

Getting more flavor out of one-flavor QCD

Tom Melia*

CERN Theory Division, CH-1211 Geneva 23, Switzerland

(Received 17 January 2014; published 3 April 2014)

We argue that no notion of flavor is necessary when performing amplitude calculations in perturbative QCD with massless quarks. We show this explicitly at tree level, using a flavor recursion relation to obtain multiflavored QCD from one-flavor QCD. The method relies on performing a color decomposition, under which the one-flavor primitive amplitudes have a structure which is restricted by planarity and cyclic ordering. An understanding of $SU(3)_c$ group theory relations between QCD primitive amplitudes and their organization around the concept of a Dyck tree is also necessary. The one-flavor primitive amplitudes are effectively $\mathcal{N} = 1$ supersymmetric, and a simple consequence is that all of tree-level massless QCD can be obtained from Drummond and Henn's closed form solution to tree-level $\mathcal{N} = 4$ super Yang-Mills theory.

DOI: 10.1103/PhysRevD.89.074012

PACS numbers: 12.38.-t, 11.15.-q, 12.38.Bx

I. INTRODUCTION

Fixed-order theoretical predictions for jet cross sections at the Large Hadron Collider (LHC) require the calculation of scattering amplitudes in perturbative QCD involving light QCD partons—gluons and quarks of different flavor. For inclusive quantities, all possible partonic contributions must be computed and summed; for example, for 6 jet production at tree level, amplitudes for $gg \rightarrow gggggg$, $d\bar{d} \rightarrow ggggu\bar{u}$, $gg \rightarrow u\bar{u}u\bar{u}u\bar{u}$, $gg \rightarrow d\bar{d}s\bar{s}c\bar{c}$, etc. are all needed.

What is the difference between the amplitudes describing $gg \rightarrow u\bar{u}u\bar{u}u\bar{u}$ and $gg \rightarrow d\bar{d}s\bar{s}c\bar{c}$ in pure QCD, in the limit that all quarks can be treated as massless? Since all flavors of quark have the same color interactions, the only difference comes from the restriction that flavor conservation imposes on the possible factorization channels of the amplitudes. The amplitudes for the one-flavor case can be obtained from the amplitudes for the distinct-flavor case via a permutation sum over quark indices,

$$\mathcal{M}(gg \rightarrow u\bar{u}u\bar{u}u\bar{u}) = \sum_{\mathcal{P}(d,s,c)} (-1)^{\text{sgn}\mathcal{P}} \mathcal{M}(gg \rightarrow d\bar{d}s\bar{s}c\bar{c}), \quad (1)$$

where \mathcal{P} gives permutations of momentum and color indices of the quarks, and the $(-1)^{\text{sgn}\mathcal{P}}$ accounts for Fermi statistics. For this reason, it is taken that the distinct-flavor case is the more general one which should be calculated.

In this paper we show that it is possible to reverse Eq. (1) for massless QCD amplitudes at tree level—more specifically, for tree-level primitive amplitudes. That is, we can recover the distinct-flavor case using primitives involving only one flavor of quark line. The result follows from an understanding of $SU(3)_c$ group theory relations between the QCD primitives, and a generalization of the Dyck basis

for quark amplitudes found in Ref. [1]. It means that all of tree-level massless QCD can be rendered effectively $\mathcal{N} = 1$ supersymmetric and can as such be obtained from the known solution to $\mathcal{N} = 4$ SYM at tree level. It is interesting that no notion of flavor is needed at the level of the field theory amplitude calculation (we will argue this is true at any loop order, using unitarity based methods to obtain loop amplitudes from tree-level amplitudes); rather it can be reinstated through combinatorics of the one-flavor tree-level amplitudes alone.

A crucial step is to use a color decomposition to define primitive amplitudes. By considering a theory where all particles, including quarks, are in the adjoint representation, they can be defined at tree level as [2,3]

$$\mathcal{M}^{\text{tree}} = \sum_{\sigma \in S_{n-1}} \text{tr}(\lambda^1 \lambda^{\sigma_1} \dots \lambda^{\sigma_{n-1}}) \mathcal{A}(1\sigma_1 \dots \sigma_{n-1}), \quad (2)$$

where the λ^a denote $SU(3)_c$ fundamental representation matrices, and the \mathcal{A} denote purely kinematic primitive amplitudes. Since only the color part of the quark interactions change when putting them into the adjoint representation, the purely kinematic primitive amplitudes are the same as in usual QCD, so we shall refer to them simply as QCD primitives—algorithms exist to relate these primitives back to the full amplitude with fundamental quarks [4–7] (at both tree and one-loop levels). The primitive amplitudes inherit a number of properties from the color decomposition [Eq. (2)]. They are gauge invariant and, with reference to their Feynman diagram representation, they only receive contributions from graphs (once drawn in a planar fashion) with a cyclic ordering of the external legs that is the same as the labeling of the primitive, giving them a simplified kinematic structure. The color decomposition with all particles in the adjoint representation makes it clear that general QCD primitive amplitudes satisfy the same group theory relations as do all-gluon amplitudes—these are known as Kleiss-Kuijf (KK) relations [8]. However, as

*thomas.melia@cern.ch

was explained in [1], the quark lines present in the amplitude impart further structure on these relations, modifying them with the effect that fewer than $(n-2)!$ primitives are independent for n particle scattering. In the pure quark and antiquark case considered, with all-distinct-flavor quark lines, the number of independent primitives was shown to be $(n-2)!/(n/2)!$. In this paper we generalize this statement to QCD primitive amplitudes with any number of gluons as well as quarks. Independence for most of this paper will be taken to mean independent over the field of real numbers [Bern-Carrasco-Johansson (BCJ) relations [9] take into account the further possibility of multiplying primitive amplitudes by kinematic invariants].

In one of the landmarks of the recent progress in the uncovering of the structure of scattering amplitudes, a solution to tree-level $\mathcal{N} = 4$ SYM was written down by Drummond and Henn [10], which provides a closed form for the primitive amplitudes of this theory. That such a remarkable formula exists is made possible by the high degree of symmetry present in the theory. But it has long been understood that some amplitudes in QCD are effectively supersymmetric at tree level [11,12], and, in a recent paper [13], it was shown that all QCD amplitudes with up to four quark lines of distinct flavor (four also being the number of gluino flavors in $\mathcal{N} = 4$ SYM) can be obtained from the formula of Drummond and Henn. In this paper we show that in fact all of massless QCD can be obtained from $\mathcal{N} = 4$ SYM, even though there are only four flavors of gluino in $\mathcal{N} = 4$ SYM (seemingly less than what is needed to describe QCD) as well as possible contributions from scalar particle exchange (seemingly more than what is needed to describe QCD).

A main motivation for the study of these amplitudes is to be able to make fixed-order theoretical predictions with which to compare to collider data. The ATLAS and CMS experiments have presented data with up to as many as 10 hard QCD jets [14,15], and this presents an enormous challenge for such computations (at around this number of jets, they are also expected to break down). At leading order (LO) in QCD, calculations with up to 8 jets are available [16], and there are next-to-leading order (NLO) descriptions of processes involving up to 5 jets [17,18]. QCD tree-level primitive amplitudes can be considered as fundamental gauge-invariant building blocks of such calculations—as well as describing tree-level contributions, they arise naturally in unitarity based methods for obtaining the one-loop part of the NLO corrections to jet cross sections at the LHC [19–26] (see e.g. [4] for a review), and also in calculations of QCD amplitudes at higher loop orders [27–33]. Primitive amplitudes can also be used explicitly in subtraction schemes, as formulated within the colorful [34] Frixione-Kunszt-Signer framework [35]. Both a knowledge of a general basis and the flavor recursion described in this paper should be useful for multileg QCD calculations at leading and next-to-leading orders.

Although, for inclusive light jets, multi-quark contributions are less important numerically than the more gluonic ones, more exclusive multi-quark final states can be experimentally defined by tagging the quark-like contributions. Bottom quarks (and, with less efficiency, charm quarks) can be identified by displaced vertices—as many as 4 b -tags are used in ongoing experimental analyses—and jet substructure techniques [36,37] can statistically distinguish between jets originating from a light quark and those originating from gluons [38–41], and are increasingly being used, for example, in searches for supersymmetric particle hadronic cascade decays. The definition of jet flavor can be made infrared safe for calculations at the parton level [42]. For LHC phenomenology, the addition of electroweak bosons into amplitudes, as well as considering massive top (and possibly bottom) quarks, is important.

The outline of this paper is as follows. Sec. 2 describes the construction of a general basis of QCD primitive amplitudes for $g^{n-2k}(\bar{q}q)^k$ scattering with all-distinct quark flavors, which is of the size $(n-2)!/k!$. The notion of a rooted oriented Dyck tree is introduced and the connection with quark line structure is made. Sec. 3 presents the flavor recursion, which enables a k -flavor primitive to be expressed in terms of one-flavor primitives, and discusses how $\mathcal{N} = 4$ SYM amplitudes can be used to completely specify all of tree-level massless QCD. We discuss moving away from massless QCD, the role of BCJ relations, and further directions in Sec. 4, and we conclude in Sec. 5.

II. A GENERAL TREE-LEVEL QCD PRIMITIVE BASIS

The use of Dyck words in understanding the number of independent primitives for purely multi-quark scattering with distinct flavors of quarks was presented in [1]. In this section we will generalize these results to QCD primitives of the form $g^{n-2k}(\bar{q}q)^k$, i.e. with k distinct-flavor quark lines and $n-2k$ gluons. Throughout this section, we will use the convention that the quarks and antiquarks are labeled by the numbers $1\dots 2k$, with antiquarks given odd labels, quarks given even labels, and flavor pairs labeled by consecutive numbers: $(1, 2)$, $(3, 4)$, ..., $(2k-1, 2k)$. The gluons are labeled by the numbers $2k+1\dots n$. There is no loss of generality in using this convention for the purposes of this section. It will, however, be necessary to introduce additional notation in the following section to distinguish different flavor pairings.

We will first state the result: for k -flavor quark QCD primitive amplitudes $g^{n-2k}(\bar{q}q)^k$, a basis of primitives is the set

$$\mathcal{A}(1\dots\sigma_1\dots\sigma_2\dots\dots\sigma_{2k-2}\dots 2), \quad (3)$$

where the \dots stand for all possible insertions of the gluons $2k+1, \dots, n$ in between the labels $\sigma = \{\sigma_1, \dots, \sigma_{2k-2}\}$

which are ‘‘Dyck permutations’’ of the set of quark indices $\{3, 4, 5, 6, \dots, 2k - 1, 2k\}$.

The Dyck permutations needed are obtained as follows (each step is described in detail below): (i) consider all Dyck topologies arising from Dyck words of length $2k - 2$, (ii) for each Dyck topology, consider all possible assignments of flavor pairs, and (iii) for each flavor pair assignment, include only one orientation of each of the flavor pairs.

The Dyck topologies needed for step (i) are obtained from Dyck words. Dyck words of length $2k - 2$ are strings of $k - 1$ letter X s and $k - 1$ letter Y s with the requirement that the number of X s is always greater than or equal to the number of Y s in any initial segment of the string. For each of these words, the Dyck topology is obtained by identifying pairs of (X_i, Y_i) in the following way: reading the word left to right, pairing each Y that you come across with the closest unpaired X to the left. There are $k - 1$ pairs constructed in this way for each Dyck word, labeled with $i = 1, \dots, k - 1$.

The different flavor pair assignments needed for step (ii) are simply the $(k - 1)!$ different possible ways of assigning the flavor pairs $(3,4), (5,6)$, etc. to the $k - 1$ Dyck topology pairs (X_i, Y_i) . This is done for each Dyck topology.

Finally, in step (iii) above, the Dyck permutations are obtained for each of these flavor pair assignments by choosing the orientation of the flavor pair. This means placing the label of each quark and antiquark pair into the Dyck word either as $\bar{q} \rightarrow X_i, q \rightarrow Y_i$, or $q \rightarrow X_i, \bar{q} \rightarrow Y_i$, where (\bar{q}, q) is whichever flavor pair is assigned to the Dyck topology pair (X_i, Y_i) . Only one orientation of each quark line is included for each of the flavor assignments arising from the Dyck topologies. We can introduce the concept of the *signature* of the permutation, which is a string of $k - 1$ $+/-$ signs depending on whether (\bar{q}, q) is assigned as $\bar{q} \rightarrow X, q \rightarrow Y$ (a $+$ sign) or $q \rightarrow X, \bar{q} \rightarrow Y$ (a $-$ sign). So we obtain a vector (\pm, \pm, \dots, \pm) , where the i th entry corresponds to the orientation for the pair (X_i, Y_i) .

As an example, for $k = 3$, there are two Dyck words of the length $4 (= 2k - 2)$: $XYXY, XXYY$. For the first Dyck word, the pair assignment is $X_1Y_1X_2Y_2$. For the flavor assignments $(3, 4) \rightarrow (X_1Y_1), (5, 6) \rightarrow (X_2Y_2)$, the permutations

$$(3, 4, 5, 6), (4, 3, 5, 6), (3, 4, 6, 5), (4, 3, 6, 5) \quad (4)$$

are obtained with signatures $(+, +), (-, +), (+, -), (-, -)$, respectively; from the second possible flavor assignment $(5, 6) \rightarrow (X_1Y_1), (3, 4) \rightarrow (X_2Y_2)$, the permutations

$$(5, 6, 3, 4), (6, 5, 3, 4), (5, 6, 4, 3), (6, 5, 4, 3) \quad (5)$$

are obtained with signatures $(+, +), (-, +), (+, -), (-, -)$, respectively. For the second Dyck word the pair assignment is $X_2X_1Y_1Y_2$. For the flavor assignments $(3, 4) \rightarrow (X_1Y_1), (5, 6) \rightarrow (X_2Y_2)$, the permutations

$$(5, 3, 4, 6), (5, 4, 3, 6), (6, 3, 4, 5), (6, 4, 3, 5) \quad (6)$$

are obtained with signatures $(+, +), (-, +), (+, -), (-, -)$, respectively; from the second possible flavor assignment $(5, 6) \rightarrow (X_1Y_1), (3, 4) \rightarrow (X_2Y_2)$, the permutations

$$(3, 5, 6, 4), (3, 6, 5, 4), (4, 5, 6, 3), (4, 6, 5, 3) \quad (7)$$

are obtained with signatures $(+, +), (-, +), (+, -), (-, -)$, respectively. A basis then consists of four permutations, one chosen from each of Eqs. (4)–(7). The four permutations do not have to have the same signature as each other.

A diagram showing the quark line structure of the primitive can be drawn for each of these permutations; see Fig. 1. A quark line is drawn between each of the identified pairs (X_i, Y_i) . The $(k - 1)!$ flavor pair allocations change the flavor of these quark lines, and the 2^{k-1} different signatures for each of these allocations are all possible ways of drawing the direction of the arrow on each of the quark lines. Gluons can be inserted in any position between the quark lines, so at all points around the circle, except between 1 and 2 which have been fixed using the KK relations as always consecutive, as described below. These gluons are not drawn on the quark line graphs. A particularly useful feature of these graphs is that if the quark lines cross then the primitive has to vanish, since with only planar contributions, the quark lines are forced to intersect, which results in a flavor-violating vertex.

This is a generalization of what was found in [1] in two directions. There is the addition of gluons, and also the freedom in choosing the signature of the Dyck permutations describing the quark line structure. We now discuss each of these generalizations in turn.

A. Addition of gluons

Considering first just the quark line structure, the number of permutations σ is given by the number of Dyck words of length $2(k - 1)$, which is given by the Catalan number $C_r = (2r)!/r!(r + 1)!$, with $r = k - 1$, multiplied by the number of flavor allocations, $(k - 1)!$. Next, the counting of all possible insertions of the gluons can be done by considering all possible distributions of an ordered set of the $p = n - 2k$ gluons into $2k - 1$ slots between the quark lines [excluding between $(1 \rightarrow 2)$], which is given by a binomial coefficient $(2k - 2 + p)!/(2k - 2)!/p!$, and then considering all ordered sets, given by the $p!$ possible gluon permutations. That is, the number of independent k -flavor $g^{n-2k}(\bar{q}q)^k$ QCD primitive amplitudes is

$$\begin{aligned} & (\#\text{Dyck words})(\#\text{flavor pair allocations}) \\ & (\#\text{gluon distributions})(\#\text{gluon permutations}) \\ &= \frac{(2k - 2)!}{(k - 1)!k!} (k - 1)! \frac{(p + 2k - 2)!}{(2k - 2)!p!} p! \\ &= \frac{(p + 2k - 2)!}{k!} = \frac{(n - 2)!}{k!}. \end{aligned} \quad (8)$$

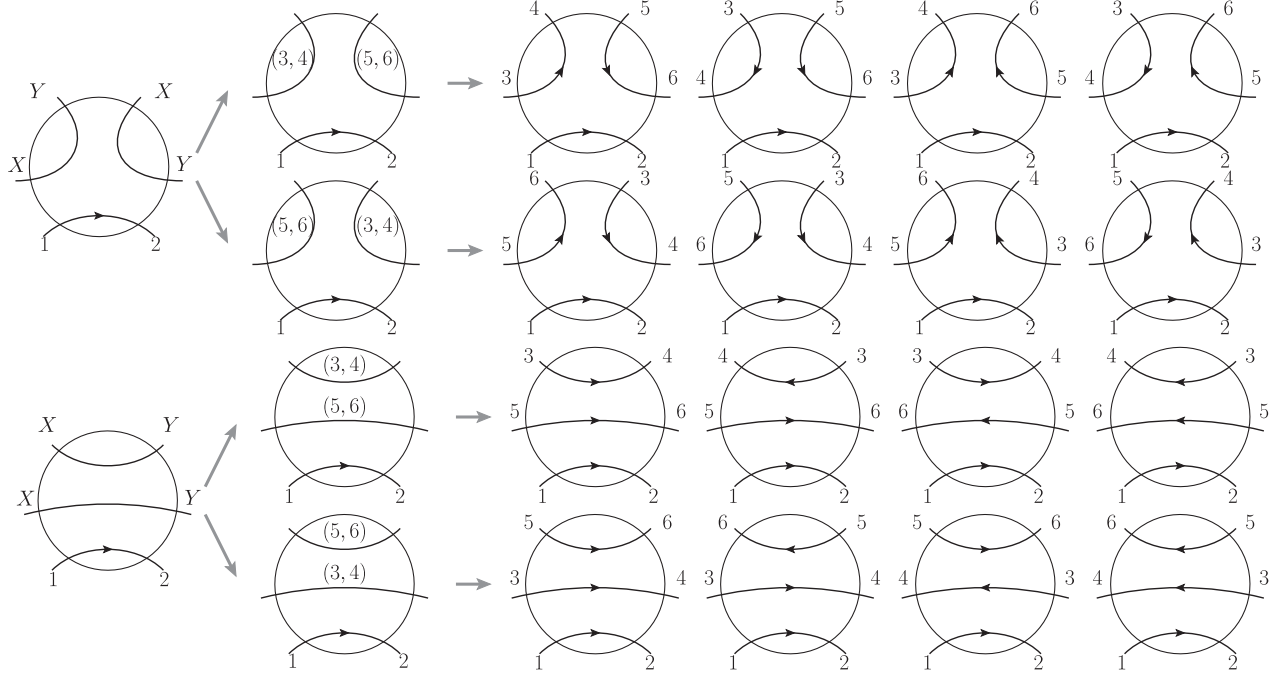


FIG. 1. Constructing quark line graphs based around the Dyck words $XYXY$ and $XXYY$. The first column shows the two Dyck topologies, and the second column shows the two possible flavor pair allocations for each topology. Following the direction of the arrows, each row then depicts the four possible choices of signature for each flavor pair allocation. Four permutations corresponding to one graph out of each of these rows constitute the Dyck permutations which are then used to construct a basis of QCD primitives.

This generalization coming from the addition of gluons to amplitude is straightforward, since they impart no further structure which can affect KK relations further than beyond the impact of the quark line structure. The KK relations allow for two labels to be fixed in consecutive order—as in Ref. [1] we chose here to fix 1 and 2 to be

cyclically next to each other (i.e. we always consider primitive amplitudes of the form $\mathcal{A}(1\dots 2)$). All of the further information concerning relations between the remaining $(n - 2)!$ primitives (after fixing 1 and 2) is contained in the quark line structure, to which we now turn.

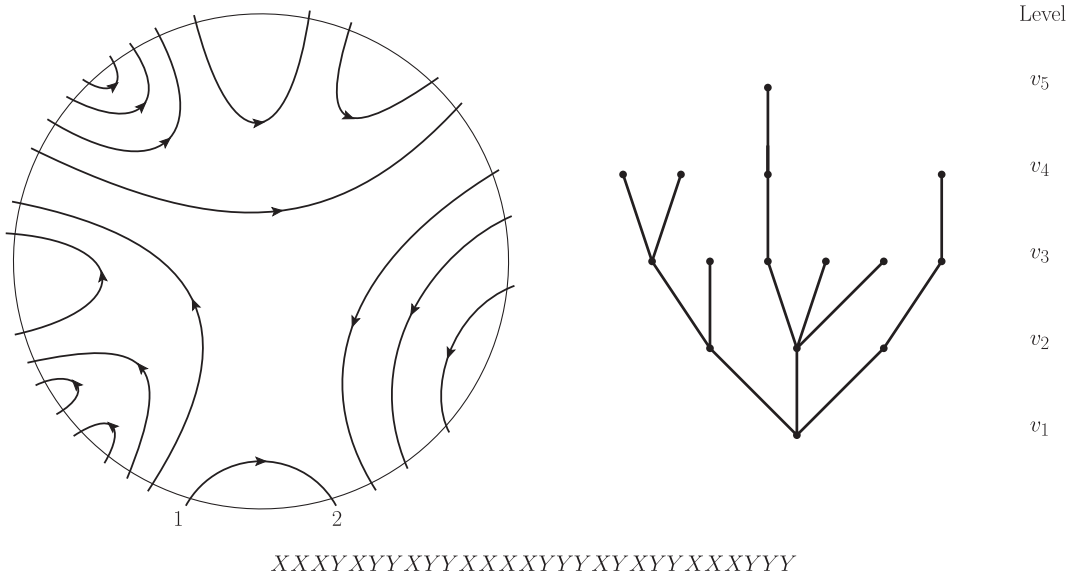


FIG. 2. The quark line graph for the above Dyck word is shown on the left. The quark line directions have been chosen in accordance with an all-positive signature. On the right is the associated rooted oriented Dyck tree, which can be seen as a dual graph to the quark line graph once the line $(1 \rightarrow 2)$ is removed and the circle representing the edge of the plane is identified as a node. The levels v_i of the tree are given on the far right.

B. A rooted oriented tree and the general signature

The second generalization concerns the signature of the Dyck permutation—which way the quark lines are directed in the quark line diagram for each primitive in the basis. With the prescription described here, there are 2^{k-1} different bases; in [1], only the basis in which each permutation has an all-positive signature $(+, +, \dots, +)$ was proven. The more general result comes from an iterative proof based around rooted oriented trees. These trees are also used in the organization of the flavor recursion in the following section. A rooted oriented tree is a dual graph to the quark line graphs with the edge $(1 \rightarrow 2)$ removed, with the circle identified as a node to which the edges of the quark line graph (the quark lines) attach. Alternatively, a rooted oriented tree can be drawn directly from the Dyck word via a “snail climbing up a tree” path—every time there is an X , the snail crawls up a branch, and every time there is a Y , it crawls down the other side; see Fig. 2. The number of rooted oriented trees for a given number of nodes is given by the Catalan number, the same as the number of Dyck words. The rooted oriented tree is composed of nodes at different levels, labeled by v_i , where $1 \leq i \leq k$, and with $i = k$ only being achieved for the Dyck word of the form $XX\dots XYY\dots Y$.

An iterative procedure can be set up around the concept of the “maturity” of a tree, which is determined by the number of nodes n_i the tree has at each level v_i , with a more mature tree having more nodes at higher levels. Concretely, we can define a tree A to be more mature than a tree B if, comparing the number of nodes at each level v_i , starting at $i = 1$, some i is reached where $n_i^A < n_i^B$. It is possible that distinct trees can have the same maturity.

We can show that the signature of a basis primitive can be chosen at will. For clarity of presentation, we will make a couple of simplifications. First, we will present the proof for the pure-quark case, and point out the straightforward generalizations to include gluons where necessary. Second, we will ignore the effect of Fermi statistics, which generates negative signs for odd permutations of the labels from the canonical form $1\dots n$ —these signs can be put in by hand after all relations are taken into account. The main ingredient is being able to show the following:

$$\mathcal{A}(1, \dots, j, \beta, i, \dots, 2) = -\mathcal{A}(1, \dots, i, \beta^T, j, \dots, 2) + (\text{trees of higher maturity}), \quad (9)$$

where β is some set of quark and antiquark labels, and where β^T denotes the set β with the ordering reversed. The “+(trees of higher maturity)” refers to primitives which have a more mature rooted oriented tree than the primitive on the left-hand side. The first primitive on the right-hand side has equal maturity to the primitive on the left-hand side—the orientation of the quark pair (i, j) has been reversed, as has the orientation of all the quark pairs contained within β , but

no other quark lines have had their orientation changed. Eq. (9) follows from Eq. (A2) given in the appendix [with gluons, this equation generalizes to Eq. (A4)], and is a $SU(3)_c$ group theory relation—it follows from KK relations modified by the quark line structure.

If the quark lines of a primitive are oriented in a different way to the signature chosen for our basis vector for this particular Dyck topology and flavor assignment, we can apply Eq. (9) to each of the wrongly oriented flavor pairs, starting with the pair sitting at the lowest level on the rooted oriented tree (if there is more than one pair at this level, they can each be oriented in turn, before orienting pairs further up the tree). We can then work upwards through the tree, each time rewriting the primitive up to higher maturity terms, which we assume we can solve, until all of the quark lines are oriented in the desired way, dictated by the signature of this particular flavor assignment that we chose for our basis. This iterative procedure terminates at amplitudes which have a tree of the highest possible maturity—these are based on the Dyck word of the form $XX\dots XYY\dots Y$ (see Fig. 3). These primitives can be ordered simply, since Eq. (A2) in this case becomes [see Eq. (A4) for the gluonic case]

$$\mathcal{A}(1, \dots, j, \beta, i, \dots, 2) = -\mathcal{A}(1, \dots, i, \beta^T, j, \dots, 2). \quad (10)$$

This completes the proof that the signature of each flavor pair assignment can be chosen independently. As a simple example of how the iteration works, consider the pure six quark case, and choose as a basis from Eqs. (4)–(7) where the first two have an all-positive signature, and the second two have an all-negative signature:

$$\begin{aligned} \mathcal{A}(1, 3, 4, 5, 6, 2), & \quad \mathcal{A}(1, 5, 6, 3, 4, 2), \\ \mathcal{A}(1, 6, 4, 3, 5, 2), & \quad \mathcal{A}(1, 4, 6, 5, 3, 2). \end{aligned} \quad (11)$$

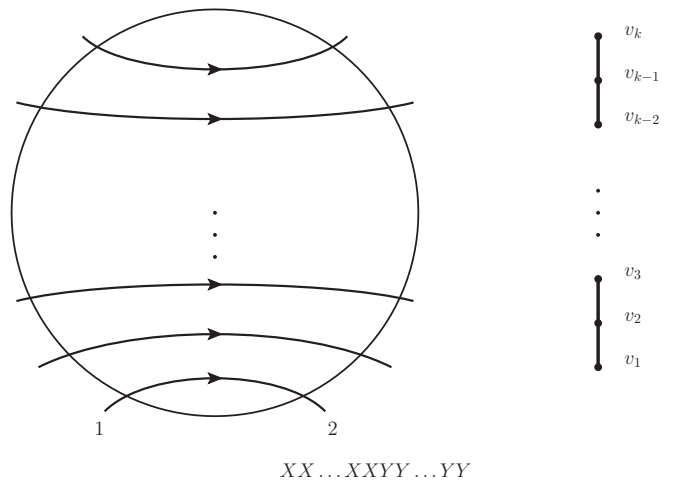


FIG. 3. The Dyck word and quark line graph of a primitive with k quark lines, for which the rooted oriented Dyck tree has maximum height, or maturity.

We demonstrate how the primitive amplitude $\mathcal{A}(1, 4, 3, 6, 5, 2)$ can be expressed in terms of this basis. First, applying Eq. (A2),

$$\mathcal{A}(1, 4, 3, 6, 5, 2) = -\mathcal{A}(1, 3, 4, 6, 5, 2) - \mathcal{A}(1, 3, 6, 5, 4, 2). \quad (12)$$

Neither of the resulting terms are yet in our chosen basis. The iteration continues on the first of the primitives on the right-hand side of Eq. (12) as

$$\begin{aligned} \mathcal{A}(1, 3, 4, 6, 5, 2) &= -\mathcal{A}(1, 3, 4, 5, 6, 2) - \mathcal{A}(1, 5, 3, 4, 6, 2) \\ &= -\mathcal{A}(1, 3, 4, 5, 6, 2) - [-\mathcal{A}(1, 6, 4, 3, 5, 2)], \end{aligned} \quad (13)$$

where in the second equality the recursion acts on the second primitive obtained after the first equality. The iteration continues on the second of the primitives on the right-hand side of Eq. (12) as

$$\begin{aligned} \mathcal{A}(1, 3, 6, 5, 4, 2) &= -\mathcal{A}(1, 4, 5, 6, 3, 2) \\ &= -[-\mathcal{A}(1, 4, 6, 5, 3, 2)]. \end{aligned} \quad (14)$$

At this point the recursion has terminated, and we have expressed the primitive amplitude $\mathcal{A}(1, 4, 3, 5, 6, 2)$ in terms of our basis:

$$\begin{aligned} \mathcal{A}(1, 4, 3, 6, 5, 2) &= \mathcal{A}(1, 3, 4, 5, 6, 2) - \mathcal{A}(1, 6, 4, 3, 5, 2) \\ &\quad - \mathcal{A}(1, 4, 6, 5, 3, 2). \end{aligned} \quad (15)$$

The primitive amplitudes in the bases described in this section are independent, since we have shown how to express any primitive in terms of a set of size $(n-2)!/k!$ —this is the minimum size for a basis, since the one-flavor case with $(n-2)!$ independent primitives must be recovered using a sum over the $k!$ momentum permutations of the distinct-flavor case. That is, since

$$\mathcal{M}_{\text{one-flavor}} = \sum_{\mathcal{P}(p_2, p_4, \dots, p_{2k})} (-1)^{\text{sgn}\mathcal{P}} \mathcal{M}_{\text{distinct-flavor}}(\{p_i\}), \quad (16)$$

the number of independent primitives is $\geq (n-2)!/k!$. The momentum permutations acting on the basis primitives of $\mathcal{M}_{\text{distinct-flavor}}$ bring them outside of the basis for a single distinct-flavor amplitude. For instance, the momentum swap $p_2 \leftrightarrow p_4$ would mean that the antiquark with momentum p_1 (p_3) is no longer connected to the quark with momentum p_2 (p_4). This is different to swapping the cyclic positions of the quark with momentum p_2 and the quark with momentum p_4 , but keeping the same flavor pairing so that p_1 still connects to p_2 and p_3 still connects to p_4 . The latter is what happens under the KK relations, and was the

subject of this section; different quark pairings, on the other hand, are tied up with what happens when quark lines have an identical flavor, and it is to these considerations we now turn.

III. FLAVOR RECURSION AND ALL MASSLESS QCD TREES FROM $\mathcal{N} = 4$ SYM

The idea behind the flavor recursion is simple—write a k -flavor primitive of the form $\mathcal{A}^{n_f=k}(1, \sigma, 2)$ as a one-flavor primitive with the same labeling $\mathcal{A}^{n_f=1}(1, \sigma, 2)$, and then subtract any wrong quark line contributions using k -flavor primitives with a different flavor pairing:

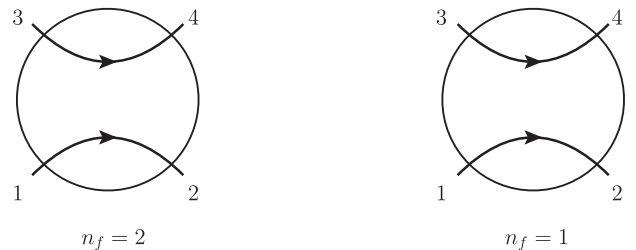
$$\mathcal{A}^{n_f=k}(1, \sigma, 2) = \mathcal{A}^{n_f=1}(1, \sigma, 2) - \sum \mathcal{A}_{\text{wrong flav}}^{n_f=k}(1, \sigma, 2). \quad (17)$$

We can then iterate this procedure on each of the subtraction terms, but, if this is to work, then these subtraction amplitudes must be further down an iterative direction which must eventually terminate. It follows that there must exist some k -flavor primitive amplitudes that are simply equal to one-flavor primitives,

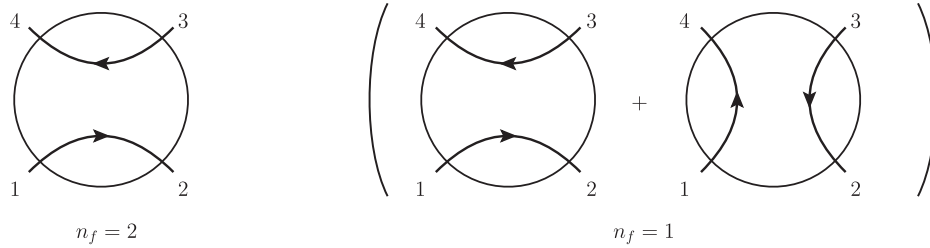
$$\mathcal{A}^{n_f=k}(1, \sigma, 2) = \mathcal{A}^{n_f=1}(1, \sigma, 2), \quad (18)$$

and this happens for the permutations based around the Dyck tree of the highest maturity (Fig. 3) and which have an all-positive signature. We shall also see that the direction in which to iterate is up the rooted oriented tree introduced in the previous section. The reason that such a relation as Eq. (18) exists is down to the role of planarity and cyclic ordering in restricting the pole structure of the one-flavor amplitude to be the same as that of the k flavor. Throughout this section, we use one flavor to implicitly specify that all quark lines have the same helicity, which is the most general case (other helicity cases, and obtaining two-, three- ... $(k-1)$ -flavor amplitudes from one-flavor amplitudes follows trivially from this).

As a simple example of this planar-cyclic restriction, consider the four-quark primitive amplitudes $\mathcal{A}^{n_f=k}(1, 3, 4, 2)$ and $\mathcal{A}^{n_f=1}(1, 3, 4, 2)$. Although it is possible, by flavor considerations, for a quark line to run from $(1 \rightarrow 4)$ in the latter primitive, this is forbidden by planarity and the specified cyclic ordering:



The situation is different if we compare the k -flavor amplitude $\mathcal{A}^{n_f=k}(1, 4, 3, 2)$ with the one-flavor amplitude with the same cyclic labeling, $\mathcal{A}^{n_f=1}(1, 4, 3, 2)$:



Now there is a difference between the amplitudes—in the one-flavor case there is both an s - and a t -channel pole, since nothing restricts the quark line to run from $(1 \rightarrow 4)$. Again, for the sake of clarity, a simplification has been made to ignore the minus signs arising from Fermi statistics when the flavor pairing in the one-flavor amplitude permutes from $(1,2)(3,4)$ to $(1,4)(2,3)$ —that is, a relative $+$ sign has been used between the quark line graphs of the one-flavor primitive above, rather than a $-$ sign. Doing this will make the structure of the flavor recursion in the next section more transparent—the minus signs can easily be reinstated in whatever method is used to calculate the one-flavor amplitudes.

It is easy to see that planarity requires the one-flavor primitive amplitudes with a quark line structure based on $XX...XYY...Y$ which has an all-positive signature to be identical to k -flavor primitives which have the same highest maturity tree:

$$\mathcal{A}^{n_f=k}(1\bar{q}\bar{q}...q_1q_2...q_2) = \mathcal{A}^{n_f=1}(1\bar{q}\bar{q}...q_1q_2...q_2). \quad (19)$$

There is no planar way in which to connect an antiquark to a quark in the one-flavor case that is not the one it would

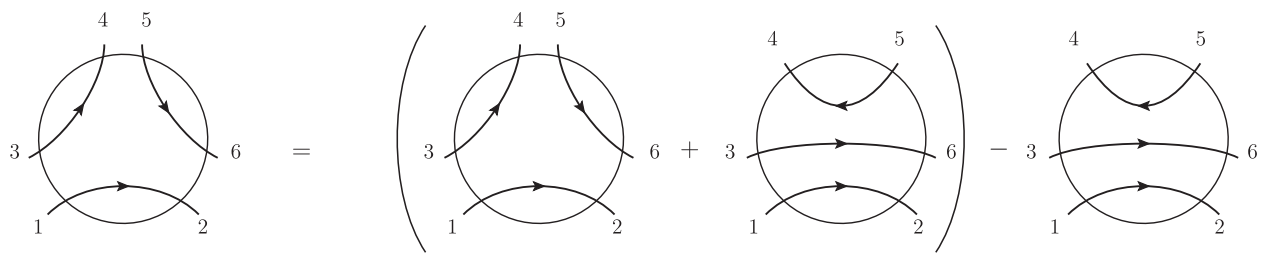
have connected to in the k -flavor case without giving rise to crossed quark lines.

The all-positive signature basis will play a special role in the following. We now present the flavor recursion, which starts by acting on k -flavor primitives of an all-positive signature:

- (1) Express an all-positive signature k -flavor primitive as a one-flavor primitive which has the same cyclic ordering of external particles, minus subtraction k -flavor primitives which serve to remove the wrong quark line contributions [see Eq. (17)].
- (2) Re-express each of the subtraction primitives in terms of all-positive signature k -flavor primitives as discussed in the previous section.
- (3) Repeat from step 1 on each of these k -flavor primitives.

Step 2 is important in order to allow the recursion to terminate.

As a warm-up to the general case for the recursion, consider again the purely six-quark primitive $\mathcal{A}_{(3,4)(5,6)}^{n_f=k}(1, 3, 4, 5, 6, 2)$, which is an all-positive signature primitive. We have used the subscripts $(3,4)(5,6)$ to indicate the flavor pairing used (in the previous section this pairing would have been implicit). This primitive can be expressed as a one-flavor primitive minus a subtraction as in the following way:



$$\mathcal{A}_{(3,4)(5,6)}^{n_f=k}(1, 3, 4, 5, 6, 2) = \mathcal{A}^{n_f=1}(1, 3, 4, 5, 6, 2) - \mathcal{A}_{(3,6)(5,4)}^{n_f=k}(1, 3, 4, 5, 6, 2). \quad (20)$$

Again, we are ignoring the effect of Fermi statistics. This completes step 1. We now need to re-express the subtraction primitive $\mathcal{A}_{(3,6)(5,4)}^{n_f=k}(1, 3, 4, 5, 6, 2)$ (which has a different flavor pairing as indicated by the subscripts) in terms of primitives of an all-positive signature, as required by step 2 above. This is achieved through the group theory relation Eq. (A2) [or with gluons Eq. (A4)],

$$\mathcal{A}_{(3,6)(5,4)}^{n_f=k}(1, 3, 4, 5, 6, 2) = -\mathcal{A}_{(3,6)(5,4)}^{n_f=k}(1, 3, 5, 4, 6, 2). \quad (21)$$

The recursion now iterates (step 3) by acting with step 1 on this primitive:

$$\mathcal{A}_{(3,6)(5,4)}^{n_f=k}(1, 3, 5, 4, 6, 2) = \mathcal{A}^{n_f=1}(1, 3, 5, 4, 6, 2). \quad (22)$$

There are no further subtractions—this is an amplitude with a Dyck tree of highest maturity—so the recursion terminates here and the full result is obtained, via Eqs. (22) and (21) into Eq. (20):

$$\mathcal{A}_{(3,4)(5,6)}^{n_f=k}(1, 3, 4, 5, 6, 2) = \mathcal{A}^{n_f=1}(1, 3, 4, 5, 6, 2) + \mathcal{A}^{n_f=1}(1, 3, 5, 4, 6, 2). \quad (23)$$

Using this method, we have succeeded in expressing a three-flavor primitive in terms of one-flavor primitives.

As discussed at the end of the previous section, the intermediate primitive $\mathcal{A}_{(3,6)(5,4)}^{n_f=k}(1, 3, 5, 4, 6, 2)$ is outside the basis of k -flavor primitives with flavor pairs (3,4)(5,6), since it specifies that pairs of equal flavor are (3 → 6) and (4 → 5).

A. The flavor recursion for the general case

The direction in which the general flavor recursion will iterate is based around the rooted oriented tree introduced in Sec. II B. We introduce a more streamlined notation to indicate the flavor structure of each primitive—as in the above example, we will be dealing with k -flavor primitives with different flavor pairings to the usual convention. We can label a primitive with a subscript $f = \{f_1, f_2, \dots, f_k\}$ being a permutation of $\{2, 4, \dots, 2k\}$ to denote the flavor pairing $(1 \rightarrow f_1), (3 \rightarrow f_2), \dots, (2k - 1 \rightarrow f_k)$. Under this notation, we can rewrite Eq. (17) more precisely as

$$\mathcal{A}_f^{n_f=k}(1, \sigma, 2) = \mathcal{A}^{n_f=1}(1, \sigma, 2) - \sum_{f' \in S_k} (1 - \delta_{ff'}) \mathcal{A}_{f'}^{n_f=k}(1, \sigma, 2), \quad (24)$$

where the $\delta_{ff'}$ removes the amplitude on the left-hand side from the sum on the right-hand side, and where we have again ignored minus signs coming from Fermi statistics. In considering the full sum over S_k we are taking into account all possible flavor pairings, but they will not all be nonzero, since some of these pairings will give rise to crossed quark lines.

Now consider a general all-positive signature k -flavor primitive, for example, the one shown in Fig. 2. Which primitives in the sum on the right-hand side of Eq. (24) are zero? When we are dealing with an all-positive signature permutation on the left-hand side, the reader can convince

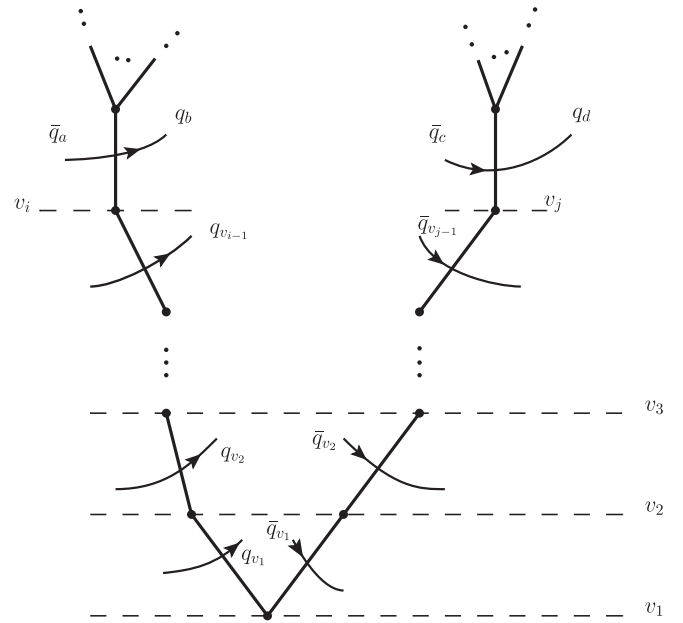


FIG. 4. Path back to the lowest node below two different nodes of the tree, one at level v_i , associated with quark line $\bar{q}_1 \rightarrow q_2$, and the other at level v_j , associated with the quark line $\bar{q}_3 \rightarrow q_4$.

herself or himself that if the pairing $(1 \rightarrow 2)$ is not present, then the amplitude must be zero: if the flavor pairing is $(1 \rightarrow f_i)$, then there will be either an odd number of quarks or an odd number of antiquarks between the position f_i and 1 in the cyclic order, $\mathcal{A}^{n_f=k}(1, \dots, f_i, \dots, 2)$.

A more careful consideration reveals the following result: for the k -flavor subtraction primitives to be nonzero in Eq. (24), the flavor permutations can only swap quark flavors at the same level of the Dyck tree. Consider any two nodes on the Dyck tree corresponding to the all-positive signature primitives on the left-hand side of Eq. (24), one node at level v_i and the other at level v_j . Next, consider the path to each node from the root node at level v_1 (these paths are unique, but they could overlap up to some node $v_{\text{overlap}}^{\text{max}}$ from which point onwards they differ—in this case in everything that follows, the node at v_1 should be taken to represent $v_{\text{overlap}}^{\text{max}}$); see Fig. 4. Let $\bar{q}_a \rightarrow q_b$ denote the quark line which is the dual line passing through the edge above the node v_i and $\bar{q}_c \rightarrow q_d$ be the quark line which is the dual line passing through the edge above the node v_j . The primitive amplitude has the form

$$\mathcal{A}_{..(a,b)..(c,d)..}(\dots \bar{q}_a \dots q_b \dots \bar{q}_c \dots q_d \dots). \quad (25)$$

The second set of ellipses in this primitive contains quark and antiquark labels which are dual quark lines to edges in the tree above the node v_i . They stand for an equal number of quarks and antiquarks. Similarly, the fourth set of ellipses in the primitive contains an equal number of quarks and antiquarks with quark lines dual to edges above the node v_j . In between this, the third set of ellipses could be

any number of quarks and antiquarks coming from two possible origins. First, they can come from lines dual to other edges of the Dyck tree which branch off to the right of the path to v_i or lines dual to edges which branch off to the left of the path to v_j —this origin gives rise to an equal number of quarks and antiquarks. Second, they can be quarks from the quark lines which cross the path to v_i (labeled $q_{v_1}, \dots, q_{v_{i-1}}$ in Fig. 4) or they can be antiquarks from the quark lines which cross the path to v_j (labeled $\bar{q}_{v_1}, \dots, \bar{q}_{v_{j-1}}$). The fact that it is always quarks crossing the path to v_i whereas it is antiquarks crossing the path to v_j is a consequence of having a primitive with an all-positive signature—in this way it is an important feature of the recursion. There are $i - 1$ quarks and $j - 1$ antiquarks from this second origin.

Now consider what happens when the quark lines $\bar{q}_a \rightarrow q_b$ and $\bar{q}_c \rightarrow q_d$ mix, so that we have $\bar{q}_a \rightarrow q_d$ and $\bar{q}_c \rightarrow q_b$. The number of quarks and antiquarks in between \bar{q}_a and q_d needs to be equal for the primitive $\mathcal{A}_{..(a,d)..(c,b)..}(\dots\bar{q}_a\dots q_b\dots\bar{q}_c\dots q_d\dots)$ to be nonzero, so it is a necessary condition that $v_i = v_j$.

We can use this information to split up the sum over all mixings in Eq. (24) as follows. Separate the labeling of the flavor structure into lists of those at equal levels of the tree: $f_1 \dots f_h$, where h is the highest level of the tree, so that the primitive can be labeled $\mathcal{A}_{f_1 \dots f_h}^{n_f=k}$. We can then write the permutation sum by splitting it up so that only permutations within each level are considered:

$$\begin{aligned} \mathcal{A}_{f_1 f_2 \dots f_h}^{n_f=k}(1, \sigma, 2) &= \mathcal{A}^{n_f=1}(1, \sigma, 2) - \sum_{f'_1 \in S_{n_1}} \sum_{f'_2 \in S_{n_2}} \dots \\ &\times \sum_{f'_h \in S_{n_h}} \left(1 - \prod_{i=1}^h \delta_{f_i f'_i} \right) \mathcal{A}_{f'_1 f'_2 \dots f'_h}^{n_f=k}(1, \sigma, 2). \end{aligned} \quad (26)$$

Again, the product of delta functions serves to remove the amplitude on the left-hand side from the sum. It is just the quark line structure which is important for the recursion; no reference to the gluon position is needed. In the sum on the right-hand side some of the amplitudes are still zero, but splitting the sum up explicitly as above makes it clear that a quark at a higher level of the tree cannot pair with an antiquark at a lower level of the tree, and as such all of the k -flavor amplitudes appearing on the right-hand side of Eq. (26) have a Dyck tree of higher maturity than the primitive on the left-hand side. Again, this feature is a consequence of having an all-positive signature primitive on the left-hand side.

Equation (26) is the master equation for the recursion. After it is applied, the recursion continues by next expressing each of the subtracted k -flavor amplitudes on the right-hand side of Eq. (26) in their all-positive signature Dyck basis. As discussed in the previous section, in doing this the only primitives of higher maturity are created. This

completes one iteration of the recursion. After it, all remaining k -flavor amplitudes are of higher maturity than the one of the left-hand side of Eq. (26). The recursion eventually terminates at the amplitude of maximum height, Fig. 3, through Eq. (19).

If a different helicity configuration is chosen from the one where all quarks have equal helicity, the recursion proceeds in exactly the same way, so that the same expressions are reached, except now some of the “one-flavor” amplitudes can be zero (owing to the fact that opposite helicity quark lines cannot mix). Finally, since two-, three-, $\dots(k - 1)$ -flavor amplitudes can be expressed as permutation sums over k -flavor amplitudes, it follows that they can also all be expressed in terms of one-flavor amplitudes, via the above recursion.

B. Obtaining QCD from $\mathcal{N} = 4$ SYM

We do not present the details of Drummond and Henn’s formula for the solution to $\mathcal{N} = 4$ SYM at tree level, and instead refer the reader to the original publication [10]. The solution is given in terms of a super wave function Φ , which in terms of on-shell gluon (g^+, g^-), gluino (\tilde{g}_A, \tilde{g}^A), and scalar (ϕ_{AB}) states, and Grassmann variables η^A , is

$$\begin{aligned} \Phi(\lambda, \bar{\lambda}, \eta) &= g^+(\lambda, \bar{\lambda}) + \eta^A \tilde{g}_A(\lambda, \bar{\lambda}) + \frac{1}{2} \eta^A \eta^B \phi_{AB}(\lambda, \bar{\lambda}) \\ &+ \eta^A \eta^B \eta^C \epsilon_{ABCD} \tilde{g}^D(\lambda, \bar{\lambda}) \\ &+ \eta^A \eta^B \eta^C \eta^D \epsilon_{ABCD} g^-(\lambda, \bar{\lambda}), \end{aligned} \quad (27)$$

with $A = 1, 2, 3, 4$. In [10] it was discussed how to perform the Grassmann integrations so as to project the formula onto specific external states, and, in [13], all- n formulas were derived for the projection onto n external particles—gluons and gluinos of the four possible flavors, $A = 1, 2, 3, 4$. These amplitudes then begin to resemble QCD amplitudes, with gluinos identified as quarks—because they are color stripped primitives, gluinos have the same interactions with gluons as quarks do. The problem in relating these amplitudes to the QCD amplitudes lies in avoiding non-QCD interactions involving scalars, which couple two gluinos of different flavors—see Fig. 5. This was achieved for specific cases in [13] by careful choices of external flavor and by summing over different flavor permutations so that scalar contributions were eliminated—either not being present or canceling against each other—so that all QCD amplitudes with up to four distinct-flavor quark lines were shown to be obtainable from Drummond and Henn’s solution. It was left as an open problem as to whether amplitudes with more quarks could also be obtained—clearly a necessary issue to deal with is the one of flavor, since there are only four flavors of gluino and if the $\mathcal{N} = 4$ SYM solution were able to describe five-flavor amplitudes then these would have to be obtainable from at least four-flavor amplitudes.

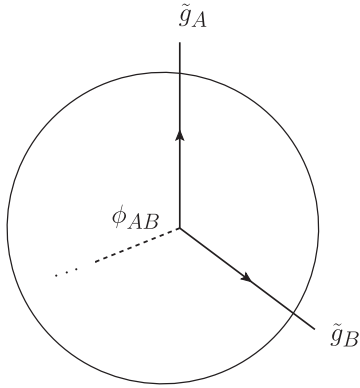


FIG. 5. The termination of a scalar ϕ_{AB} results in two external gluinos \tilde{g}_A and \tilde{g}_B (to be identified as quarks) of different flavors, since the coupling requires $A \neq B$.

It follows directly from the flavor recursion described in the previous section that in fact the whole of massless QCD at tree level *is* obtainable from $\mathcal{N} = 4$ SYM, since one-flavor amplitudes are identical in QCD and $\mathcal{N} = 4$ SYM. This is because if a scalar exchange is created, we can trace it to its termination point at some quark line (it has to terminate, because no external scalars are specified in the amplitude); see Fig. 5. However, for this to be nonzero, it necessarily creates two different flavored quarks which will eventually leave the amplitude, contradicting the original specification of a one-flavor amplitude. In other words, when all quark lines are of the same flavor, no scalars can be exchanged between the quark lines, since they couple to different flavored quarks as $\phi_{AB}\tilde{g}_A\tilde{g}_B, A \neq B$. (This is similar reasoning to that which asserts that all-gluon tree amplitudes are identical in QCD and $\mathcal{N} = 4$ SYM).

So, in order to obtain all of massless QCD at tree level, the most difficult primitives to obtain, the all-distinct-flavor cases, are expressed through the flavor recursion in terms of a particular set of one-flavor QCD primitives. These can then be directly obtained from the one-flavor $\mathcal{N} = 4$ SYM amplitudes, once projected onto gluon and one-flavor gluino external states, as described in [13]. One feature of this particular set of one-flavor primitives is that the external quark and antiquark labels are fixed into a Dyck word, with X s identified with antiquark labels and Y s identified with quark labels. The discussion above has implicitly assumed that the helicities of the quark lines in the k -flavor primitives are equal. However, this is all that is needed to obtain massless QCD from the $\mathcal{N} = 4$ SYM formulas, since all other helicity cases can be obtained from them via the charge parity equation,

$$\mathcal{A}(\dots i_q^\pm \dots j_q^\mp \dots) = -\mathcal{A}(\dots i_q^\mp \dots j_q^\pm \dots). \quad (28)$$

It is also possible to set up a second recursion from which to obtain QCD trees from $\mathcal{N} = 4$ SYM trees. This is a two-flavor recursion, and is described briefly in Appendix B.

IV. DISCUSSION

Before concluding, we address issues surrounding amplitudes beyond massless QCD, as well as the role of BCJ relations between the one-flavor primitives, and the counting of the number of one-flavor primitives needed to produce the k -flavor primitives.

A. Beyond massless QCD

In this subsection we describe the behavior of amplitudes involving electroweak bosons and massive quarks specifically under the KK relations and the flavor recursion.

Amplitudes involving W s, Z s, and photons all require external quark lines in tree-level amplitudes to couple to. The electroweak particles do not carry color charge, and these amplitudes satisfy the KK relations described in previous sections as if there were no electroweak particle present. When there is more than one weak boson present, the amplitudes can be further ordered under their $SU(2)$ charge—see Ref. [43]. However, because quarks of different flavors carry different electroweak charges, the flavor recursion cannot apply once electroweak bosons are present—they are sensitive to flavor in contrast to gluons which are not.

Amplitudes involving massive quarks, top quarks for example, also satisfy the same KK relations as those with purely massless quarks—the presence of a massive quark line does not affect the group theory factors from which the relations follow. However, here too the flavor recursion relation described above will not work, since a one-flavor amplitude with external quark lines of differing mass is not well defined.

B. BCJ relations and QCD primitives

Another interesting thing about being able to express all of massless QCD in terms of one-flavor amplitudes is that these amplitudes can be written in BCJ form, with kinematic numerator factors satisfying Jacobi relations such that the number of independent primitive amplitudes is reduced to $(n-3)!$ (relations hold between primitives multiplied by kinematic factors).

This motivates a question as to how BCJ relations apply in the distinct-flavor case. This would have to be reconciled with the counting of the number of one-flavor primitives needed to reconstruct a basis for the distinct-flavor case. We perform this counting in Table I, where this number is compared to the number of k -flavor independent primitives under KK relations alone, and the $(n-3)!$ one-flavor primitives under BCJ. The second column was constructed using examples up to the pure quark case $n = 10$ —i.e. using the flavor recursion to find the one-flavor primitives needed for every k -flavor amplitude and then taking the union of this set—and the general formula is conjectured to hold beyond this.

TABLE I. Counting of the number of independent primitive amplitudes under KK relations for the distinct-flavor case, the number of one-flavor amplitudes needed to construct a basis of these amplitudes, and the number of independent primitive amplitudes under BCJ relations for the one-flavor case.

n, k	# k -flavor (no BCJ)	# one-flavor for k -flavor (no BCJ)	BCJ one-flavor
4,2	1	1	1
6,3	4	6	6
8,4	30	90	120
10,5	336	2520	5040
...
n, k	$(n-2)!/k!$	$(n-2)!/2^{k-1}$	$(n-3)!$

Are there BCJ relations between the primitives in this particular set of all the one-flavor primitives? If so, how many independent amplitudes remain? These questions will be interesting to address in future work.

C. Further discussion

We have shown that any tree-level massless QCD amplitude can be expressed in terms of amplitudes that possess effective $\mathcal{N} = 1$ supersymmetry. These amplitudes might be expected to have nicer properties than the distinct-flavor case, owing to this fact. Of course, one such realization of this is that a closed form solution for them is known, this in turn owing to the fact that $\mathcal{N} = 1$ SYM is a closed subset of $\mathcal{N} = 4$ SYM, for which this tree-level solution is known.

The flavor recursion singles out a particular set of one-flavor amplitudes needed to reconstruct the distinct-flavor amplitudes. These are the ones with the external quarks and antiquarks ordered as Dyck words, with antiquarks at X locations and quarks at Y locations. One might wonder whether these amplitudes could take a simpler form, given that the underlying Feynman diagram representation for k -flavor amplitudes is considerably simpler than the one-flavor case. Interestingly, in Drummond and Henn's proof of their solution to $\mathcal{N} = 4$ SYM trees, the concept of a rooted (but not oriented) tree was also used to define a direction in which to perform Britto-Cachazo-Feng-Witten (BCFW) recursion [44], and there is some freedom in this choice of recursive direction. It would be interesting if some alignment of the BCFW recursion with the Dyck tree could result in expressions where the quark line structure was easily identified in the analytic form. Another avenue would be to investigate the use of the momentum twistor variables [45,46] (λ, μ, χ) rather than the $(\lambda, \tilde{\lambda}, \eta)$ of Ref. [10], under which expressions for amplitudes involving fermions can take simpler form.

A further feature of the Kleiss-Kuijff relations worth mentioning is that they do not depend on the dimensionality of spacetime. Studies of amplitudes in different spacetime dimensions is of theoretical interest, in particular in $d = 3$ and $d = 6$ (see e.g. [47] for a recent review), and dimensional

regularization is a common way of isolating infrared and ultraviolet singularities in amplitudes. For phenomenological applications, the technique of D -dimensional generalized unitarity [25] is an approach to build one-loop amplitudes out of tree-level primitive amplitudes in dimensions $d > 4$, and is particularly suited to numerical evaluation.

Since then, the flavor recursion also does not depend on the dimensions of spacetime, and given that unitarity based techniques are able to construct any loop-order amplitude purely out of trees (evaluated, in general, in higher dimensions), it follows that any loop-order amplitude in massless QCD is obtainable from tree-level amplitudes calculated in one-flavor QCD. This means that the only objects that actually require a field theory calculation (i.e. the tree-level amplitudes) are done in a theory where there is no notion of flavor. Rather, flavor comes about only in the way we combine these objects, as we have seen for the tree-level case studied in this paper.

V. CONCLUSION

In conclusion, we have studied the $SU(3)_c$ group theory relations between tree-level QCD primitive amplitudes involving k quark lines of distinct flavor and $n - 2k$ gluons, and shown that they reduce the number of independent primitives to $(n-2)!/k!$. We described how bases can be constructed using the concept of a rooted, oriented Dyck tree. Exploiting the planarity and cyclic ordering of a class of these primitives which have an all-positive signature, and using the $SU(3)_c$ relations, we derived a flavor recursion relation, also based around a Dyck tree. This flavor recursion can express a multicolor tree-level QCD primitive in terms of one-flavor tree-level QCD primitives, which possess effective $\mathcal{N} = 1$ supersymmetry. In turn, this makes it possible to use known formulas from $\mathcal{N} = 4$ SYM to obtain all of massless QCD at tree level. An interesting aspect is that, using these results, no notion of flavor is needed when making amplitude calculations in perturbative, massless QCD, given that unitarity based techniques can construct higher loop amplitudes out of tree-level amplitudes.

We expect that a knowledge of a minimal general QCD basis, along with implementation of one-flavor $\mathcal{N} = 4$ SYM results, will increase the performance of computational efforts to describe multijet events at fixed order in perturbation theory (see [48] for a discussion of the phenomenological use of the $\mathcal{N} = 4$ SYM analytic formulas, although the use of purely one-flavor formulas should be investigated further).

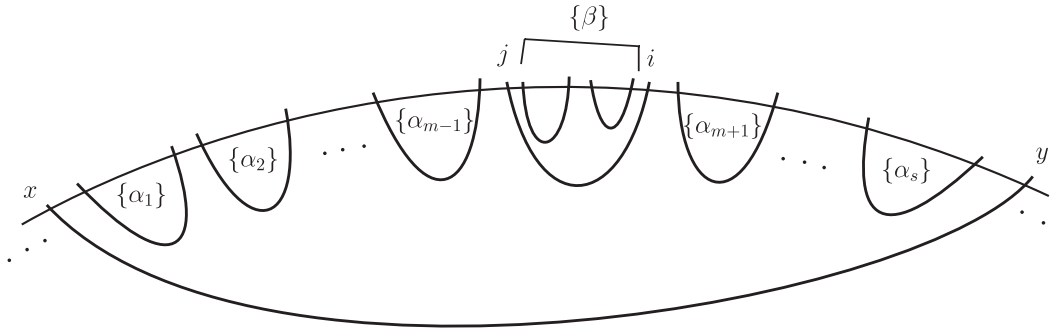
In the final stages of the preparation of this script, the preprint [49] appeared, which provides an independent, multicolor approach to obtaining QCD trees from $\mathcal{N} = 4$ SYM, similar to the two-flavor one outlined in Appendix B of this work. It would be interesting to fully explore this connection, and the one between the fermion flip identities defined in [49] and the $SU(3)_c$ relations presented in Ref. [1] and used here.

ACKNOWLEDGMENTS

This research was supported by ERC Grant No. 291377 “LHCtheory—Theoretical predictions and analyses of LHC physics: advancing the precision frontier.” T.M. thanks Mat Bullimore, Lance Dixon, James Drummond, Johannes Henn, and Henrik Johansson for interesting and helpful conversations and suggestions during the course of this work.

APPENDIX A: GROUP THEORY IDENTITIES BETWEEN QCD PRIMITIVES

A group theory identity for the pure multiquark case, $2k = n$, which can be used to further relate primitive amplitudes of the form $\mathcal{A}(1, \dots, 2)$ was given in [1], where a proof is also given—see this reference for a detailed discussion. The identity is given with reference to a primitive of the form



$$\mathcal{A}(\dots x\{\alpha_1\}\{\alpha_2\}\dots\{\alpha_{m-1}\}j\{\beta\}i\{\alpha_{m+1}\}\dots\{\alpha_s\}y\dots). \quad (\text{A1})$$

Here, the boundaries of the sets $\{\alpha_i\}$ are quark lines corresponding to a level one higher than the level of the quark line $(x - y)$ in the Dyck tree for the above primitive. Each set $\{\alpha_i\}$ may contain further quark lines (i.e. they may be sets of length greater than two). The quark line $(j - i)$ is also at a level one higher in the Dyck tree than the quark line $(x - y)$. There can be further quark lines between the positions of the labels j and i in the primitive—these are denoted as $\{\beta\}$ (they are all at a higher level in the Dyck tree than $(j - i)$). A group theory identity satisfied by this primitive is

$$\begin{aligned} & \mathcal{A}(\dots x\{\alpha_1\}\dots\{\alpha_{m-1}\}j\{\beta\}i\{\alpha_{m+1}\}\dots\{\alpha_s\}y\dots) \\ &= - \sum_{c=1}^m \left[\sum_{\text{OP}\{D_c\}\{E\}} \left(\sum_{\text{OP}\{A_c\}\{B\}} \mathcal{A} \left(\dots x\{\alpha_1\}\dots\{\alpha_{c-1}\}i \underbrace{\{\alpha_c\}\dots\{\alpha_{m-1}\}}_{\{A_c\}} \underbrace{\{\beta^T\}}_{\{B\}} j \underbrace{\{\alpha_{m+1}\}\dots\{\alpha_s\}}_{\{E\}} y\dots \right) \right) \right], \quad (\text{A2}) \end{aligned}$$

where $\text{OP}\{A\}\{B\}$ stands for “ordered permutations” and is the shuffle product of the sets $\{A\}$ and $\{B\}$, and where $\{\beta^T\}$ is the set $\{\beta\}$ with the ordering reversed. In all of the primitives on the right-hand side, the direction of the quark line $(j = i)$ has been reversed. Not all of the primitives generated by the ordered permutations are nonzero—the shuffle product can give rise to some configurations with crossed quark lines. As with the expressions in the main body of this paper, Eq. (A2) should be dressed with appropriate minus signs as dictated by Fermi statistics.

If gluons are also present, the identity generalizes as follows. Let g_1 be the set of gluons in between x and the set $\{\alpha_1\}$, g_2 be the set of gluons in between the set $\{\alpha_1\}$ and $\{\alpha_2\}$, and so on. The most general primitive amplitude takes the form

$$\mathcal{A}(\dots x\mathbf{g}_1\{\alpha_1\}\mathbf{g}_2\{\alpha_2\}\dots\mathbf{g}_{m-1}\{\alpha_{m-1}\}\mathbf{g}_m j\{\beta\}i\mathbf{g}_{m+1}\{\alpha_{m+1}\}\dots\mathbf{g}_s\{\alpha_s\}\mathbf{g}_{s+1}y\dots). \quad (\text{A3})$$

The first and last elements of the sets $\{\alpha_i\}$ are by construction quark labels, but we now allow for the possibility that any other element of the sets $\{\alpha_i\}$ as well as any element of the set $\{\beta\}$ is a gluon label. Denote the first i elements in the set \mathbf{g}_c as the set \mathbf{g}_c^i , and the remaining $n_c - i$ elements as $\mathbf{g}_c^{n_c-i}$, where n_c is the number of gluons in the set \mathbf{g}_c . The generalization of Eq. (A2) is

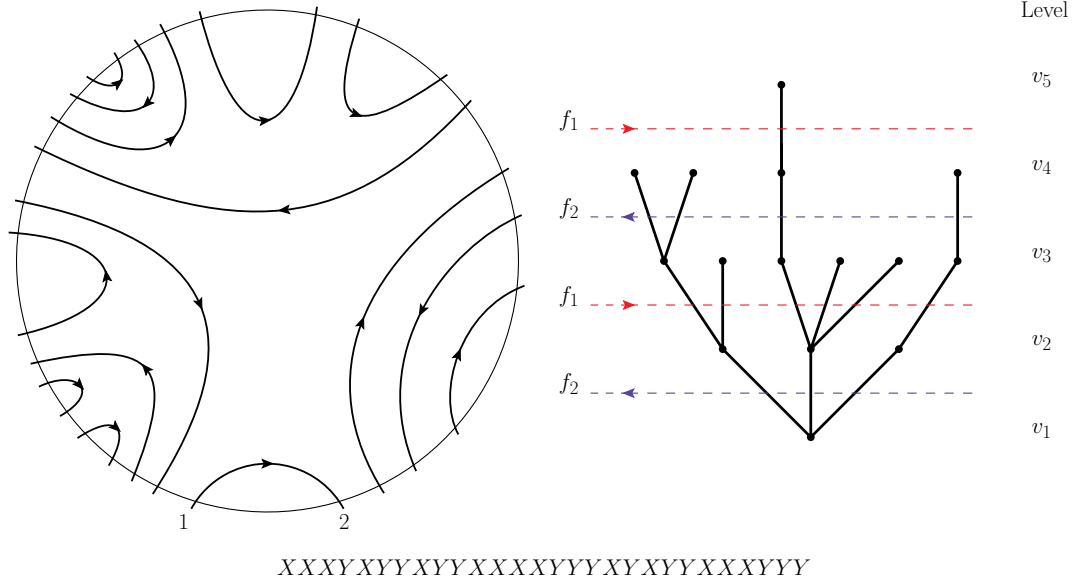


FIG. 6 (color online). The quark line graph on the left-hand side shows how to choose the signature of the Dyck permutation (i.e. quark line direction) so as to create alternating quark and antiquark external states. The right-hand side is the corresponding Dyck tree. The dotted lines drawn across the edges of the Dyck tree have an arrow to indicate the direction of the arrows on the quark lines which cross them to create the quark line graph shown on the left, and have a flavor label f_1 or f_2 to indicate which flavor the quark lines should have in the two-flavor recursion described in the text.

$$\begin{aligned}
& \mathcal{A}(\dots x \mathbf{g}_1 \{\alpha_1\} \mathbf{g}_2 \{\alpha_2\} \dots \mathbf{g}_{m-1} \{\alpha_{m-1}\} \mathbf{g}_m j \{\beta\} i \mathbf{g}_{m+1} \{\alpha_{m+1}\} \dots \mathbf{g}_s \{\alpha_s\} \mathbf{g}_{s+1} y \dots) \\
&= (-1)^{1+n_{g\beta}} \sum_{c=1}^m \left[\sum_{i=0}^{n_{g_c}} \left[\sum_{\text{OP}\{D_c^i\}\{E\}} \left(\sum_{\text{OP}\{A_c^i\}\{B\}} \mathcal{A}(\dots x \mathbf{g}_1 \{\alpha_1\} \dots \mathbf{g}_{c-1} \{\alpha_{c-1}\} \right. \right. \right. \\
& \quad \left. \left. \left. \times \underbrace{\mathbf{g}_c^i i \mathbf{g}_c^{n_c-i} \{\alpha_c\} \dots \mathbf{g}_{m-1} \{\alpha_{m-1}\} \mathbf{g}_m \{\beta^T\} j}_{\{A_c^i\} \{B\}} \underbrace{\mathbf{g}_{m+1} \{\alpha_{m+1}\} \dots \mathbf{g}_s \{\alpha_s\} \mathbf{g}_{s+1} y \dots}_{\{E\}} \right) \right] \right], \tag{A4}
\end{aligned}$$

where $n_{g\beta}$ is the number of gluon labels contained in the set $\{\beta\}$. The proof of this equation is a straightforward generalization of the proof of Eq. (A2), given in [1]. The quark line structure generated in Eq. (A4) is exactly the same as the quark line structure generated in Eq. (A2), the only difference being there are more terms involving each quark line structure, with different placings of gluons which now enter the shuffle. As in Eq. (A2), some of the terms generated by the shuffle products will be zero because of quark line constraints.

APPENDIX B: A TWO-FLAVOR RECURSION TO OBTAIN QCD FROM $\mathcal{N} = 4$ SYM

In this appendix we briefly outline a flavor recursion based on a different basis than the all-positive signature recursion of Sec. III. The choice of signature for each Dyck topology is the one which results in alternating quarks and antiquarks for the external particles. In this way the signature of each primitive will depend on its Dyck topology. An

example graph is shown in Fig. 6. The reason for choosing such a basis is that it is another way to prevent any possible scalar exchange between gluino lines in the $\mathcal{N} = 4$ SYM amplitudes. Consider following an internal scalar line up until the point at which it terminates, as shown in Fig. 5. Now we can relax the requirement that all external quarks (we use quark in this section, but this should be identified with a gluino in the $\mathcal{N} = 4$ SYM context) are the same flavor, so that such a scalar contribution to the primitive could exist. However, if the alternating basis is chosen, then since the termination point is two quarks (or two antiquarks) there will remain either an odd number of quarks or an odd number of antiquarks between the two shown in Fig. 5. These cannot possibly be joined up in a planar way, and it follows that there can be no such scalar contributions.

The flavor recursion proceeds in a similar fashion as the one described in Sec. III. Since the direction of the quark lines alternate between levels in the Dyck tree, now two flavors must be used to stop the flavor mixing between odd

and even levels of the Dyck tree. That is, the distinct-flavor primitive in the alternating basis is expressed as a two-flavor primitive of the same labeling, with one flavor, f_1 , for quarks corresponding to odd levels of the Dyck tree, and the other flavor, f_2 , assigned to even levels of the Dyck tree—see the right-hand side of Fig. 6—minus distinct-flavor subtraction primitives to subtract the wrong flavor running. It is clear that flavor pairing cannot occur between odd and even levels. Considering now just the odd levels, the same arguments as explained in Sec. III (see Fig. 4) apply to these quark lines of flavor f_1 , that is, the nonzero

permutations only occur at the same level of the Dyck tree. The same applies to quark lines at even levels. Each of the subtraction primitives is of higher maturity, and needs to be re-expressed in terms of primitives of alternating basis form. In this way, the recursion terminates at primitives of the form

$$\mathcal{A}(1_{f_1} q_{f_2} \bar{q}_{f_1} q_{f_2} \bar{q}_{f_1} \dots q_{f_1} \bar{q}_{f_2} q_{f_1} \bar{q}_{f_2} 2_{f_1}). \quad (\text{B1})$$

These are the primitives which have a Dyck tree of the highest maturity (shown on the right-hand side of Fig. 3).

-
- [1] T. Melia, *Phys. Rev. D* **88**, 014020 (2013).
 [2] F. A. Berends and W. Giele, *Nucl. Phys.* **B294**, 700 (1987).
 [3] M. L. Mangano, S. J. Parke, and Z. Xu, *Nucl. Phys.* **B298**, 653 (1988).
 [4] R. K. Ellis, Z. Kunszt, K. Melnikov, and G. Zanderighi, *Phys. Rep.* **518**, 141 (2012).
 [5] H. Ita and K. Ozeren, *J. High Energy Phys.* **02** (2012) 118.
 [6] S. Badger, B. Biedermann, P. Uwer, and V. Yundin, *Comput. Phys. Commun.* **184**, 1981 (2013).
 [7] C. Reuschle and S. Weinzierl, *Phys. Rev. D* **88**, 105020 (2013).
 [8] R. Kleiss and H. Kuijf, *Nucl. Phys.* **B312**, 616 (1989).
 [9] Z. Bern, J. Carrasco, and H. Johansson, *Phys. Rev. D* **78**, 085011 (2008).
 [10] J. Drummond and J. Henn, *J. High Energy Phys.* **04** (2009) 018.
 [11] S. J. Parke and T. Taylor, *Phys. Lett. B* **157**, 81 (1985).
 [12] Z. Kunszt, *Nucl. Phys.* **B271**, 333 (1986).
 [13] L. J. Dixon, J. M. Henn, J. Plefka, and T. Schuster, *J. High Energy Phys.* **01** (2011) 035.
 [14] G. Aad *et al.* (ATLAS Collaboration), *J. High Energy Phys.* **10** (2013) 130.
 [15] CMS Collaboration, Report No. CMS-PAS-SUS-13-012, 2013.
 [16] T. Gleisberg and S. Hoeche, *J. High Energy Phys.* **12** (2008) 039.
 [17] Z. Bern, L. J. Dixon, F. Febres Cordero, S. Höche, H. Ita, D. A. Kosower, D. Maître, and K. J. Ozeren, *Phys. Rev. D* **88**, 014025 (2013).
 [18] S. Badger, B. Biedermann, P. Uwer, and V. Yundin, *Phys. Rev. D* **89**, 034019 (2014).
 [19] Z. Bern, L. J. Dixon, D. C. Dunbar, and D. A. Kosower, *Nucl. Phys.* **B425**, 217 (1994).
 [20] Z. Bern, L. J. Dixon, and D. A. Kosower, *Nucl. Phys.* **B513**, 3 (1998).
 [21] R. Britto, F. Cachazo, and B. Feng, *Nucl. Phys.* **B725**, 275 (2005).
 [22] D. Forde, *Phys. Rev. D* **75**, 125019 (2007).
 [23] G. Ossola, C. G. Papadopoulos, and R. Pittau, *Nucl. Phys.* **B763**, 147 (2007).
 [24] R. K. Ellis, W. Giele, and Z. Kunszt, *J. High Energy Phys.* **03** (2008) 003.
 [25] W. T. Giele, Z. Kunszt, and K. Melnikov, *J. High Energy Phys.* **04** (2008) 049.
 [26] C. Berger, Z. Bern, L. Dixon, F. Cordero, D. Forde, H. Ita, D. Kosower, and D. Maître, *Phys. Rev. D* **78**, 036003 (2008).
 [27] P. Mastrolia and G. Ossola, *J. High Energy Phys.* **11** (2011) 014.
 [28] D. A. Kosower and K. J. Larsen, *Phys. Rev. D* **85**, 045017 (2012).
 [29] S. Badger, H. Frellesvig, and Y. Zhang, *J. High Energy Phys.* **04** (2012) 055.
 [30] Y. Zhang, *J. High Energy Phys.* **09** (2012) 042.
 [31] S. Badger, H. Frellesvig, and Y. Zhang, *J. High Energy Phys.* **08** (2012) 065.
 [32] H. Johansson, D. A. Kosower, and K. J. Larsen, arXiv:1308.4632.
 [33] M. Sogaard and Y. Zhang, *J. High Energy Phys.* **12** (2013) 008.
 [34] S. Frixione, *J. High Energy Phys.* **09** (2011) 091.
 [35] S. Frixione, Z. Kunszt, and A. Signer, *Nucl. Phys.* **B467**, 399 (1996).
 [36] J. Gallicchio and M. D. Schwartz, *Phys. Rev. Lett.* **107**, 172001 (2011).
 [37] A. J. Larkoski, G. P. Salam, and J. Thaler, *J. High Energy Phys.* **06** (2013) 108.
 [38] ATLAS Collaboration, Report No. ATLAS-CONF-2011-053, 2011.
 [39] S. Chatrchyan *et al.* (CMS Collaboration), *J. High Energy Phys.* **04** (2012) 036.
 [40] CMS Collaboration, Report No. CMS-PAS-JME-13-002, 2013.
 [41] F. Pandolfi, *Eur. Phys. J. Plus* **128**, 117 (2013).
 [42] A. Banfi, G. P. Salam, and G. Zanderighi, *Eur. Phys. J. C* **47**, 113 (2006).
 [43] L. Dai, K. Melnikov, and F. Caola, *J. High Energy Phys.* **04** (2012) 095.
 [44] R. Britto, F. Cachazo, B. Feng, and E. Witten, *Phys. Rev. Lett.* **94**, 181602 (2005).
 [45] A. Hodges, *J. High Energy Phys.* **05** (2013) 135.
 [46] J. Drummond and L. Ferro, *J. High Energy Phys.* **07** (2010) 027.
 [47] H. Elvang and Y.-t. Huang, arXiv:1308.1697.
 [48] S. Badger, B. Biedermann, L. Hackl, J. Plefka, T. Schuster, and P. Uwer, *Phys. Rev. D* **87**, 034011 (2013).
 [49] T. Schuster, arXiv:1311.6296.