_5 = 8. \tag{86}$$

Table 3. Remaining parameters in the ansätze for the (MHV, NMHV) amplitude after each constraint is applied, at each loop order. The superscript “*” (“* n ”) denotes an additional ambiguity (n ambiguities) which arises only due to lack of knowledge of the cosmic normalization constant ρ at the given stage. The ‘?’ indicates an ambiguity about the number of weight 12 odd functions that are ‘dropouts’; they are allowed at symbol level but not function level. Reproduced from [15]. [CC BY 4.0](#).

Constraint	$L = 1$	$L = 2$	$L = 3$	$L = 4$	$L = 5$	$L = 6$
1. \mathcal{H}_6	6	27	105	372	1214	3692?
2. Symmetry	(2, 4)	(7, 16)	(22, 56)	(66, 190)	(197, 602)	(567, 1795?)
3. Final-entry	(1, 1)	(4, 3)	(11, 6)	(30, 16)	(85, 39)	(236, 102)
4. Collinear	(0, 0)	(0, 0)	(0*, 0*)	(0*, 2*)	(1* ³ , 5* ³)	(6* ² , 17* ²)
5. LL MRK	(0, 0)	(0, 0)	(0, 0)	(0, 0)	(0*, 0*)	(1* ² , 2* ²)
6. NLL MRK	(0, 0)	(0, 0)	(0, 0)	(0, 0)	(0*, 0*)	(1*, 0* ²)
7. NNLL MRK	(0, 0)	(0, 0)	(0, 0)	(0, 0)	(0, 0)	(1, 0*)
8. N ³ LL MRK	(0, 0)	(0, 0)	(0, 0)	(0, 0)	(0, 0)	(1, 0)
9. Full MRK	(0, 0)	(0, 0)	(0, 0)	(0, 0)	(0, 0)	(1, 0)
10. T^1 OPE	(0, 0)	(0, 0)	(0, 0)	(0, 0)	(0, 0)	(1, 0)
11. T^2 OPE	(0, 0)	(0, 0)	(0, 0)	(0, 0)	(0, 0)	(0, 0)

Congratulations, you have just bootstrapped the two-loop correction to the planar six-particle amplitude, or equivalently lightlike hexagon Wilson loop in planar $\mathcal{N} = 4$ SYM theory! Its initial computation in terms of Feynman diagrams required a rather nontrivial effort, and resulted in a 17-page long sum of multiple polylogarithms, that is equivalent to the expressions (82) and (86).

Let us close this section with a few remarks on how the bootstrap ideas we have presented in this example apply more generally.

- For the hexagon bootstrap we have considered here, it is possible to further reduce the size of $\mathcal{H}_{6,m}$, and thus to facilitate the identification of the amplitude, by only including the constants ζ_{2n} with $n \geq 2$ as independent functions. This requires a further modification of the amplitude normalization by a coupling-dependent constant, dictated by what is known as a (cosmic Galois) coaction principle [21], see also the review [38], initially carried out order by order in perturbation theory, and later conjectured to all loops in [169]. The number of functions remaining after applying consecutive constraints, so as to uniquely determine the amplitude together with this ‘cosmic’ normalization through six loops, is summarized in table 3. The difference between our count of five unknowns in our ansatz (82) and the four unknowns quoted for the MHV case in the $L = 2$ column and final-entry row, is precisely due to our redundant inclusion of ζ_2 as an independent constant. Though it does not appear in the table, $\mathcal{E}_{6,0}^{(7)}$ has also been determined in [15], and $\mathcal{E}_{6,1}^{(7)}$ is also known [181].
- Similarly to what we did for the more complicated functions encountered in our example, in general it proves more economical to recursively represent each weight- m MPL in terms of its differential or $\Delta_{m-1,1}$ coproduct, together with its value at a point [10]. As is discussed in chapter 3 [79] of the SAGEX review [42], this is not a restriction however, since explicit G -function representations may be found algorithmically when there exists a choice of variables such that the symbol letters are rational functions thereof. This is even simpler when these functions are further restricted to be linear, where algorithmic integration via *fibration bases* [182, 183], see also [184, 185], has been implemented in

the software packages `HyperInt` [186], `MPL` [187] and `PolyLogTools` [91] (with the first of the three in fact based on a further refinement of this algorithm).

- Building on the aforementioned coproduct representation, a significant efficiency upgrade that becomes necessary at higher loops is to encode all the information on integrable symbols and functions, as well as of the equations needed to construct them, in terms of tensors with purely numeric entries [17], see also [21]. The latter tensor has already appeared when expressing the integrability and Steinmann relations as we did in equation (60), and the former tensor simply relates a function basis element at weight m to a basis element at weight $m - 1$ as well as to the position of a letter of an ordered alphabet. The advantage of this approach is that it not only provides the most compact way of storing all function data, but most importantly that it reduces the iterative construction of the function space exclusively to matrix operations. At the level of the symbol, it has been implemented in the package `SymBuild` [148].

5. New frontiers

5.1. $n > 7$ singularities from tropical Grassmannians

In the previous sections, we have seen the spectacular success of the bootstrap program in determining scattering amplitudes in planar $\mathcal{N} = 4$ SYM theory at multiplicity $n = 6, 7$. A key idea is that given a finite set of singularities, or more precisely symbol letters, the space of polylogarithmic functions containing the amplitude is also finite at each loop order, and thus one can efficiently construct it and single out the actual amplitude. For the aforementioned multiplicities, this set of symbol letters exactly matches the variables of a $\text{Gr}(4, n)$ cluster algebra, thus lending support to the expectation that it should remain stable as the loop order increases.

However even in the ideal setting of the simplest interacting gauge theory, the following significant conceptual and practical challenges prevented the application of the bootstrap in order to efficiently compute amplitudes at higher multiplicity n in general kinematics:

- (a) $\text{Gr}(4, n)$ cluster algebras with $n \geq 8$ become infinite [97], and thus provide no predictability on what the symbol alphabet should be.
- (b) By construction, cluster \mathcal{A} -coordinates are rational functions of the Plücker coordinates $\langle ijkl \rangle$ ¹³. Yet for $n \geq 8$, symbol letters that also contain square roots thereof are known to appear, and can hence not be captured by cluster algebras.

A prototypical example of an integral yielding square-root letters is the one-loop four-mass box depicted in figure 7, which in particular contains $\sqrt{\Delta_{ijkl}}$ with

$$\Delta_{ijkl} \equiv (f_{ij}f_{kl} - f_{ik}f_{jl} + f_{il}f_{jk})^2 - 4f_{ij}f_{jk}f_{kl}f_{il}, \quad f_{ij} \equiv \langle ii + 1 jj + 1 \rangle, \quad (87)$$

see for example [188]. While one could hope that individual Feynman diagram contributions would cancel out so as to yield a simpler result for the amplitude, it can be shown that this is the only diagram contributing to a particular component of $A_{8,2}^{(1)}$ [32].

Very similar natural resolutions of these longstanding problems were simultaneously proposed in [61, 189, 190] based on the relation of cluster algebras with geometric objects known as positive tropical Grassmannians [60, 191], or equivalently their duals as constructed by

¹³This is a direct consequence of the mutation rule (32) or (37), although in fact all denominators cancel, and in the end \mathcal{A} -coordinates simplify to homogeneous polynomials of the four-brackets or Plücker coordinates, as seen e.g. in equation (42).

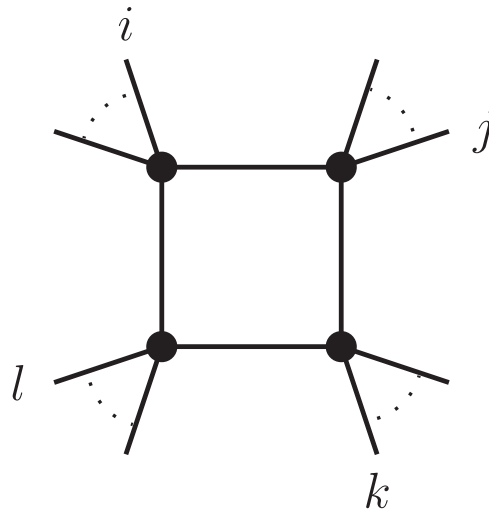


Figure 7. The one-loop four-mass box formed by $n \geq 8$ cyclically ordered momentum twistors, distributed in the four corners as indicated by four of their labels.

‘stringy canonical form’ integrals [125]. In the first instance, this resolution, which we will describe in more detail in the next subsections, may be pictorially represented as in figure 8. It boils down to explicit predictions for the $n = 8$ alphabet, and more recently it has been generalized in principle to any n , and in practice to $n = 9$ [62]. These predictions are in agreement with all currently known data for amplitudes at these multiplicities [142–144] and for $n = 8$ they are also backed by a related but distinct approach based on scattering diagrams and wall-crossing [192]¹⁴. They thus pave the way for bootstrapping new results.

5.1.1. The positive tropical Grassmannian. Let us begin by defining this object, before explaining how it leads to finite alphabet predictions. The simplest way to define the positive part [60] of the tropical Grassmannian [191] $\text{Tr}_+(k, n)$, is by first expressing all $\text{Gr}(k, n)$ Plücker variables in terms of the \mathcal{X} -coordinates of the initial cluster of the corresponding cluster algebra, which can be constructed with the web algorithm presented in the former paper. Then, one *tropicalizes* this parametrization of the Plücker variables, which practically means one replaces

$$\begin{array}{rcl}
 & \text{addition} & \longrightarrow \text{minimum} \\
 \text{tropicalization :} & \text{multiplication} & \longrightarrow \text{addition} \\
 & \mathbb{C}^* \text{ constants} & \longrightarrow 0 \\
 & 0 & \longrightarrow \infty
 \end{array} \tag{88}$$

¹⁴It has been observed that an alternative means for reproducing the square-root letters found in the known data, is by solving polynomial equations associated to certain plabic graphs [193–195]. As soon as one attempts to also incorporate rational letters in this approach, however, non-plabic graphs are required as well [196]. In this case the solution space includes all cluster variables of $G(4, n)$, that is the alphabet becomes infinite again.

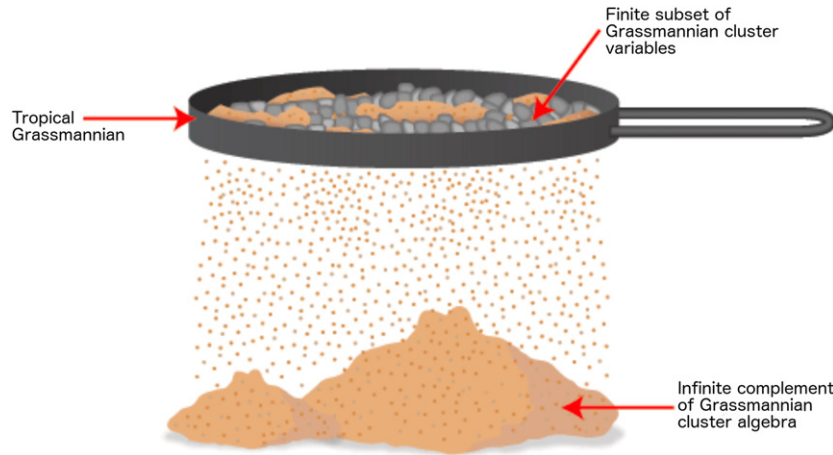


Figure 8. Rough sketch of how a finite subset of variables of $\text{Gr}(k, n)$ cluster algebras may be selected with the help of their associated positive tropical Graßmannians.

We have already seen the $\text{Gr}(4, 6)$ \mathcal{X} -coordinate parametrization in equation (9), so to be concrete let us tropicalize the Plücker variable shown in equation (10),

$$\langle 1346 \rangle = 1 + x_1 + x_1x_2 \rightarrow \min(0, x_1, x_1 + x_2). \tag{89}$$

The *tropical hypersurface* for any such tropicalized polynomial is the $(d - 1)$ -dimensional surface in \mathbb{R}^d where the minimum is attained twice simultaneously¹⁵, see the left-hand side of figure 9 for the case of $\langle 1346 \rangle$. Since the (in)equalities enforcing these are invariant under rescaling, we may equivalently describe the hypersurface with integer-valued vectors

$$g = (-1, 0, 0) \quad g' = (0, 1, 0), \quad g'' = (1, -1, 0). \tag{90}$$

The *positive tropical Graßmannian* $\text{Tr}_+(4, n)$ is then defined as the union of tropical hypersurfaces for all $\langle ijkl \rangle$. Being a solution of linear (in)equalities, it is inherently finite-dimensional. Its building blocks are one-dimensional intersections of tropical hypersurfaces emanating from the origin, known as *rays*. A positive span of certain sets of rays then yields the regions in \mathbb{R}^d bounded by the tropical hypersurfaces, where all tropicalized $\langle ijkl \rangle$ are continuous. These are the *cones*, and the set of all cones then forms a (tropical) *fan*. The full $\text{Tr}_+(4, 6)$ fan is illustrated on the right-hand side of figure 9.

One may also consider further generalizations of $\text{Tr}_+(4, n)$ where any subset of $\langle ijkl \rangle$ (or in fact any \mathcal{A} -coordinate of the corresponding cluster algebra) is tropicalized, leading to different fans but with similar properties with what we have described above. Since $\text{Tr}_+(4, n)$ is not invariant under parity, which is a symmetry of the MHV amplitudes, a particularly natural choice is to tropicalize the maximal parity-invariant subset of Plücker variables $\langle i - 1j - 1j \rangle, \langle ij - 1jj + 1 \rangle$. It is therefore this choice of partial tropicalization of the positive Graßmannian that we will adopt from now on, and we will denote it as $\text{pTr}_+(4, n)$. Further choices and their implications are discussed in [197].

¹⁵ For $\text{Gr}(k, n)$, $d = (k - 1)(n - k - 1)$, but from this point onwards we will specialize to $k = 4$.

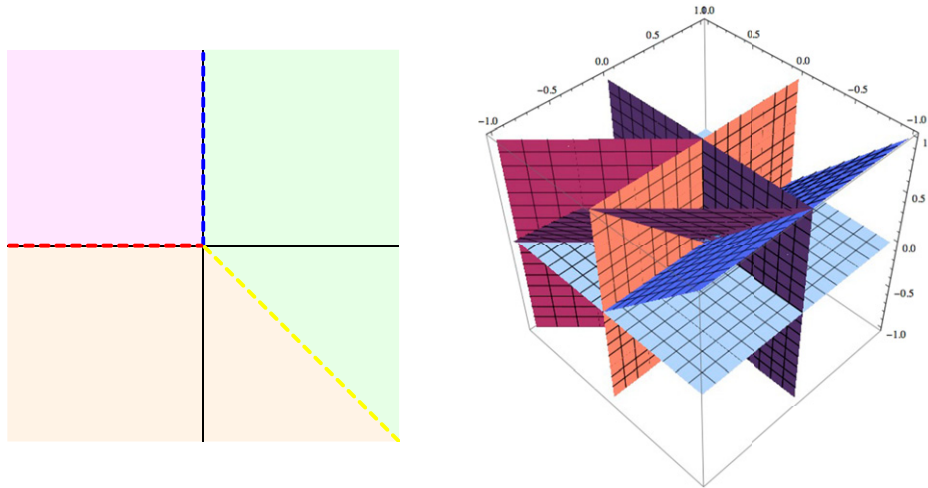


Figure 9. Left: the $\langle 1346 \rangle$ tropical hypersurface, given by the half-planes (in the transverse x_3 direction) $(x_1 = 0, x_2 > 0)$, $(x_1 = -x_2, x_2 < 0)$, $(x_2 = 0, x_1 < 0)$ where $\min(0, x_1, x_1 + x_2)$ is attained twice simultaneously. From reference [190]. Right: the $\text{Tr}_+(4, 6)$ fan. Reproduced from [190]. CC BY 4.0.

5.1.2. *A tropical sieve for rational letters.* Now that we have defined $(\text{p})\text{Tr}_+(4, n)$, we can describe their relation to cluster algebras. Starting with the $\text{pTr}_+(4, 6) = \text{Tr}_+(4, 6)$ fan to the right of figure 9, careful observation reveals that it is dual to the $\text{Gr}(4, 6)$ cluster polytope shown in figure 3! This can be seen by drawing a vertex inside each cone, and connecting them with lines if they are separated by a plane. More rigorously, it can be proven that any cluster \mathcal{A} -coordinate a may be uniquely written as [101]

$$a = \prod_{i=1}^d a_i^{g_i} \cdot \frac{F(x_1, \dots, x_d)}{F_{\mathbb{T}}(y_1, \dots, y_d)}, \tag{91}$$

where a_i, x_i are the \mathcal{A} -coordinates and \mathcal{X} -coordinates of the initial cluster, and the precise definition of the so called coefficients y_i and (tropical) $F_{\mathbb{T}}$ -polynomial will not be important for our purposes. What matters in the above formula is that each \mathcal{A} -coordinate is in one-to-one correspondence with an integer vector (g_1, \dots, g_d) that defines a ray similarly to equation (90), each cluster then defines a cone spanned by the rays of its \mathcal{A} -coordinates, and finally the entire cluster algebra defines a *cluster fan*.

So in the $\text{Gr}(4, 6)$ example the cluster fan coincides with the tropical fan, and more generally the former *triangulates* the latter¹⁶ [60]. This fact has already been used to compute tree-level amplitudes of generalized biadjoint scalar theory [199], defined as a natural extension of the Cachazo–He–Yuan formulation [200, 201] for the corresponding amplitudes in ordinary biadjoint scalar theory, which are essentially given by the volume of the tropical Grassmannian [202]. There, it was also pointed out that the fan of finite cluster algebras may contain not only

¹⁶ More precisely, due to different choice of conventions it is the fan of the *dual* cluster algebra, obtained by transposing its exchange matrices or inverting the arrow direction in its quivers [198], that triangulates the tropical fan. Combinatorially, the cluster fan and its dual are equivalent.

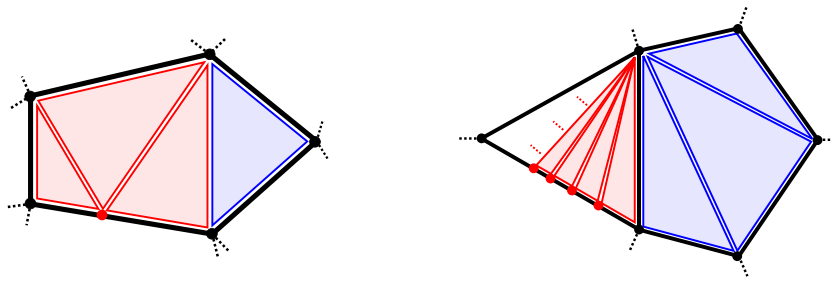


Figure 10. Illustrative examples of the (redundant) triangulation of a tropical fan by a (a) finite and (b) infinite cluster algebra. Each of the figures depicts two cones of a three-dimensional fan intersected with the unit sphere S^2 in black. The cones and the redundant rays from the redundant triangulation are drawn in red, those from the non-redundant triangulation in blue. Reproduced from [190]. CC BY 4.0.

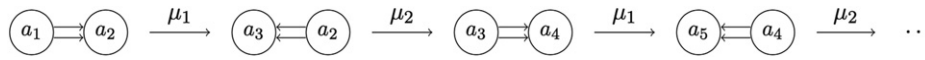


Figure 11. The $A_1^{(1)}$ cluster algebra, where up to orientation each quiver is mapped back to itself after one mutation. Reproduced from [62]. CC BY 4.0.

additional boundaries compared to the tropical fan, but also additional or *redundant rays*, as shown on the left of figure 10.

A key insight behind the works [61, 189, 190] is to turn this logic around in the case of infinite cluster algebras, where it is natural to expect that their infinities are due to infinitely redundant triangulations, as shown on the right of figure 10. Therefore we may prevent this from happening by selecting the *finite subset* of cluster variables whose rays coincide with the tropical rays! Concretely, starting from the initial cluster of the $\text{Gr}(4, n)$ cluster algebra, where all cluster variables have non-redundant rays, after every mutation we compare the ray of the resulting cluster variable to the set of all $\text{pTr}(4, n)$ rays, which may be computed independently e.g. with the program `polymake` [203]. We then stop mutating whenever a redundant ray is reached. It is in this sense that $\text{pTr}(4, n)$ acts as a sieve as shown in figure 8.

5.1.3. Square-root letters from infinite mutation sequences. At this point, we have resolved the first challenge we have discussed at the beginning of this section. How about the second challenge, associated to the appearance of square-root letters in the alphabet of the amplitude? To this end, an important observation comes from existing studies in the mathematical literature, of sequences of cluster mutations that map a quiver back to itself, and thereby lead to recursion relations between the \mathcal{A} -coordinates. In figure 11 we present the simplest example of an infinite cluster algebra where this occurs, of affine rank-2, or $A_1^{(1)}$ in the extended Dynkin diagram classification, type. Very interestingly, the limit of consecutive \mathcal{A} -coordinates along this infinite mutation sequence becomes [204],

$$\lim_{i \rightarrow \infty} \frac{a_i}{a_{i-1}} = \frac{a_2}{2a_1} \left(1 + x_1 + x_1 x_2 + \sqrt{(1 + x_1 + x_1 x_2)^2 - 4x_1 x_2} \right) \quad (92)$$

with $x_1 = 1/a_2^2, x_2 = a_1^2$.

The second main idea of references [61, 189, 190] was to thus also consider infinite mutation sequences, so as to obtain generalized cluster variables of the form (92), which should

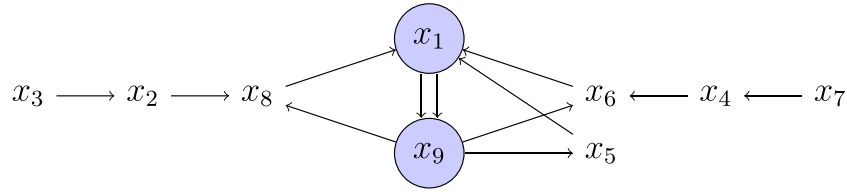


Figure 12. Quiver of a $\text{Gr}(4, 8)$ cluster containing an $A_1^{(1)}$ subalgebra highlighted by the blue nodes. The x_i labels denote the \mathcal{X} -coordinates of the cluster. Reproduced from [190]. CC BY 4.0.

correspond to square-root symbol letters of amplitudes! Indeed, $\text{Gr}(4, n)$ cluster algebras with $n \geq 8$ do contain $A_1^{(1)}$ subalgebras¹⁷, as shown in figure 12. The fine print is that the analogous to equation (92) limit value also depends on all the cluster variables held frozen in the infinite mutation sequence, hence in the language of frozen variables one would have to separately analyze these for every cluster of $\text{Gr}(4, n)$ containing an $A_1^{(1)}$ subalgebra.

Nevertheless there exists a framework for simultaneously describing any choice of frozen variables [101]: it involves grouping them into *coefficients*,

$$y_i = \frac{\prod_{\text{arrows } \boxed{j} \rightarrow i} a_j}{\prod_{\text{arrows } \boxed{j} \leftarrow i} a_j}, \tag{93}$$

associated to the unfrozen variable they connect to, and defining mutation rules for them analogous to those of the \mathcal{A} - or \mathcal{X} coordinates, independently of their constituent frozen coordinates. The simplest case of *principal coefficients* amounts to $y_i = a_{d+i}$, $i = 1, \dots, d$, namely one frozen variable attached to every unfrozen one.

In [190], infinite $A_1^{(1)}$ mutation sequences with principal coefficients were analyzed as a proof of concept, see also [205]. References [61, 189] additionally found the generating functional of the mutation sequences for the particular case of frozen variables needed to embed them inside the $\text{Gr}(4, 8)$ cluster algebra. In [61] it was furthermore noticed that when embedding $A_1^{(1)}$ in a larger cluster algebra, it is possible to also take the direction of approach to the limit ray into account, so as to associate many square-root letters to each limit ray¹⁸. It was also proposed how to do this in a particular fashion, that was subsequently supported by a complementary approach based on scattering diagrams [192]. In this manner, one obtains a candidate eight-particle alphabet consisting of 272 rational and 18 square-root dual conformal invariant letters, which as a highly nontrivial check contains all those found in explicit computations of $A_{8,1}^{(2)}$ [143] and more recently $A_{8,0}^{(3)}$ [144].

Finally, infinite higher-rank $A_1^{(m)}$ mutation sequences with general coefficients were worked out in [62], and this provides the missing link for predicting finite symbol alphabets in principle at any multiplicity n . In the latter reference these general results were also specialized

¹⁷ A subalgebra of a cluster algebra is obtained by freezing, i.e. not mutating certain variables in one of its clusters. For the $A_1^{(1)}$ subalgebra of figure 12, this would be all variables but x_1, x_9 .

¹⁸ As with the selection of cluster variables/rational letters, also the limit ray must coincide with a $\text{pTr}(4, n)$ ray so as not to discard the associated square-root letters.

to the $n = 9$ case, yielding 3078 rational and 2349 square-root letters expected to appear in the amplitude. Support for the correctness of this proposal comes again from the fact that it contains the alphabet appearing in the independent determination of $A_{9,1}^{(2)}$ [142], as well as from the agreement of the entire rational part with an alternative proposal based on tensor diagrams [206]. The $n = 8$ and $n = 9$ alphabets are too long to quote here, but they may be found in the ancillary files of [62].

While resolving the longstanding issues we presented at the beginning of this subsection, the line of research we have described has also led to very interesting open questions that deserve further inquiry. For example, at $n = 9$ we have for the first time the appearance of square-root letters whose radicand does not correspond to that of the one-loop box (87), and it would thus be worthwhile to identify specific Feynman integrals giving rise to them. Perhaps more importantly, a new qualitative feature starting at this multiplicity is the existence of $\text{pTr}(4, n)$ rays which are inaccessible from the $\text{Gr}(4, n)$ cluster algebra even when enlarging the latter so as to also include limits of infinite mutation sequences. It is currently unclear if the missing rays are associated to more intricate algebraic letters beyond square roots, or point towards the need for more complicated, elliptic generalizations of MPLs starting to contribute at $n = 9$. Indeed, while it is known that such functions certainly appear at $n = 10$ [207], the possibility that these also appear at lower multiplicity cannot be excluded at the moment.

Even in the latter case, as we have stressed the key prerequisite for the bootstrap approach is the finiteness of the expected space of functions at each loop order, which is not necessarily restricted to MPLs. For example, the symbol calculus has been developed also for elliptic generalizations of MPLs [208], and its application on the elliptic double box in fact reveals that it is more similar to the non-elliptic case than previously expected [209]. The coaction and hence also the symbol contained in it has in fact been defined for even more general classes of periods and Feynman integrals [210, 211], see additionally the recent review [212] on Feynman integrals involving special functions beyond MPLs. Hence, also for quantities expressible in terms of these types of functions, what is needed to render them amenable to the bootstrap is a principle dictating a finite set of integration kernels contributing to them. In light of this, it would be very exciting to find a means to associate this kind of generalized symbol letters to the $\text{pTr}(4, n)$ missing rays, and explore their relevance for scattering amplitudes and Feynman integrals.

5.2. Bootstrapping Feynman integrals

While the analytic bootstrap approach to perturbative quantum field theory has been initiated and more extensively developed in the context of $\mathcal{N} = 4$ SYM amplitudes, that we have presented so far, the same methodology is applicable in many other situations as well. This in particular includes individual Feynman integrals that belong to the class of multiple polylogarithms defined in section 3.1, especially if they are *pure*, meaning that they have uniform weight, and their rational coefficients do not depend on the kinematics. Identifying such integrals is possible at the level of the integrand by examining their leading singularities [214], and if necessary modifying the integrals by taking out any overall factors, or by choosing the numerators appropriately, such that these leading singularities are constant [58].

5.2.1. Survey of explicit results. The constant leading singularity criterion was originally understood in the realm of $\mathcal{N} = 4$ SYM as well, and so the first pure weight integrals that were thus identified, and later on bootstrapped, were planar, finite and dual conformal in strictly four dimensions. These most notably contain the six-point double pentagon integral $\Omega^{(2)}$, that has already appeared in the construction of our hexagon function space in subsection 4.4, as

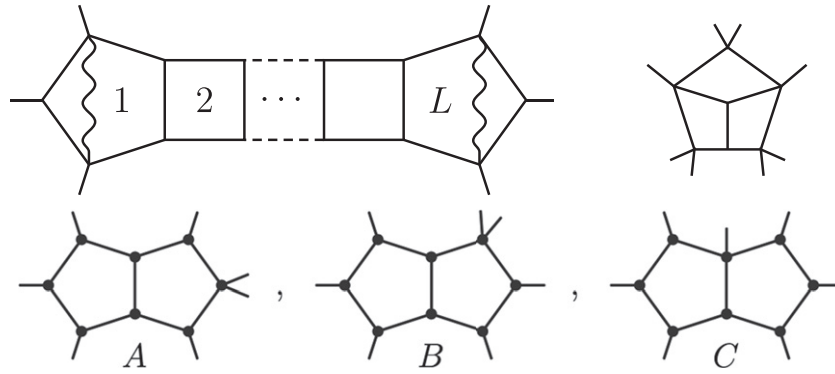


Figure 13. Examples of planar, four-dimensional finite dual conformal integrals that have been bootstrapped. Top left: six-point double pentagon ladder $\Omega^{(L)}$ [9, 31] (see also [33] for its generalization to massive left- and right-most legs through $L = 4$). Bottom: seven-point double pentagons [20]((B) and (C), the latter also at $L = 3$), [32] ((A) and (C)). Top right: eight-point $L = 3$ wheel [34]. Reproduced from [213]. CC BY 4.0.

well as its generalizations to higher loops or legs, some of which are depicted in figure 13¹⁹. The wavy lines denote numerators carefully chosen so as to render the integrals pure, see also chapter 7 [114] of the SAGEX review [42] for a concrete one-loop example. The fact that at consecutive loop orders these integrals are related by differential equations [216] or equivalently recursive integral representations [217], allows one to easily locate them inside the expected function space. Combined with integrability, bootstrap methods have also led to a closed form expression for a doubly infinite class of four-point fishnet integrals [218].

As with many computational tools first developed in the laboratory of the simplest gauge theory, the leading singularity analysis for finding pure weight integrals is generally applicable, and for integrands only having simple poles in the integration variables [219], it has in fact been automated in the `DLogBasis` package [220]. With the help of this analysis, the Feynman integral bootstrap has also been applied to the nonplanar cases relevant for massless five-point scattering shown in figure 14 [30], under the additional assumption that the corresponding symbol alphabet may be obtained from permutations of the known planar two-loop alphabet [49]. In order to fix an ansatz for the integral, additional information may be generically obtained by taking limits where it reduces to other simpler integrals that are already known or can be computed simply, either exactly or as series expansions, or by taking discontinuities that decrease the weight and hence also the complexity.

Finally, it is worth mentioning another type of bootstrap that is closely related to the differential equations obeyed by a basis of master integrals contributing to a particular process, and represented by the vector \mathbf{f} , with respect to the independent kinematic variables \vec{x} . If the integrals in question are expressible in terms of MPLs, then finding the transformation that brings the differential equations to canonical form [221]²⁰,

¹⁹ Many of these integrals may be also computed by direct methods, see for example [215] for $\Omega^{(L)}$ through $L = 10$, and [213] for the type A, B seven-point double pentagon ladders through $L = 4$, as well as their eight-point generalizations through $L = 3$ and $L = 2$, respectively.

²⁰ The notion of master integrals and the (canonical) differential equations they obey are discussed in more detail in chapter 3 [79] of the SAGEX review [42]. Different strategies for transforming differential equations to canonical form have been implemented in publicly available software such as `epsilon` [222], `Fuchsia` [223], `Canonica` [224], `Initial` [225] and `Libra` [226].

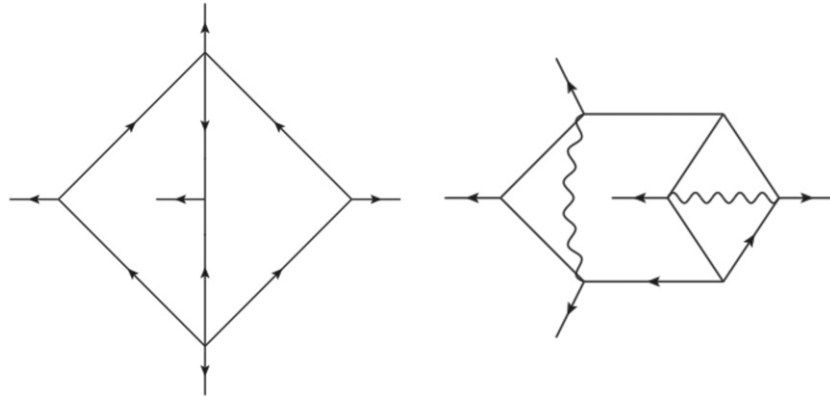


Figure 14. Nonplanar two-loop five-point massless integrals that have been bootstrapped up to $\mathcal{O}(\epsilon^0)$ in dimensional regularization. Reproduced from [30]. CC BY 4.0.

$$d\mathbf{f}(\vec{x}; \epsilon) = \epsilon \left[\sum_i \mathbf{A}_i d \log \phi_i(\vec{x}) \right] \mathbf{f}(\vec{x}; \epsilon), \tag{94}$$

where $d = \sum_j dx_j \partial_{x_j}$, ϕ_i are the letters and \mathbf{A}_i are constant matrices, is a very powerful method for algorithmically evaluating them in closed form to arbitrary order in the dimensional regulator $\epsilon = (4 - D)/2$. That is, when supplemented with boundary conditions, and provided that the letters are rational functions of the kinematic variables. If algebraic letters such as the square roots we saw in the previous section appear, under certain conditions variable transformations that rationalize them may also be found algorithmically [227], and have been implemented in the package `RationalizeRoots` [228]. Given that this however does not always work, an alternative presented in [35, 36], as also reviewed in the textbook [81], is to match \mathbf{f} and its derivatives to an ansatz built out of MPLs with the known alphabet, constructed by searching for admissible arguments of these MPLs. In particular, this search is guided by the requirement that the letters of the candidate MPLs should factorize over the known alphabet, and builds on analogous methods previously developed for rational alphabets [229].

5.2.2. The role of cluster algebras. It is important to note that a prior knowledge (or educated guess) of the symbol alphabet is either strictly necessary for the amplitude or integral bootstrap, or tremendously helpful for bringing the differential equations of master integrals to canonical form, as the remaining dependence of (94) on the purely numeric matrices \mathbf{A}_i can be determined much more easily. We have seen in subsections 3.3 and 5.1 that cluster algebras and their generalizations may give strong clues about the right alphabet, however until recently their appearance was confined to the realm of $\mathcal{N} = 4$ SYM theory. Excitingly, this changed with a recent publication [23], where it was discovered that cluster algebras underlie the analytic structure of a host of Feynman integrals in dimensional regularization. This most notably includes all known four-point integrals with one off-shell (or equivalently massive) leg, two-loop planar and nonplanar [230, 231], L -loop ladders [232, 233], and more recently the three-loop tennis court [234, 235], see also figure 15 for some examples. Specifically, the alphabet of this entire class of integrals is described by a C_2 cluster algebra.

In the remainder of this subsection, let us discuss how the latter connection was established, and what it implies. The C_2 cluster algebra is the first non-simply laced case we encounter, so given that it has no standard quiver representation, we will define it directly in terms of

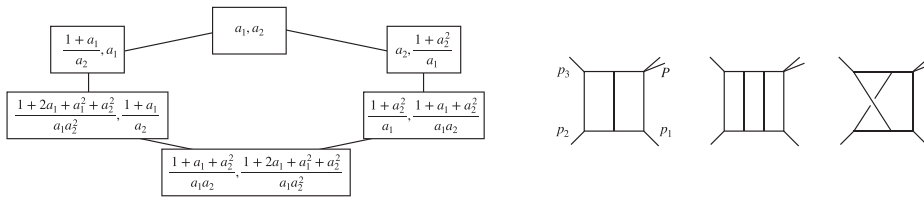


Figure 15. Left: the exchange graph of the C_2 cluster algebra, with cluster coordinates ordered as a_i, a_{i+1} . Right: examples of four-point one-mass integrals whose alphabet is given by the C_2 cluster coordinates. Reproduced from [23]. CC BY 4.0.

the exchange matrix B of its initial cluster. The elements b_{ij} of (the principal part of) B are essentially given by those of the associated Cartan matrix C by [99]

$$b_{ij} = (-1)^{\text{sign}(i-j)} (c_{ij} - \delta_{ij}c_{ii}), \tag{95}$$

therefore for the C_2 case without frozen variables we have,

$$C = \begin{pmatrix} 2 & -1 \\ -2 & 2 \end{pmatrix} \Rightarrow B = \begin{pmatrix} 0 & 1 \\ -2 & 0 \end{pmatrix}. \tag{96}$$

Mutating the initial \mathcal{A} -coordinates $\{a_1, a_2\}$ according to the rules (37) and (38), then yields in total six \mathcal{A} -coordinates, arranged in six clusters as depicted in the corresponding exchange graph on the left of figure 15.

One the other hand, the aforementioned integrals are expressible in terms of the well-studied class of two-dimensional harmonic polylogarithms (2dHPLs) [85, 230], which have the alphabet, in the dimensionless variables $z_1 = 2(p_1 \cdot p_2)/P^2$ and $z_2 = 2(p_2 \cdot p_3)/P^2$ where the momentum labels as shown in the above figure,

$$\Phi_{2\text{dHPL}} = \{z_1, z_2, z_3, 1 - z_1, 1 - z_2, 1 - z_3\}, \tag{97}$$

with $z_1 + z_2 + z_3 = 1$. Remarkably, the 2dHPL alphabet (97) is equivalent to the C_2 alphabet given in the left of figure 15, as can be readily verified by applying the variable transformation

$$z_1 = -\frac{a_2^2}{1+a_1}, \quad z_2 = -\frac{1+a_1+a_2^2}{a_1(1+a_1)}. \tag{98}$$

Transformations of this type between equivalent alphabets may be searched for systematically in a fashion analogous to the search for admissible MPL arguments, discussed a few paragraphs above. For example, the fact that z_1, z_2 are both variables and symbol letters implies that their logarithm should be a linear combination of logarithms of the \mathcal{A} -coordinate alphabet of figure 15, and that $1 - z_i$ should factorize over this alphabet.

It is very interesting to note that this cluster algebra connection also extends to a variety of processes in quantum chromodynamics, where the aforementioned integrals contribute as master integrals at two loops. These include for example three-jet production in electron–positron annihilation [236], as well as vector boson plus jet [237], or perhaps more importantly, Higgs boson plus jet production [238] in proton–proton collisions (in the massless and the heavy-top limit, respectively). As is also reviewed in chapter 1 [41] of the SAGEX review [42], the latter amplitude is in fact a three-particle *form factor*. These quantities are defined as vacuum expectation values of local operators between the vacuum and an n -particle external state, so

they are between the fully on-shell amplitudes and the fully off-shell correlators²¹. In the case at hand, the operator $H \text{Tr} F^2$, where H is the Higgs boson and F the gauge field strength, is the leading effective vertex in the effective field theory that arises when integrating out the top mass. The analogous quantity in $\mathcal{N} = 4$ SYM, which will be the focus of the next subsection, replaces this operator with any component of the stress-tensor multiplet of the theory, can also be expressed in terms of the same set of master integrals, and was first computed in [28].

Finally, we turn to the implications of the C_2 cluster algebra for the 2dHPL master integrals and the physical quantities they compute. We have seen in subsection 3.3 that in $\mathcal{N} = 4$ SYM theory cluster algebras additionally restrict which symbol letters can appear next to each other, could we hope for something similar here as well? Surprisingly, it turns out that the following subset of cluster adjacency restrictions hold,

$$\dots \otimes a_{2l+1} \otimes a_{2m+1} \otimes \dots \quad \Leftrightarrow \quad \dots \otimes 1 - z_l \otimes 1 - z_m \otimes \dots, \quad l \neq m, \tag{99}$$

to all orders in ϵ ! The same restrictions were also independently observed when bootstrapping the $\mathcal{N} = 4$ SYM three-particle form factor [29], and as in the case of amplitudes they considerably reduce the size of the corresponding function space.

Can we understand why only this subset and not all of the adjacencies occur? To this end, it is very suggestive that C_2 is the parity-invariant surface of the A_3 cluster algebra, relevant for six-particle scattering: this readily follows by inspecting their cluster polytopes, shown in figures 3 and 15²². In more detail, with the help of \mathcal{X} -coordinates it can be shown that the six nonvanishing A_3 letters of equation (40) (to avoid clash of notation, in the latter equation we switch $a_1 \rightarrow a, a_2 \rightarrow b, a_3 \rightarrow c$), are related to the C_2 cluster variables as

$$a_1 = \sqrt{a}, \quad a_3 = \sqrt{c}, \quad a_5 = \sqrt{b}, \quad a_{2i} = \sqrt{m_{3-i}}, \quad i = 1, 2, 3. \tag{100}$$

With this identification, we observe that the adjacency restrictions (99) precisely become the extended Steinmann relations for six-particle massless scattering (53)! This observation may point towards the right formulation of the (extended) Steinmann relations at lower multiplicity $n < 6$, which is currently not well understood²³. Further aspects of cluster-algebraic structures and their generalizations to Feynman integrals have been discussed in [27, 33, 34, 139, 240], and providing a first-principle derivation of their presence would be a very interesting goal for the future.

5.3. Bootstrapping a three-particle form factor

In the previous subsection, we saw that two-loop three-particle form factors are expressible in terms of the 2dHPL alphabet (97), whose equivalent description in terms of a C_2 cluster algebra makes plausible that the same holds true also at higher loops. In $\mathcal{N} = 4$ SYM theory, while the computation of the two-loop form factor of the stress tensor multiplet was carried out by unitarity methods, at the same time it was also shown that its symbol may also be uniquely

²¹ Protected correlators, their connection to scattering amplitudes in $\mathcal{N} = 4$ SYM, and related bootstrap approaches for computing them are also discussed in chapter 8 [239] of the SAGEX review [42].

²² This is an instance of a more general folding procedure, which carries over from Dynkin diagrams to cluster algebras, and allows one to embed B_n, C_n, F_4 and G_2 inside A_{2n-1}, D_{n+1}, E_6 and D_4 , respectively [102].

²³ We emphasize however that the adjacency conditions (99) do not correspond to discontinuities with respect to two-particle Mandelstam invariants.

determined by a bootstrap approach very much alike the one we have described for scattering amplitudes [28]. However what prevented the application of this approach at higher loops was the absence of enough independent information on the behavior of the form factor in kinematics limits, in order to uniquely identify it inside the expected space of functions. This limitation has been recently overcome, with the extension of the integrability-based pentagon OPE for predicting the near-collinear limit expansion of amplitudes and Wilson loops, mentioned in subsection 4.3, also to form factors [241–243]. As a result, the form factor in question has been computed through five loops in [29], and results through eight loops also appeared very recently [244].

While a detailed exposition of the three-particle form factor bootstrap would be out of the scope of this review, let us present some of the main features, to also illustrate the close similarity to the amplitudes case. The quantity of interest is the infrared-finite part of the form factor, which was originally obtained by factoring out the exponentiated one-loop form factor [28], in other words the corresponding BDS ansatz,

$$\mathcal{F}_3 = \mathcal{F}_3^{\text{BDS}} \exp[R]. \tag{101}$$

In this normalization, by convention the finite part was chosen to be represented by its logarithm, or remainder function R . However it turns out to be advantageous to pull out the finite kinematic-dependent part of the one-loop form factor, so as to define the BDS-like normalized form-factor E [29]²⁴,

$$E = e^{\frac{1}{4}\Gamma_{\text{cusp}}E^{(1)}+R}, \tag{102}$$

where its one-loop contribution is given by

$$E^{(1)} = -\sum_{i=1}^3 [2\text{Li}_2(1-z_i) + \ln z_i \ln z_{i+1}] + 4\zeta_2, \tag{103}$$

in terms of the kinematic variables we have defined above equation (97). Then, at weight 1 the space of functions F_1 containing E and its derivatives is dictated by the first entry condition,

$$F_1 = \{\log z_i\}, \quad i = 1, 2, 3, \tag{104}$$

and at higher weight it is constructed from the alphabet (97), ensuring that it obeys the integrability condition (59) or (60), the adjacency condition (99), which may be equivalently formulated as

$$F^{1-z_l, 1-z_m} = 0, \quad l \neq m, \tag{105}$$

as well a branch cut condition analogous to equation (62),

$$F^{1-z_i} \Big|_{\substack{z_i \rightarrow 1 \\ z_{i+1} \rightarrow 0}} = 0, \quad i = 1, 2, 3. \tag{106}$$

An important difference between amplitudes and form factors, however, is that in the latter case the branch cut condition also removes symbol-level functions.

²⁴ Further refinement of the normalization so as to reduce the number of independent constants is possible, but we will not describe this here.

Once the function space at the desired weight has been constructed, then an ansatz for E is uniquely fixed by imposing its full $z_i \rightarrow z_j$ permutation symmetry, the restriction on the

$$\text{final symbol entry of } E \in \frac{1 - z_i}{z_i}, \quad (107)$$

the fact that in the strict collinear limit in one orientation, e.g.

$$\lim_{z_2 \rightarrow 0} R = 0, \quad (108)$$

and finally constraints coming from the aforementioned near-collinear OPE.

A few final remarks are in order: first, in [28] the surprising observation was made, that when normalized in the same fashion, the maximal transcendental part of the leading-color term of the two-loop Higgs amplitude [238] discussed in the previous subsection, coincides with $R^{(2)}$ at symbol level (but not quite at function level [85]). It would be very interesting to find out if this agreement persists also at higher loops. Second, in the previous subsection we saw that the alphabet of the six-particle amplitude reduces to the alphabet of the form factor in the parity invariant surface. Quite remarkably, this kinematic relation also extends to the level of dynamics: from the existing data on both sides, it was observed that the MHV six-particle amplitude on the parity-even surface coincides with the form factor, up to certain variable substitution and the reversal of the order of letters in its symbol [245]! More precisely, the latter symbol-level order reversal also extends to functions up to factors of π , and is captured by the antipode operation on the Hopf algebra structure of MPLs, as is reviewed e.g. in [85]. It would be very interesting to understand the physical origin of this correspondence, and explore whether it persists for amplitudes and form factors with more legs or different MHV degree.

Acknowledgments

This work was supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No. 764850 'SAGEX'. We also acknowledge support from the Deutsche Forschungsgemeinschaft under Germany's Excellence Strategy EXC 2121 Quantum Universe 390833306.

Data availability statement

No new data were created or analyzed in this study.

ORCID iDs

Georgios Papathanasiou  <https://orcid.org/0000-0002-2627-9906>

References

- [1] White A R 2000 arXiv:[hep-ph/0002303](https://arxiv.org/abs/hep-ph/0002303)
- [2] Dixon L J, Drummond J M and Henn J M 2011 *J. High Energy Phys.* **JHEP11(2011)023**
- [3] Bern Z, Dixon L J and Smirnov V A 2005 *Phys. Rev. D* **72** 085001
- [4] Alday L F and Maldacena J 2007 *J. High Energy Phys.* **JHEP06(2007)064**

- [5] Alday L F and Maldacena J 2009 *J. High Energy Phys.* **JHEP11(2009)082**
- [6] Alday L F, Gaiotto D and Maldacena J 2011 *J. High Energy Phys.* **JHEP09(2011)032**
- [7] Alday L F, Maldacena J, Sever A and Vieira P 2010 *J. Phys. A: Math. Theor.* **43** 485401
- [8] Maldacena J 1998 *Adv. Theor. Math. Phys.* **2** 231–52
- [9] Dixon L J, Drummond J M and Henn J M 2012 *J. High Energy Phys.* **JHEP01(2012)024**
- [10] Dixon L J, Drummond J M, von Hippel M and Pennington J 2013 *J. High Energy Phys.* **JHEP12(2013)049**
- [11] Dixon L J, Drummond J M, Duhr C and Pennington J 2014 *J. High Energy Phys.* **JHEP06(2014)116**
- [12] Dixon L J and von Hippel M 2014 *J. High Energy Phys.* **JHEP10(2014)065**
- [13] Dixon L J, von Hippel M and McLeod A J 2016 *J. High Energy Phys.* **JHEP01(2016)053**
- [14] Caron-Huot S, Dixon L J, McLeod A and von Hippel M 2016 *Phys. Rev. Lett.* **117** 241601
- [15] Caron-Huot S, Dixon L J, Dulat F, von Hippel M, McLeod A J and Papathanasiou G 2019 *J. High Energy Phys.* **JHEP08(2019)016**
- [16] Drummond J M, Papathanasiou G and Spradlin M 2015 *J. High Energy Phys.* **JHEP03(2015)072**
- [17] Dixon L J, Drummond J, Harrington T, McLeod A J, Papathanasiou G and Spradlin M 2017 *J. High Energy Phys.* **JHEP02(2017)137**
- [18] Drummond J, Foster J, Gurdogan O and Papathanasiou G 2019 *J. High Energy Phys.* **JHEP03(2019)087**
- [19] Dixon L J and Liu Y T 2020 *J. High Energy Phys.* **JHEP10(2020)031**
- [20] Drummond J, Foster J and Gürdoğan Ö 2018 *Phys. Rev. Lett.* **120** 161601
- [21] Caron-Huot S, Dixon L J, Dulat F, Von Hippel M, McLeod A J and Papathanasiou G 2019 *J. High Energy Phys.* **JHEP09(2019)061**
- [22] Abreu S, Ita H, Moriello F, Page B, Tschernow W and Zeng M 2020 *J. High Energy Phys.* **JHEP11(2020)117**
- [23] Chicherin D, Henn J M and Papathanasiou G 2021 *Phys. Rev. Lett.* **126** 091603
- [24] Almelid Ø, Duhr C, Gardi E, McLeod A and White C D 2017 *J. High Energy Phys.* **JHEP09(2017)073**
- [25] Dixon L J, Duhr C and Pennington J 2012 *J. High Energy Phys.* **JHEP10(2012)074**
- [26] Chestnov V and Papathanasiou G 2021 *J. High Energy Phys.* **JHEP09(2021)007**
- [27] He S, Li Z, Tang Y and Yang Q 2021 *J. High Energy Phys.* **JHEP10(2021)084**
- [28] Brandhuber A, Travaglini G and Yang G 2012 *J. High Energy Phys.* **JHEP05(2012)082**
- [29] Dixon L J, McLeod A J and Wilhelm M 2021 *J. High Energy Phys.* **JHEP04(2021)147**
- [30] Chicherin D, Henn J and Mitev V 2018 *J. High Energy Phys.* **JHEP05(2018)164**
- [31] Caron-Huot S, Dixon L J, von Hippel M, McLeod A J and Papathanasiou G 2018 *J. High Energy Phys.* **JHEP07(2018)170**
- [32] Henn J, Herrmann E and Parra-Martinez J 2018 *J. High Energy Phys.* **JHEP10(2018)059**
- [33] He S, Li Z and Yang Q 2021 *J. High Energy Phys.* **JHEP12(2021)110**
- [34] He S, Li Z and Yang Q 2021 arXiv:2112.11842
- [35] Heller M, von Manteuffel A and Schabinger R M 2020 *Phys. Rev. D* **102** 016025
- [36] Heller M 2021 arXiv:2105.08046
- [37] Dixon L J, Drummond J M, Duhr C, von Hippel M and Pennington J 2014 *Proc. Sci.* **LL2014** 077
- [38] Caron-Huot S, Dixon L J, Drummond J M, Dulat F, Foster J, Gürdoğan O, von Hippel M, McLeod A J and Papathanasiou G 2020 *Proc. Sci.* **CORFU2019** 003 arXiv:2005.06735
- [39] Brink L, Schwarz J H and Scherk J 1977 *Nucl. Phys. B* **121** 77–92
- [40] Gliozzi F, Scherk J and Olive D 1977 *Nucl. Phys. B* **122** 253–90
- [41] Brandhuber A, Plefka J and Travaglini G 2022 arXiv:2203.13012
- [42] Travaglini G *et al* 2022 arXiv:2203.13011
- [43] 't Hooft G 1974 *Nucl. Phys. B* **72** 461
- [44] Dixon L J 2011 *J. Phys. A: Math. Theor.* **44** 454001
- [45] Srednicki M 2007 *Quantum Field Theory* (Cambridge: Cambridge University Press) <https://cds.cern.ch/record/1019751>
- [46] Nair V P 1988 *Phys. Lett. B* **214** 215–8
- [47] Elvang H, Freedman D Z and Kiermaier M 2010 *J. High Energy Phys.* **JHEP10(2010)103**
- [48] Zee A 2010 *Quantum Field Theory in a Nutshell* 2nd edn (Princeton, NJ: Princeton University Press)
- [49] Gehrman T, Henn J M and Lo Presti N A 2016 *Phys. Rev. Lett.* **116** 062001
Gehrman T, Henn J M and Lo Presti N A 2016 *Phys. Rev. Lett.* **116** 189903 (erratum)
- [50] Grisaru M T and Pendleton H N 1977 *Nucl. Phys. B* **124** 81–92

- [51] Drummond J M 2012 *Lett. Math. Phys.* **99** 481–505
- [52] Drummond J M, Henn J, Smirnov V A and Sokatchev E 2007 *J. High Energy Phys.* **JHEP01(2007)064**
- [53] Drummond J M, Korchemsky G P and Sokatchev E 2008 *Nucl. Phys. B* **795** 385–408
- [54] Drummond J M, Henn J, Korchemsky G P and Sokatchev E 2010 *Nucl. Phys. B* **826** 337–64
- [55] Drummond J, Henn J and Plefka J 2009 *J. High Energy Phys.* **JHEP05(2009)046**
- [56] Eden B, Heslop P, Korchemsky G P and Sokatchev E 2012 *Nucl. Phys. B* **862** 450–503
- [57] Hodges A 2013 *J. High Energy Phys.* **JHEP05(2013)135**
- [58] Arkani-Hamed N, Bourjaily J, Cachazo F and Trnka J 2012 *J. High Energy Phys.* **JHEP06(2012)125**
- [59] Bullimore M R 2013 Scattering amplitudes and Wilson loops in twistor space *PhD Thesis* Oxford University, New York
- [60] Speyer D and Williams L 2005 *J. Algebr. Comb.* **22** 189–210
- [61] Drummond J, Foster J, Gürdogan O and Kalousios C 2021 *J. High Energy Phys.* **JHEP04(2021)002**
- [62] Henke N and Papathanasiou G 2021 *J. High Energy Phys.* **10** 007
- [63] Golden J, Goncharov A B, Spradlin M, Vergu C and Volovich A 2014 *J. High Energy Phys.* **JHEP01(2014)091**
- [64] Arkani-Hamed N, Bourjaily J L, Cachazo F, Goncharov A B, Postnikov A and Trnka J 2016 *Grassmannian Geometry of Scattering Amplitudes* (Cambridge: Cambridge University Press)
- [65] Stermen G and Tejada-Yeomans M E 2003 *Phys. Lett. B* **552** 48–56
- [66] Beisert N, Eden B and Staudacher M 2007 *J. Stat. Mech.* **P01021**
- [67] Freyhult L 2012 *Lett. Math. Phys.* **99** 255–76
- [68] Chicherin D and Korchemsky G P 2022 arXiv:2203.13020
- [69] Alday L F and Roiban R 2008 *Phys. Rep.* **468** 153–211
- [70] Yang G 2011 *J. High Energy Phys.* **JHEP03(2011)087**
- [71] Papadopoulos C G, Tommasini D and Wever C 2016 *J. High Energy Phys.* **JHEP04(2016)078**
- [72] Canko D D, Papadopoulos C G and Syrrakos N 2021 *J. High Energy Phys.* **JHEP01(2021)199**
- [73] Abreu S, Ita H, Page B and Tschernow W 2021 arXiv:2107.14180
- [74] Arkani-Hamed N, Bourjaily J L, Cachazo F, Goncharov A B, Postnikov A and Trnka J 2016 *Grassmannian Geometry of Scattering Amplitudes* (Cambridge: Cambridge University Press) <http://cambridge.org/us/academic/subjects/physics/theoretical-physics/grassmannian-geometry-scattering-amplitudes?format=HB&isbn=9781107086586>
- [75] Brown F and Duhr C 2020 A double integral of dlog forms which is not polylogarithmic (arXiv:2006.09413)
- [76] Chen K-T 1977 *Bull. Am. Math. Soc.* **83** 831–79
- [77] Goncharov A B 1995 *Adv. Math.* **114** 197–318
- [78] Goncharov A B 1998 *Math. Res. Lett.* **5** 497–516
- [79] Abreu S, Britto R and Duhr C 2022 arXiv:2203.13014
- [80] Blümlein J and Schneider C 2022 arXiv:2203.13015
- [81] Weinzierl S 2022 arXiv:2201.03593
- [82] Goncharov A B 2001 arXiv:math/0103059
- [83] Goncharov A B 2005 *Duke Math. J.* **128** 209–84
- [84] Brown F 2012 *Adv. Stud. Pure Math.* **63** 31–58 (arXiv: 1102.130)
- [85] Duhr C 2012 *J. High Energy Phys.* **JHEP08(2012)043**
- [86] Goncharov A B 2009 arXiv:0908.2238
- [87] Goncharov A B, Spradlin M, Vergu C and Volovich A 2010 *Phys. Rev. Lett.* **105** 151605
- [88] Remiddi E and Vermaseren J A M 2000 *Int. J. Mod. Phys. A* **15** 725–54
- [89] Bauer C, Frink A and Kreckel R 2002 *J. Symb. Comput.* **33** 1–12
- [90] Maitre D 2012 *Comput. Phys. Commun.* **183** 846
- [91] Duhr C and Dulat F 2019 *J. High Energy Phys.* **JHEP08(2019)135**
- [92] Frellesvig H 2018 arXiv:1806.02883
- [93] Del Duca V, Duhr C and Smirnov V A 2010 *J. High Energy Phys.* **JHEP03(2010)099**
- [94] Del Duca V, Duhr C and Smirnov V A 2010 *J. High Energy Phys.* **JHEP05(2010)084**
- [95] Del Duca V, Duhr C and Smirnov V A 2011 *Phys. Lett. B* **703** 363–5
- [96] Dixon L J, Drummond J M and Henn J M 2011 *J. High Energy Phys.* **JHEP06(2011)100**
- [97] Scott J S 2003 arXiv:math/0311148
- [98] Fomin S and Zelevinsky A 2002 *J. Am. Math. Soc.* **15** 497–529
- [99] Fomin S and Zelevinsky A 2003 *Invent. Math.* **154** 63–121

- [100] Berenstein A, Fomin S and Zelevinsky A 2003 *Duke Math. J.* **126** 1–52
- [101] Fomin S and Zelevinsky A 2007 *Compos. Math.* **143** 112–64
- [102] Fomin S and Zelevinsky A 2001 arXiv:[hep-th/0111053](https://arxiv.org/abs/hep-th/0111053)
- [103] Gaiotto D, Moore G W and Neitzke A 2013 *Adv. Theor. Math. Phys.* **17** 241–397
- [104] Alim M, Cecotti S, Córdova C, Espahbodi S, Rastogi A and Vafa C 2013 *Commun. Math. Phys.* **323** 1185–227
- [105] Alim M, Cecotti S, Córdova C, Espahbodi S, Rastogi A and Vafa C 2014 *Adv. Theor. Math. Phys.* **18** 27–127
- [106] Arkani-Hamed N, Bai Y, He S and Yan G 2018 *J. High Energy Phys.* **JHEP05(2018)096**
- [107] Arkani-Hamed N, He S, Salvatori G and Thomas H 2019 arXiv:[1912.12948](https://arxiv.org/abs/1912.12948)
- [108] Vergu C 2015 Polylogarithm identities, cluster algebras and the $N = 4$ supersymmetric theory (arXiv:[1512.08113](https://arxiv.org/abs/1512.08113))
- [109] Drummond J, Foster J and Gürdoğan Ö 2018 arXiv:[1810.08149](https://arxiv.org/abs/1810.08149)
- [110] Feng B, Hanany A, He Y-H and Urra A M 2001 *J. High Energy Phys.* **JHEP12(2001)035**
- [111] Keller B 2012 arXiv:[1202.4161](https://arxiv.org/abs/1202.4161)
- [112] Felikson A, Shapiro M and Tumarkin P 2011 arXiv:[1111.3449](https://arxiv.org/abs/1111.3449)
- [113] Fock V V and Goncharov A B 2003 arXiv:[math/0311245](https://arxiv.org/abs/math/0311245)
- [114] Herrmann E and Trnka J 2022 arXiv:[2203.13018](https://arxiv.org/abs/2203.13018)
- [115] Caron-Huot S 2011 *J. High Energy Phys.* **JHEP12(2011)066**
- [116] Golden J and Spradlin M 2013 *J. High Energy Phys.* **JHEP09(2013)111**
- [117] Golden J, Paulos M F, Spradlin M and Volovich A 2014 *J. Phys. A: Math. Theor.* **47** 474005
- [118] Parker D E, Scherlis A, Spradlin M and Volovich A 2015 *J. High Energy Phys.* **JHEP11(2015)136**
- [119] Golden J and McLeod A J 2019 *J. High Energy Phys.* **JHEP01(2019)017**
- [120] Gekhtman M, Shapiro M and Vainshtein A 2002 arXiv:[math/0208033](https://arxiv.org/abs/math/0208033)
- [121] Golden J and Spradlin M 2014 *J. High Energy Phys.* **JHEP08(2014)154**
- [122] Golden J and Spradlin M 2015 *J. High Energy Phys.* **JHEP02(2015)002**
- [123] Golden J and McLeod A J 2021 *J. High Energy Phys.* **JHEP06(2021)142**
- [124] Harrington T and Spradlin M 2017 *J. High Energy Phys.* **JHEP07(2017)016**
- [125] Arkani-Hamed N, He S and Lam T 2021 *J. High Energy Phys.* **JHEP02(2021)069**
- [126] Arkani-Hamed N, He S, Lam T and Thomas H 2019 arXiv:[1912.11764](https://arxiv.org/abs/1912.11764)
- [127] Gates S J, Hazel Mak S N, Spradlin M and Volovich A 2021 arXiv:[2111.08186](https://arxiv.org/abs/2111.08186)
- [128] Steinmann O 1960 *Helv. Phys. Acta* **33** 257
- [129] Steinmann O 1960 *Helv. Phys. Acta* **33** 347
- [130] Cahill K E and Stapp H P 1975 *Ann. Phys.* **90** 438
- [131] Cutkosky R E 1960 *J. Math. Phys.* **1** 429–33
- [132] Papathanasiou G 2017 *Amplitudes* https://indico.ph.ed.ac.uk/event/26/contributions/342/attachments/284/319/Papathanasiou_Amplitudes2017.pdf
- [133] Mago J, Schreiber A, Spradlin M and Volovich A 2019 *J. High Energy Phys.* **JHEP10(2019)099**
- [134] Lippstreu L, Mago J, Spradlin M and Volovich A 2019 *J. High Energy Phys.* **JHEP09(2019)093**
- [135] Lukowski T, Parisi M, Spradlin M and Volovich A 2019 *J. High Energy Phys.* **JHEP10(2019)158**
- [136] Mago J, Schreiber A, Spradlin M and Volovich A 2021 *J. High Energy Phys.* **JHEP01(2021)084**
- [137] Abreu S, Dormans J, Febres Cordero F, Ita H and Page B 2019 *Phys. Rev. Lett.* **122** 082002
- [138] Golden J, McLeod A J, Spradlin M and Volovich A 2019 *J. High Energy Phys.* **JHEP03(2019)195**
- [139] He S, Li Z and Yang Q 2022 *J. High Energy Phys.* **JHEP01(2022)073**
- [140] Gaiotto D, Maldacena J, Sever A and Vieira P 2011 *J. High Energy Phys.* **JHEP12(2011)011**
- [141] Caron-Huot S and He S 2012 *J. High Energy Phys.* **JHEP07(2012)174**
- [142] He S, Li Z and Zhang C 2021 *J. High Energy Phys.* **JHEP03(2021)278**
- [143] He S, Li Z and Zhang C 2020 *Phys. Rev. D* **101** 061701
- [144] Li Z and Zhang C 2021 *J. High Energy Phys.* **JHEP12(2021)113**
- [145] Chen Z and Storjohann A 2005 A BLAS based C library for exact linear algebra on integer matrices *Proc. of the 2005 Int. Symp. on Symbolic and Algebraic Computation ISSAC '05* (New York: ACM) pp 92–9
- [146] Stein W and Joyner D 2005 *SIGSAM Bull.* **39** 61–4
- [147] The SpaSM Group 2017 SpaSM: a sparse direct solver modulo p v1.2 <http://github.com/cbouilla/spasm>
- [148] Mitev V and Zhang Y 2018 arXiv:[1809.05101](https://arxiv.org/abs/1809.05101)
- [149] Peraro T 2019 *J. High Energy Phys.* **JHEP07(2019)031**

- [150] Alday L F, Gaiotto D, Maldacena J, Sever A and Vieira P 2011 *J. High Energy Phys.* **JHEP04(2011)088**
- [151] Basso B, Sever A and Vieira P 2013 *Phys. Rev. Lett.* **111** 091602
- [152] Basso B, Sever A and Vieira P 2014 *J. High Energy Phys.* **JHEP01(2014)008**
- [153] Basso B, Sever A and Vieira P 2014 *J. High Energy Phys.* **JHEP08(2014)085**
- [154] Basso B, Sever A and Vieira P 2014 *J. High Energy Phys.* **JHEP09(2014)149**
- [155] Belitsky A V 2015 *Nucl. Phys. B* **896** 493–554
- [156] Belitsky A V 2015 *Nucl. Phys. B* **894** 108–35
- [157] Basso B, Caetano J, Cordova L, Sever A and Vieira P 2015 *J. High Energy Phys.* **JHEP08(2015)018**
- [158] Basso B, Caetano J, Cordova L, Sever A and Vieira P 2015 *J. High Energy Phys.* **JHEP12(2015)088**
- [159] Basso B, Sever A and Vieira P 2016 *J. Phys. A: Math. Theor.* **49** 41LT01
- [160] Belitsky A V 2017 *Nucl. Phys. B* **923** 588–607
- [161] Papathanasiou G 2013 *J. High Energy Phys.* **JHEP11(2013)150**
- [162] Papathanasiou G 2014 *Int. J. Mod. Phys. A* **29** 1450154
- [163] Drummond J M and Papathanasiou G 2016 *J. High Energy Phys.* **JHEP02(2016)185**
- [164] Cordova L 2017 *J. High Energy Phys.* **JHEP01(2017)051**
- [165] Lam H T and von Hippel M 2016 *J. High Energy Phys.* **JHEP12(2016)011**
- [166] Belitsky A V 2018 *Nucl. Phys. B* **929** 113–36
- [167] Belitsky A V 2018 *Phys. Lett. B* **780** 66–73
- [168] Bork L and Onishchenko A 2019 arXiv:1909.13675
- [169] Basso B, Dixon L J and Papathanasiou G 2020 *Phys. Rev. Lett.* **124** 161603
- [170] Del Duca V and Dixon L J 2022 arXiv:2203.13026
- [171] Bartels J, Lipatov L N and Sabio Vera A 2010 *Eur. Phys. J. C* **65** 587–605
- [172] Fadin V S and Lipatov L N 2012 *Phys. Lett. B* **706** 470–6
- [173] Bartels J, Kormilitzin A, Lipatov L N and Prygarin A 2012 *Phys. Rev. D* **86** 065026
- [174] Lipatov L, Prygarin A and Schnitzer H J 2013 *J. High Energy Phys.* **JHEP01(2013)068**
- [175] Bartels J, Kormilitzin A and Lipatov L 2014 *Phys. Rev. D* **89** 065002
- [176] Del Duca V, Druc S, Drummond J, Duhr C, Dulat F, Marzucca R, Papathanasiou G and Verbeek B 2016 *J. High Energy Phys.* **JHEP08(2016)152**
- [177] Del Duca V, Druc S, Drummond J, Duhr C, Dulat F, Marzucca R, Papathanasiou G and Verbeek B 2018 *J. High Energy Phys.* **JHEP06(2018)116**
- [178] Basso B, Caron-Huot S and Sever A 2015 *J. High Energy Phys.* **JHEP01(2015)027**
- [179] Del Duca V, Druc S, Drummond J M, Duhr C, Dulat F, Marzucca R, Papathanasiou G and Verbeek B 2020 *Phys. Rev. Lett.* **124** 161602
- [180] Dixon L J, Liu Y-T and Miczajka J 2021 *J. High Energy Phys.* **JHEP12(2021)218**
- [181] Dixon L and Dulat F unpublished.
- [182] Brown F C S 2009 *Ann. Sci. École Norm. Sup.* **42** 371
- [183] Bogner C and Brown F 2015 *Commun. Numer. Theor. Phys.* **9** 189–238
- [184] Ablinger J, Blümlein J, Raab C, Schneider C and Wißbrock F 2014 *Nucl. Phys. B* **885** 409–47
- [185] Anastasiou C, Duhr C, Dulat F and Mistlberger B 2013 *J. High Energy Phys.* **JHEP07(2013)003**
- [186] Panzer E 2015 *Comput. Phys. Commun.* **188** 148–66
- [187] Bogner C 2016 *Comput. Phys. Commun.* **203** 339–53
- [188] Bourjaily J L, Caron-Huot S and Trnka J 2015 *J. High Energy Phys.* **JHEP01(2015)001**
- [189] Arkani-Hamed N, Lam T and Spradlin M 2021 *J. High Energy Phys.* **JHEP03(2021)065**
- [190] Henke N and Papathanasiou G 2020 *J. High Energy Phys.* **JHEP08(2020)005**
- [191] Speyer D and Sturmfels B 2003 arXiv:math/0304218
- [192] Herderschee A 2021 *J. High Energy Phys.* **JHEP07(2021)049**
- [193] Mago J, Schreiber A, Spradlin M and Volovich A 2020 *J. High Energy Phys.* **JHEP10(2020)128**
- [194] He S and Li Z 2021 *J. High Energy Phys.* **JHEP02(2021)155**
- [195] Mago J, Schreiber A, Spradlin M, Yelleshpur Srikant A and Volovich A 2021 *J. High Energy Phys.* **JHEP09(2021)002**
- [196] Mago J, Schreiber A, Spradlin M, Srikant A Y and Volovich A 2021 *J. High Energy Phys.* **JHEP04(2021)056**
- [197] Drummond J, Foster J, Gürdoğan O and Kalousios C 2021 *J. High Energy Phys.* **JHEP11(2021)071**
- [198] Nakanishi T and Zelevinsky A 2011 arXiv:1101.3736
- [199] Cachazo F, Early N, Guevara A and Mizera S 2019 *J. High Energy Phys.* **JHEP06(2019)039**
- [200] Cachazo F, He S and Yuan E Y 2014 *Phys. Rev. Lett.* **113** 171601
- [201] Cachazo F, He S and Yuan E Y 2014 *J. High Energy Phys.* **JHEP07(2014)033**

- [202] Drummond J, Foster J, Gürdoğan Ö and Kalousios C 2020 *J. High Energy Phys.* **JHEP04(2020)146**
- [203] Gawrilow E and Joswig M 2000 polymake: a framework for analyzing convex polytopes *Polytopes—Combinatorics and Computation (Oberwolfach, 1997) (DMV Seminar)* vol 29 (Basel: Birkhäuser) pp 43–73
- [204] Canakci I and Schiffler R 2018 *Compos. Math.* **154** 565–93
- [205] Reading N 2018 arXiv:1806.05094
- [206] Ren L, Spradlin M and Volovich A 2021 *J. High Energy Phys.* **JHEP12(2021)079**
- [207] Caron-Huot S and Larsen K J 2012 *J. High Energy Phys.* **JHEP10(2012)026**
- [208] Broedel J, Duhr C, Dulat F, Penante B and Tancredi L 2018 *J. High Energy Phys.* **JHEP08(2018)014**
- [209] Kristensson A, Wilhelm M and Zhang C 2021 *Phys. Rev. Lett.* **127** 251603
- [210] Brown F 2017 *Commun. Numer. Theor. Phys.* **11** 453–556
- [211] Brown F 2015 arXiv:1512.06410
- [212] Bourjaily J L *et al* 2022 Functions beyond multiple polylogarithms for precision collider physics *2022 Snowmass Summer Study*
- [213] Bourjaily J L, McLeod A J, von Hippel M and Wilhelm M 2018 *J. High Energy Phys.* **JHEP08(2018)184**
- [214] Cachazo F 2008 arXiv:0803.1988
- [215] McLeod A J, Munch H J, Papathanasiou G and von Hippel M 2020 *J. High Energy Phys.* **JHEP11(2020)122**
- [216] Drummond J M, Henn J M and Trnka J 2011 *J. High Energy Phys.* **JHEP04(2011)083**
- [217] He S, Li Z, Tang Y and Yang Q 2021 *J. High Energy Phys.* **JHEP05(2021)052**
- [218] Basso B and Dixon L J 2017 *Phys. Rev. Lett.* **119** 071601
- [219] Arkani-Hamed N, Bourjaily J L, Cachazo F and Trnka J 2014 *Phys. Rev. Lett.* **113** 261603
- [220] Henn J, Mistlberger B, Smirnov V A and Wasser P 2020 *J. High Energy Phys.* **JHEP04(2020)167**
- [221] Henn J M 2013 *Phys. Rev. Lett.* **110** 251601
- [222] Prausa M 2017 *Comput. Phys. Commun.* **219** 361–76
- [223] Gituliar O and Magerya V 2017 *Comput. Phys. Commun.* **219** 329–38
- [224] Meyer C 2018 *Comput. Phys. Commun.* **222** 295–312
- [225] Dlapa C, Henn J and Yan K 2020 *J. High Energy Phys.* **JHEP05(2020)025**
- [226] Lee R N 2021 *Comput. Phys. Commun.* **267** 108058
- [227] Besier M, Van Straten D and Weinzierl S 2019 *Commun. Numer. Theor. Phys.* **13** 253–97
- [228] Besier M, Wasser P and Weinzierl S 2020 *Comput. Phys. Commun.* **253** 107197
- [229] Duhr C, Gangl H and Rhodes J R 2012 *J. High Energy Phys.* **JHEP10(2012)075**
- [230] Gehrmann T and Remiddi E 2001 *Nucl. Phys. B* **601** 248–86
- [231] Gehrmann T and Remiddi E 2001 *Nucl. Phys. B* **601** 287–317
- [232] Di Vita S, Mastroliia P, Schubert U and Yundin V 2014 *J. High Energy Phys.* **JHEP09(2014)148**
- [233] Panzer E 2015 Feynman integrals and hyperlogarithms *PhD Thesis* Humboldt University
- [234] Canko D D, Gasparotto F, Mattiazzi L, Papadopoulos C G and Syrrakos N 2021 N^3LO calculations for $2 \rightarrow 2$ processes using simplified differential equations *15th Int. Symp. on Radiative Corrections: Applications of Quantum Field Theory to Phenomenology and LoopFest XIX: Workshop on Radiative Corrections for the LHC and Future Colliders*
- [235] Canko D D and Syrrakos N 2021 arXiv:2112.14275
- [236] Garland L W, Gehrmann T, Glover E W N, Koukoutsakis A and Remiddi E 2002 *Nucl. Phys. B* **642** 227–62
- [237] Gehrmann T, Tancredi L and Weihs E 2013 *J. High Energy Phys.* **JHEP04(2013)101**
- [238] Gehrmann T, Jaquier M, Glover E W N and Koukoutsakis A 2012 *J. High Energy Phys.* **JHEP02(2012)056**
- [239] Heslop P 2022 arXiv:2203.13019
- [240] He S, Li Z and Yang Q 2021 *J. High Energy Phys.* **JHEP06(2021)119**
- [241] Sever A, Tumanov A G and Wilhelm M 2021 *Phys. Rev. Lett.* **126** 031602
- [242] Sever A, Tumanov A G and Wilhelm M 2021 *J. High Energy Phys.* **JHEP10(2021)071**
- [243] Sever A, Tumanov A G and Wilhelm M 2021 arXiv:2112.10569
- [244] Dixon L J, Gurdogan O, McLeod A J and Wilhelm M 2022 arXiv:2204.11901
- [245] Dixon L J, Gurdogan O, McLeod A J and Wilhelm M 2021 arXiv:2112.06243