

**Dyck words and multiquark primitive amplitudes**

Tom Melia\*

*CERN Theory Division, CERN, CH-1211, Geneva 23, Switzerland*

(Received 30 April 2013; published 15 July 2013)

I study group theory (Kleiss-Kuijf) relations between purely multiquark primitive amplitudes at tree level and prove that they reduce the number of independent primitives to  $(n-2)!/(n/2)!$ , where  $n$  is the number of quarks plus antiquarks, in the case where quark lines have different flavors. I give an explicit example of an independent basis of primitives for any  $n$  which is of the form  $\mathcal{A}(1, 2, \sigma)$ , where  $\sigma$  is a permutation based on a Dyck word.

DOI: [10.1103/PhysRevD.88.014020](https://doi.org/10.1103/PhysRevD.88.014020)

PACS numbers: 12.38.Bx, 12.38.–t

**I. INTRODUCTION**

Color decompositions have proven to be an extremely useful tool in state-of-the-art calculations of multijet cross sections at particle colliders (for a review see, e.g., Ref. [1]). They allow for the definition of purely kinematic, gauge invariant objects, which in their most basic form have cyclically ordered external legs and are called primitive amplitudes [2,3]. Cyclic ordering leads to a simplified dependence on kinematic quantities, and as a result of this, primitive amplitudes are well suited for the application of on-shell techniques such as Britto-Cachazo-Feng-Witten recursion [4] and the most developed formulations of unitarity techniques (such as those based on generalized unitarity [5–8] and the numerical Ossola-Papadopoulos-Pittau reduction procedure [9–12]). Knowledge of how the color information is reintroduced can be used to calculate only those primitive amplitudes which will contribute to a particular order in a  $1/N_c$  expansion, which is particularly useful when computing one-loop high-multiplicity amplitudes since the very time consuming yet numerically small, subleading color parts can be neglected. Color decompositions have also recently been applied to non-Abelian gauge theories which are spontaneously broken [13]. Different methods of dealing with color quantum numbers, such as Monte Carlo summation over fixed external colors [14–17], and color dressed recursive techniques [18,19], are highly efficient at tree level; a recent study has developed this technique for use with virtual amplitudes [20].

In this paper I consider relations between primitive amplitudes composed of many massless antiquark-quark pairs at tree level. No all- $n$  formula is known to relate these primitives to the full amplitude, but examples have been worked out at tree and loop level (also with external gluons) by equating coefficients of Feynman diagrams [1,21]—in Ref. [21] a color decomposition in terms of one-loop primitives involving up to six quarks and one gluon was given (which is sufficiently complicated that it has to be provided in an attached text file to the paper); see also Ref. [22]. A motivation of this paper is the hope that a

better understanding of these objects and the relations they satisfy could lead to an all- $n$  color decomposition in terms of them.

As a way of introducing the multiquark primitive amplitudes, consider first the trace-based color decomposition for tree-level  $n$  gluon scattering amplitudes in terms of fundamental  $SU(N_c)$  matrices,  $\lambda^a$ , [23,24],

$$\begin{aligned} \mathcal{M}_{\text{gluon}}^{\text{tree}}(g_1, g_2, \dots, g_n) \\ = \sum_{\mathcal{P}(2\dots n)} \text{tr}(\lambda^1 \lambda^2 \dots \lambda^n) \mathcal{A}_{\text{gluon}}(1, 2, \dots, n), \end{aligned} \quad (1)$$

where the sum is over all permutations  $\mathcal{P}$  of  $2\dots n$ . The labelling of the primitive amplitudes  $\mathcal{A}_{\text{gluon}}$  attests to the fact that the only Feynman diagrams which contribute to them, when drawn in a planar fashion, have a cyclic ordering of the external legs which corresponds to the label. The  $\mathcal{A}_{\text{gluon}}$  inherit many properties from this color decomposition, and relations exist between them (which I review in Sec. II). In particular, the Kleiss-Kuijf (KK) relations [25] reduce the number of linearly independent primitive amplitudes from the  $(n-1)!$  in Eq. (1) to  $(n-2)!$ . Bern-Carrasco-Johansson relations [26] further reduce the number of independent primitives to  $(n-3)!$ , but I will not discuss this color-kinematic duality in this paper and will use the term independent to implicitly mean independent over the field of real numbers. Primitive amplitudes can be obtained directly using color-ordered Feynman rules (see, e.g., Ref. [27]) that assign purely kinematic factors to vertices between particles which are antisymmetric under exchange of two of the legs.

The multiquark primitive amplitudes I will consider in this paper can be defined by considering the scattering of  $n$  external massless antiquarks plus quarks in the adjoint representation of  $SU(N_c)$ —then the amplitude has the same color decomposition as Eq. (1), and I study relations between the corresponding primitive amplitudes  $\mathcal{A}_{\text{quark}} \equiv \mathcal{A}$ . They inherit all of the relations that the all-gluon amplitudes do from this color decomposition, but they are evaluated with different color-ordered Feynman rules, and the presence of quark lines which require quark number and flavor conservation at their interaction vertices

\*thomas.melia@cern.ch

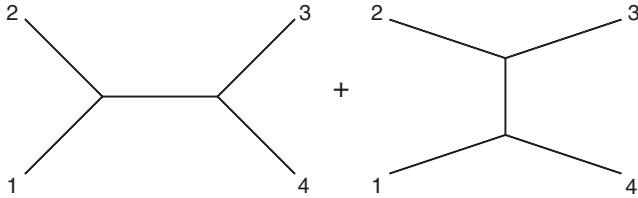


FIG. 1. The planar graphs contributing to a four particle primitive amplitude with cyclic ordering 1234.

introduces an interesting structure which gives rise to additional relations between the primitive amplitudes.

As a simple example of this, consider the planar graphs which constitute a four particle primitive amplitude with the cyclic ordering 1234 (see Fig. 1). If all four particles are gluons, then both graphs are nonzero and contribute. When swapping gluons 2 and 3 around to give the contributions to the primitive 1324, the second diagram simply picks up a negative sign through the color-ordered Feynman rules, whereas the kinematic structure of the first diagram changes to something different. Next consider the case where particle 1 is an up quark, particle 4 is an up antiquark, particle 2 is a down quark, and particle 3 is a down antiquark. The first graph is zero because it violates both flavor and quark number at its vertices, whereas the second graph is allowed (the exchanged particle is a gluon); now when particles 2 and 3 are swapped around, the first diagram is still zero (this time only due to flavor violation at the vertices), and the second diagram picks up a negative sign. This leads to a relation between the primitive amplitudes  $1234 = -1324$  which is not present in the all-gluon case. This type of relation was observed in Ref. [21] arising from nontrivial solutions to linear equations involving Feynman diagrams. As a final example based on Fig. 1, if particles 1 and 3 are an up quark and up antiquark and particles 2 and 4 are a down quark and down antiquark, then both diagrams are zero—there is no planar way to connect the quarks of equal flavor.

In this paper, I will interpret the relations described above as KK relations, and I show that in using them, an independent set of  $(n-2)!/(n/2)!$  primitive amplitudes can be found for the case when all quarks have different flavors. These amplitudes can be constructed using Dyck words (named after the German mathematician Walther von Dyck), which are strings of equal numbers of the letters  $X$  and  $Y$  such that the number of  $X$ s is greater than or equal to the number of  $Y$ s in any initial segment of the string.

This paper is organized as follows. In Sec. II I will review the familiar properties of multi-quark primitive amplitudes and introduce a useful graphical way of representing them. In Sec. III I will describe a correspondence between Dyck words and nonzero primitive amplitudes and provide a proof that the number of primitive amplitudes can be reduced to an independent set of size  $(n-2)!/(n/2)!$  using the KK relations. I conclude in Sec. IV.

## II. MULTIQUARK PRIMITIVE AMPLITUDES

Consider the  $n$  particle tree-level scattering amplitude which has  $n/2$  massless antiquark-quark pairs. I focus on the case where all the pairs have a distinct flavor (where the term flavor can be used loosely in the sense that two quark lines which have different helicities have different “flavor” even if they are both up-type, down-type, etc.) and will discuss at the end of this section the equal flavor case. I adopt the convention that odd momentum labels are assigned to antiquarks, even momentum labels are assigned to quarks, and that the pairs of  $\bar{q}q$  of equal flavor are  $(1 \rightarrow 2), (3 \rightarrow 4), \dots, (n-1 \rightarrow n)$ , where I use the notation  $(\bar{q} \rightarrow q)$  to indicate a particular quark line. A color decomposition when these quarks are in the adjoint representation is

$$\mathcal{M}_{\text{adj}}^{\text{tree}}(\bar{q}_1, q_2, \dots, \bar{q}_{n-1}, q_n) = \sum_{\mathcal{P}(2\dots n)} \text{tr}(\lambda^1 \lambda^2 \dots \lambda^n) \mathcal{A}(1, 2, \dots, n-1, n), \quad (2)$$

where as discussed in the introduction, the purely kinematic objects  $\mathcal{A}$  are planar, cyclically ordered, primitive amplitudes. This decomposition is to be compared to the one when the (anti)quarks are in the fundamental representation with color indices  $(\bar{i})i$  [24],

$$\mathcal{M}_{\text{fund}}^{\text{tree}}(\bar{q}_1, q_2, \dots, \bar{q}_{n-1}, q_n) = \sum_{\alpha \in \mathcal{S}_{n/2}} \delta_{\bar{i}_1 \alpha_2} \delta_{\bar{i}_3 \alpha_4} \dots \delta_{\bar{i}_{n-1} \alpha_n} \mathcal{B}_\alpha, \quad (3)$$

where the sum runs over all permutations  $\alpha = (\alpha_2, \alpha_4, \dots, \alpha_n)$  of quark indices  $(i_2, i_4, \dots, i_n)$ . The color factors are strings of delta functions, and the  $\mathcal{B}_\alpha$  are purely kinematic functions (often called color-ordered amplitudes, but they are not cyclically ordered). Factors of  $1/N_c$  associated with a particular permutation  $\alpha$  have been absorbed into the definition of the  $\mathcal{B}_\alpha$ . As already mentioned, the  $\mathcal{B}$  can be expressed in terms of the  $\mathcal{A}$  by solving linear systems of equations defined by Feynman diagram expansions, but no all- $n$  formula is known (low multiplicity cases can be obtained using the color decompositions described in Ref. [28]).

Some of the relations between the  $\mathcal{A}$  in Eq. (2) are familiar from gluon amplitudes. They are gauge invariant, are invariant under cyclic permutations of  $1 \dots n$ , possess a reflection symmetry  $\mathcal{A}(1, 2, \dots, n) = (-1)^n \mathcal{A}(n, \dots, 2, 1)$ , and satisfy the KK relation, which may be written in the form

$$\mathcal{A}(1, \{\beta\}, 2, \{\alpha\}) = (-1)^{n_\beta} \sum_{\text{OP}\{\alpha\}\{\beta^T\}} \mathcal{A}(1, 2, \{\alpha\}\{\beta^T\}), \quad (4)$$

where  $\{\alpha\} \cup \{\beta\} = \{3, \dots, n-1, n\}$ ,  $\{\beta^T\}$  is the set  $\{\beta\}$  with the ordering of the elements reversed,  $n_\beta$  is the number of elements in  $\{\beta\}$ , and  $\text{OP}\{\alpha\}\{\beta^T\}$  stands for “ordered permutations,” which are the shuffle product of the elements of the sets  $\{\alpha\}$  and  $\{\beta^T\}$ —all permutations of

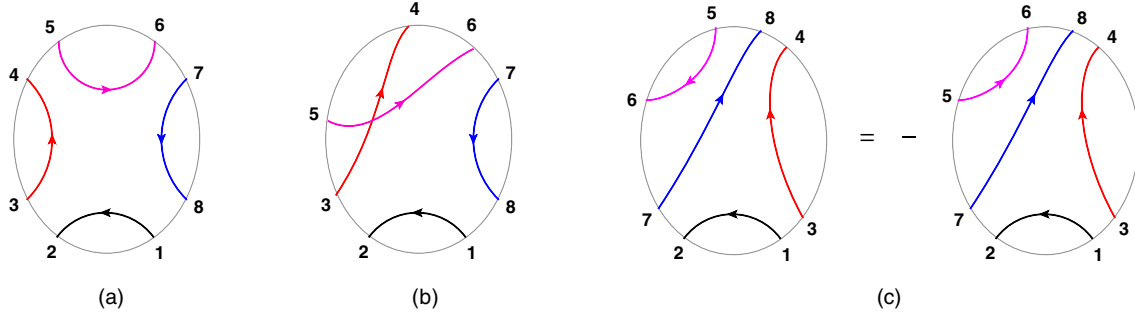


FIG. 2 (color online). (a) Quark line graph for the permutation  $\sigma = (12345678)$ ; (b) an example of a permutation giving rise to crossed quark lines—the corresponding primitive amplitude is zero; (c) a relation between primitive amplitudes.

the union of the two sets which keep fixed the ordering of the  $\alpha_i$  within  $\{\alpha\}$  and the  $\beta_i$  within  $\{\beta^T\}$ .

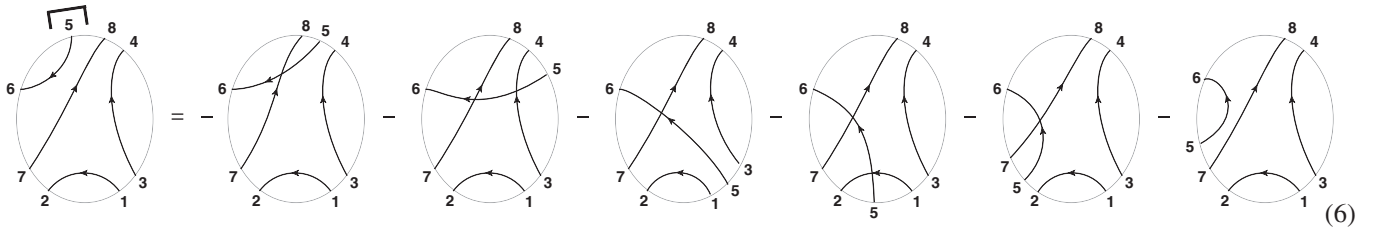
Consider the set of primitive amplitudes with the labels for antiquark 1 and quark 2 fixed adjacent to each other in the order 12, but allowing for all permutations of the labels  $3 \dots n$ —I call this set the  $\mathcal{A}(1, 2, \sigma)$  basis (I discuss the implications of fixing this choice of labels in Sec. IV). All other primitive amplitudes can be expressed in terms of elements of this set through Eq. (4). A useful way to represent the  $\mathcal{A}$  graphically is shown in Fig. 2. A light grey circle is drawn to indicate the edge of the plane (for clarity, it does not mean any kind of trace is taken), and the quark labels are written clockwise around this circle in the order dictated by the particular permutation. Quark lines are then drawn to join  $(1 \rightarrow 2)$ ,  $(3 \rightarrow 4)$ , etc.. These quark line graphs make clear the structure of the Feynman diagrams (computed using color-ordered rules) which contribute to a given

$\mathcal{A}$ —they are the tree diagrams which arise from joining the quark lines together with gluons in all distinct planar ways. They also make it easy to see that some  $\mathcal{A}(1, 2, \sigma)$  with certain permutations of  $3 \dots n$  are zero. These are the ones where quark lines cross, since there is no planar way in which to connect the crossed antiquark-quark pairs. An example of such a permutation is shown in Fig. 2(b).

Figure 2(c) shows a relation between primitive amplitudes in the  $\mathcal{A}(1, 2, \sigma)$  basis (and which does not hold for all-gluon amplitudes),

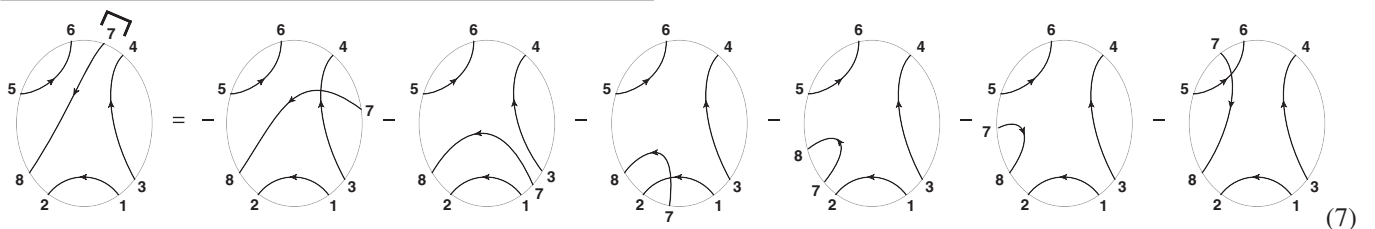
$$\mathcal{A}(1, 2, 7, 6, 5, 8, 4, 3) = -\mathcal{A}(1, 2, 7, 5, 6, 8, 4, 3). \quad (5)$$

This is really a KK relation with many of the primitive amplitudes entering it being zero, so they are not written explicitly in Eq. (5). Writing the relation graphically using quark line graphs, and explicitly showing the zero primitive amplitudes (which have crossed quark lines), Eq. (5) becomes



where the bracket on the diagram on the lhs denotes which leg is to be associated with  $\beta$  on the lhs of Eq. (4). For this particular case, it is easy to understand the relation Eqs. (5) and (6) in terms of the Feynman diagram expansion of the two nonzero primitive amplitudes using the color-ordered rules. This is because the quark line  $(5 \rightarrow 6)$  is isolated in a planar sense from the rest of the diagram, with the line  $(7 \rightarrow 8)$

acting as a boundary to a zone in which it resides. In every contributing Feynman diagram to  $\mathcal{A}(\dots 6, 5 \dots)$ ,  $(5 \rightarrow 6)$  connects only to  $(7 \rightarrow 8)$ , via a single gluon line, and the  $6g5$  vertex can be flipped trivially, picking up a minus sign through the color-ordered Feynman rules, to produce the corresponding Feynman diagram of  $\mathcal{A}(\dots 5, 6 \dots)$ . This is to be compared with the case that arises from the KK relation,



and relates four nonzero elements of the  $\mathcal{A}(1, 2, \sigma)$  basis (the ones without crossed quark lines). The interplay of the color-ordered Feynman diagrams in the expansions of each of these primitive amplitudes is now more complicated, owing to the fact that the quark line ( $7 \rightarrow 8$ ) is not isolated and has to be joined with gluons to at least two of the other quark lines and that this can be done in a number of different ways. I will not consider Feynman diagram expansions of the primitive amplitudes in this paper and will instead use the KK relations such as Eqs. (5)–(7) directly in order to address the question of how such relations impact on the number of independent multiquark primitive amplitudes.

As presented above, the extra structure that having quark lines brings is to alter the KK relations so that terms drop out, so that it is possible to reexpress a primitive amplitude in the  $\mathcal{A}(1, 2, \sigma)$  basis in terms of others which are all solely in the  $\mathcal{A}(1, 2, \sigma)$  basis. This never happens with all-gluon amplitudes, but having the additional quark line structure introduces the possibility that the amplitudes in a KK relation which are not in the  $\mathcal{A}(1, 2, \sigma)$  basis can be zero due to crossed quark lines (this is the case with the  $\mathcal{A}(1, 5, 2, \dots)$  in Eq. (6) and the  $\mathcal{A}(1, 7, 2, \dots)$  in Eq. (7). A question of interest is how many independent  $\mathcal{A}(1, 2, \sigma)$  elements are left after all such relations are taken into account. An equivalent way of posing this question is of asking which independent basis of  $(n - 2)!$  primitive amplitudes, chosen before quark line considerations, is such that as many of the primitives as possible are zero once the quark lines are taken into account. In starting with the  $\mathcal{A}(1, 2, \sigma)$  basis (chosen because it is initially a clearly independent set), the further relations between the elements of this basis detail the extent to which this is not a basis of initial  $(n - 2)!$  independent primitives which maximises the number of zero primitives when the conservation rules at quark-gluon vertices are enforced. It is this question (from the first point of view) that I will address in Sec. III.

Up until this point, and for the remainder of the paper (except for a brief discussion in Sec. IV), I am considering the case where all quark lines are of different flavor. The amplitude where  $n_e$  quark lines have the same flavor,  $\mathcal{M}_{n_e}^{\text{tree}}$ , can be obtained using the all distinct

flavor amplitude ( $\mathcal{M}_{n_e}^{\text{tree}}$  in the below) by a permutation over quarks,

$$\mathcal{M}_{n_e}^{\text{tree}} = \sum_{\mathcal{P}(q_1, q_2, \dots, q_{n_e})} (-1)^{\text{sgn}(\mathcal{P})} \mathcal{M}_{n_e=0}^{\text{tree}}, \quad (8)$$

where the sum is over all of the  $n_e!$  permutations of the equal flavor quark indices,  $q_1 \dots q_{n_e}$ , and the  $(-1)^{\text{sgn}(\mathcal{P})}$  accounts for Fermi statistics. An interesting observation is that when  $n_e = n/2$ , so that all quark lines are of equal flavor, none of the primitive amplitudes in the  $\mathcal{A}(1, 2, \sigma)$  basis are zero. One way of seeing this is that when constructing a quark line graph, for any permutation of  $3 \dots n$ , it is always possible to find a way of joining up the odd numbers with even numbers without crossing lines [any join (odd  $\rightarrow$  even) is allowed as all flavors are the same]. I prove this in Appendix A. No primitive amplitudes drop out of the KK relations, and so the number of independent  $\mathcal{A}_{n_e=n/2}$  is  $(n - 2)!$ , as it is for gluons. This fact, along with Eq. (8) gives a bound on the number of independent primitives with  $n_e = 0$ ,

$$\#(\text{independent } \mathcal{A}_{n_e=0}) \geq (n - 2)! / (n/2)!. \quad (9)$$

In the next section, I will first discuss how to count all of the possible nonzero graphs in the  $\mathcal{A}(1, 2, \sigma)$  basis, and then I will use KK relations (the additional quark line structure setting terms in them to zero) to express this set in terms of one of its subsets of size  $(n - 2)! / (n/2)!$ . Equation (9) then implies that this subset is independent.

### III. DYCK WORDS

A Dyck word is a string of length  $2r$  consisting of  $r$  Xs and  $r$  Ys, such that the number of Xs is greater than or equal to the number of Ys in any initial segment of the string. The number of Dyck words of length  $2r$  is given by the  $r$ th Catalan number,  $C_r = (2r)! / (r + 1)! r!$ . For example, for  $r = 3$  there are five Dyck words:

$$\text{XXXYYY} \quad \text{XXYXY} \quad \text{XXYYXY} \quad \text{XYXXYY} \quad \text{XYXYXY}. \quad (10)$$

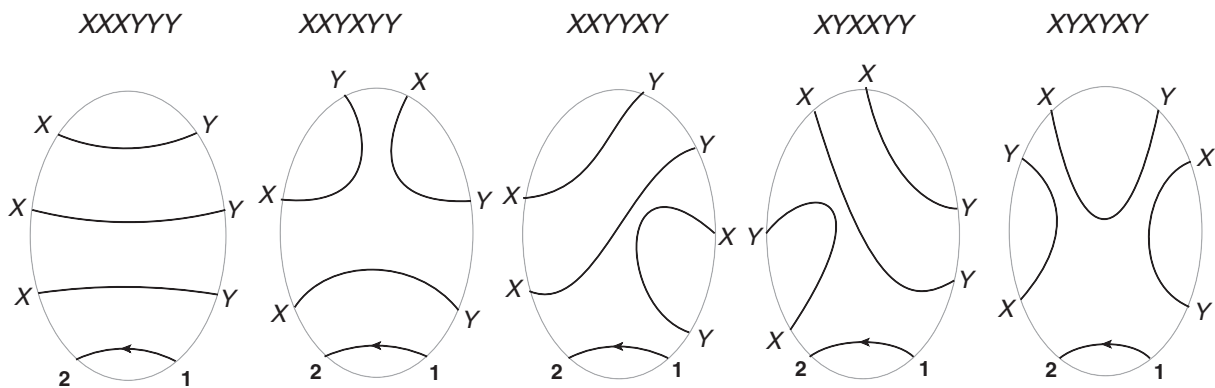


FIG. 3. Dyck words for  $r = 3$  (top row) and the quark line graph topologies they describe.



A topology for a quark line graph can be associated with a Dyck word in the following way (see Fig. 3). First draw in the line ( $1 \rightarrow 2$ ); then, moving clockwise from this line, write the Dyck word around the edge of the plane, and each time a  $Y$  is encountered, connect it with a line to the most recently written  $X$  which has not already been connected. This procedure does not require any quark lines to cross, and the Dyck words of length  $2r$  provide all possible noncrossing topologies of the  $r = n/2 - 1$  quark lines coming from the  $\sigma$  permutations in the  $\mathcal{A}(1, 2, \sigma)$  basis for  $n$  quark scattering. That this is true can be seen from the interpretation of Dyck words as strings of correctly nested parentheses:  $X \rightarrow "("$  and  $Y \rightarrow ")"$ . Identifying these parentheses with the two ends of a quark line, it is clear that the correct nesting requirement is equivalent to avoiding crossed lines.

For each of these topologies, the quarks of different flavors can be assigned in  $r!$  ways, and each quark line can be directed in one of two ways, so the number of nonzero amplitudes in the  $\mathcal{A}(1, 2, \sigma)$  basis is

$$2^r r! C_r = 2^{\frac{n}{2}-1} (n-2)! / (n/2)! \quad (11)$$

However, not all of these primitive amplitudes are independent. I shall now show that an independent subset can be chosen defined as those primitives where the quark lines are all oriented in the same direction as the quark line ( $1 \rightarrow 2$ ). That is, when reading clockwise around the quark line graph, starting at antiquark 1 [or reading the label in  $\mathcal{A}(1, 2, \sigma)$  from left to right] the antiquark of each quark line is encountered before the quark. This removes the factor of  $2^{\frac{n}{2}-1}$  in the above equation and is a consequence of using KK relations to express the other primitives in terms of members of this oriented subset [and other primitives of the form  $\mathcal{A}(1\sigma'2\sigma)$  which are zero, where  $\sigma \cup \sigma' = \{3, 4, \dots, n\}$  and  $\sigma' \neq \emptyset$ ]. As an example, the oriented primitives corresponding to the topologies shown in Fig. 3 would have antiquark labels assigned to the  $X$ s and quark labels assigned to the  $Y$ s. In the following, I will use the term orient to mean reexpress a primitive amplitude with wrongly directed quark lines in terms of primitive amplitudes where the quark lines point in the same direction as the line ( $1 \rightarrow 2$ ). I want to emphasize that this is not related to the charge parity equation,

$$\mathcal{A}(\dots i_q^\pm \dots j_q^\mp \dots) = -\mathcal{A}(\dots i_q^\mp \dots j_q^\pm \dots), \quad (12)$$

which assigns a different momentum and helicity ( $\pm$ ) to the quark  $q$  on either side of the equation.

Figure 4 depicts a zone bounded by the quark line with ends labelled  $x$  and  $y$  (direction not specified) within a generic primitive amplitude. (The ellipses outside this zone stand for any number of quark lines, which may or may not straddle this zone—a few example lines are shown, their direction is not specified.) Within the zone are  $s$  subzones  $\{\alpha_1\} \dots \{\alpha_s\}$  which consist of a quark line boundary and in general contain further substructure in the form of more quark lines (their boundaries are drawn in Fig. 4, but without specifying their direction for the subzones  $\{\alpha_1\}, \dots, \{\alpha_{m-1}\}, \{\alpha_{m+1}\}, \dots, \{\alpha_s\}$ ). The boundary of the subzone

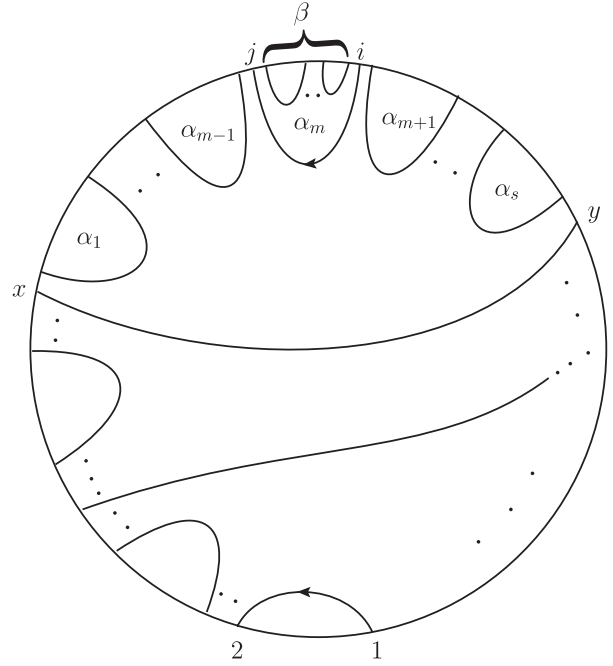


FIG. 4. A quark line graph for a primitive multi-quark amplitude, showing the general structure of a zone with a boundary of the quark line with ends labelled  $x$  and  $y$ , and with sub-zones  $\{\alpha_1\} \dots \{\alpha_s\}$ . The orientation of the boundary ( $i \rightarrow j$ ) of sub-zone  $\{\alpha_m\}$  is explicitly shown to be in the wrong direction, and the substructure of  $\{\alpha_m\}$  is also explicitly shown and labelled  $\{\beta\}$ .

$\{\alpha_m\}$  is wrongly oriented, and this subzone is shown explicitly broken down into its boundary [the line ( $i \rightarrow j$ )] and its substructure  $\{\beta\}$ .

I prove by induction that this zone can be oriented, by which I mean all of the quark lines inside the boundary ( $x - y$ ) (and not the boundary itself) can be oriented. For this, the following identity, a consequence of the KK relations and valid for multi-quark primitive amplitudes with different flavor quark lines (and which I prove in Appendix B'), is useful,

$$\begin{aligned} & \mathcal{A}(\dots x\{\alpha_1\} \dots \{\alpha_{m-1}\} j\{\beta\} i\{\alpha_{m+1}\} \dots \{\alpha_s\} y \dots) \\ &= - \sum_{c=1}^m \left[ \sum_{\text{OP}\{D_c\}\{E\}} \left( \sum_{\text{OP}\{A_c\}\{B\}} \mathcal{A}(\dots x\{\alpha_1\} \dots \{\alpha_{c-1}\} \right. \right. \\ & \quad \left. \left. \times i\{\alpha_c\} \dots \{\alpha_{m-1}\} \underbrace{\{\beta\}}_{\{B\}} j\{\alpha_{m+1}\} \dots \{\alpha_s\} y \dots \right) \right], \quad (13) \end{aligned}$$

where  $\{\alpha_i\}$  are the subzones, and it should be remembered that some of the permutations induced by  $\text{OP}\{A_c\}\{B\}$  and  $\text{OP}\{D_c\}\{E\}$  will be ones with crossed quark lines and will give rise to primitive amplitudes that are zero.

A zone which contains  $k = 1$  quark lines is oriented trivially as follows from Eq. (13) with  $s = 1$ ,  $\{\beta\} = \emptyset$ ,

$$\mathcal{A}(\dots xjiy\dots) = -\mathcal{A}(\dots xijy\dots), \quad (14)$$

which is just the situation discussed beneath Eq. (6), since the quark line ( $i \rightarrow j$ ) is completely isolated from the rest of the diagram by the boundary ( $x - y$ ). Now assume that a zone containing  $k - 1$  quark lines can be oriented. I will prove that a zone containing  $k$  quark lines can be orientated.

If these  $k$  quark lines are arranged so that the number of subzones  $s = 1$ , the zone can be oriented as follows from Eq. (13) with  $s = 1$ ,

$$\mathcal{A}(\dots xj\{\beta\}iy\dots) = -\mathcal{A}(\dots xi\{\beta^T\}jy\dots). \quad (15)$$

This has correctly oriented the boundary of the subzone, which is all that is needed here (and in the following), for any substructure inside subzones can be oriented by assumption by treating the boundary of the subzone as the boundary of a new zone containing  $k' < k$  quark lines. Now I assume that the case where the  $k$  quark lines are arranged into  $s - 1$  subzones can be oriented and show that the case with  $s$  subzones can be oriented. To do this, apply Eq. (13) in its general form,

$$\begin{aligned} &\mathcal{A}(\dots x\{\alpha_1\}\dots\{\alpha_{m-1}\}j\{\beta\}i\{\alpha_{m+1}\}\dots\{\alpha_s\}y\dots) \\ &= -\mathcal{A}(\dots x\{\alpha_1\}\dots\{\alpha_{m-1}\}i\{\beta^T\}j\{\alpha_{m+1}\}\dots\{\alpha_s\}y\dots) \\ &\quad + \text{terms with smaller } s, \end{aligned} \quad (16)$$

which correctly orients the boundary of subzone  $m$  up to other terms which by assumption can be oriented. Equation (13) can be further applied for any  $m$ ,  $1 \leq m \leq s$ , until all  $s$  subzones have correctly oriented boundaries. This concludes the proof that the quark lines contained within the full zone can all be oriented.

Finally, a special case of the above is the zone which contains  $k = n/2 - 1$  quark lines and which has boundary ( $1 \rightarrow 2$ ) which is by definition correctly oriented. Orienting this zone orients the full primitive amplitude.

The reduced set of  $(n - 2)!/(n/2)!$  amplitudes obtained with the above method are

$$\{\mathcal{A}(1, 2, \sigma) | \sigma \in \text{Dyck}_{n/2-1}[X_{\tau_1} \dots X_{\tau_{n/2-1}} Y \dots Y]\}, \quad (17)$$

where  $\text{Dyck}_r$  means all Dyck words of length  $2r$  composed of the  $X$ s and  $Y$ s, with all  $r!$  possible different labellings of the  $X$ s; i.e.  $\tau = (\tau_1, \dots, \tau_r)$  is a permutation of  $(1, \dots, r)$ . The indices of the letter  $Y$ s are determined by the  $X_{\tau_i}$  they are matched to via the method described at the beginning of this section, which depends on the particular Dyck permutation. If a  $Y$  gets matched to  $X_{\tau_i}$ , then it is labelled as  $Y_{\tau_i}$ . Finally the identification  $X_1 \rightarrow 3, Y_1 \rightarrow 4, X_2 \rightarrow 5, Y_2 \rightarrow 6, \dots, X_{n/2-1} \rightarrow n - 1, Y_{n/2-1} \rightarrow n$  should be made.

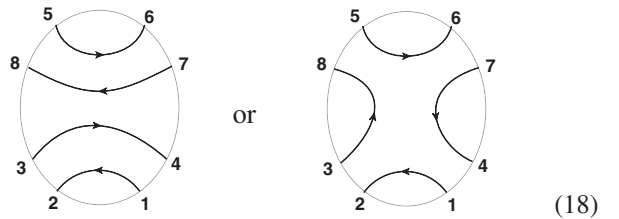
That this set of primitive amplitudes is independent follows from Eq. (9).

#### IV. DISCUSSION

At the beginning of Sec. II, I made a choice of fixing the labels 1 and 2. This choice had an effect on the way in which

the paper progressed, since I have always been considering amplitudes in the  $\mathcal{A}(1, 2, \sigma)$  basis and discussing relations between them. Of course, now that an independent set is found in Eq. (17)—call it  $\{\mathcal{A}(1, 2, \text{Dyck})\}$ —[or this set where the role of ( $1 \rightarrow 2$ ) is switched with another quark line], all of the nonzero amplitudes which result from fixing, say, antiquarks 1 and 3—a  $\mathcal{A}(1, 3, \sigma)$  basis—can be expressed in terms of this set. However, it is an interesting question as to which mixtures of primitive amplitudes from the  $\{\mathcal{A}(1, 2, \text{Dyck})\}$ ,  $\{\mathcal{A}(3, 4, \text{Dyck})\}$ , etc., bases are independent; addressing this question might help understand whether an all- $n$  color decomposition in terms of primitive multiquark amplitudes could be written down.

When carrying out the permutation sum in Eq. (8) on  $\{\mathcal{A}(1, 2, \text{Dyck})\}$  to obtain amplitudes where one or more quark lines have equal flavor, the  $n_e!$  primitive amplitudes at first sight also contain only quark lines which are correctly orientated, since Eq. (8) only permutes the quark indices. There is, however, an ambiguity since with equal flavor quark lines, the quark line graphs are not well defined—for instance if ( $3 \rightarrow 4$ ) and ( $7 \rightarrow 8$ ) have the same flavor, then the graph for the primitive amplitude  $\mathcal{A}(12385674)$  could be drawn one of two ways,



The first of these diagrams is the one which would be drawn in the  $n_e = 0$  case, and it has the line ( $7 \rightarrow 8$ ) incorrectly orientated. The second diagram has both quark lines ( $3 \rightarrow 8$ ) and ( $7 \rightarrow 4$ ) correctly orientated and is what is obtained from the quark line graph for  $\mathcal{A}(12345678)$  by switching the labels  $4 \leftrightarrow 8$  as in Eq. (8). The primitive  $\mathcal{A}(12385674)$  receives contributions from Feynman diagrams associated with both quark line graphs. It would be wrong to try and eliminate it based on the first graph above. The KK relations are affected because fewer terms drop out of them—when the quark lines ( $3 \rightarrow 4$ ) and ( $7 \rightarrow 8$ ) cross, which would produce a zero primitive in the distinct flavor case.

The addition of gluons to these pure-quark amplitudes is phenomenologically important—contributions to collider processes with  $n$  jets which involve quark lines are suppressed in color by a factor of  $1/N_c$  in comparison with the contribution, equal in  $\alpha_s$ , where the quark line is replaced by a gluon pair. The pure-quark primitive amplitudes act as skeletons upon which to fix gluons, and the flavor structure, which this paper investigates, is not altered by their presence. In the future I hope to think about mixed quark-gluon primitive amplitudes, in particular in the context of an all- $n$  color decomposition (an all- $n$  decomposition is known for the one quark line case [29,30]; explicit expressions for the

relation between primitives and the full amplitude for  $4q + 0/1g$  can be found in, e.g., Ref. [31]).

Finally, it would be interesting (and useful in practical collider physics applications) to develop these ideas at one loop. Here, a relation similar to the KK relation at tree level relates nonplanar primitive amplitudes to planar ones. There is also a further division of multiquark primitive amplitudes depending on whether quark lines turn left or right past the loop [2,21]. These amplitudes have already been shown in Ref. [21] to satisfy more relations than would all-gluon amplitudes, owing to their quark line structure. Such relations should be investigated in terms of the one-loop KK relations. Furthermore, the planar one-loop primitives retain the Dyck word structure that is seen here at tree level.

In conclusion, I have presented a proof that KK relations, along with flavor and quark number conservation, reduce the number of independent purely multiquark primitive amplitudes to  $(n-2)!/(n/2)!$  in the case when all quark lines have a different flavor. I have given an explicit set of primitive amplitudes in Eq. (17) for any (even) value of  $n$ , which have a labelling of external particles based on Dyck words.

### ACKNOWLEDGMENTS

I would like to thank Michelangelo Mangano and Franz Herzog for useful discussions. This research was supported by ERC Grant No. 291377.

### APPENDIX A: NONCROSSING QUARK LINES OF EQUAL FLAVOR

In this appendix, I prove the assertion in the text that when all quark lines are of equal flavor, a quark line graph can always be drawn to connect quarks and antiquarks without any quark lines crossing. The proof is simple—it relies on an argument which is similar to the construction of the Dyck topologies in Sec. III. The quark line graph has some permutation of  $n/2$  odd and  $n/2$  even numbers written around the circle that represents the edge of the plane. Any odd

number can be connected to any even number by a quark line, since the quarks and antiquarks all have equal flavor. Pick a starting point at an even number, and move around the circle, and each time an odd number is encountered, connect it with a quark line to the most recently encountered even number that has not already been connected. This procedure leads to no crossed quarks lines, proving the assertion.

### APPENDIX B: PROOF OF THE IDENTITY EQ. (13)

I show how to apply KK relations in order to prove Eq. (13), which is valid for purely multiquark scattering when all quark lines are of a different flavor. See also Fig. 4. First use cyclic symmetry to write the lhs of Eq. (13) so that  $x$  appears as the first label in the primitive amplitude. Then apply the KK relation to  $\{\alpha_1\} \dots \{\alpha_{m-1}\}$  (as indicated in the equation below),

$$\begin{aligned} \mathcal{A}(x \overbrace{\{\alpha_1\} \dots \{\alpha_{m-1}\}}^{KK} j \{\beta\} i \{\alpha_{m+1}\} \dots \{\alpha_s\} y \dots) \\ = \sum_{\text{OP}\{A\}\{B\}} \mathcal{A}(x j \underbrace{\{\alpha_{m-1}^T\} \dots \{\alpha_1^T\}}_{\{A\}} \underbrace{\{\beta\} i \{\alpha_{m+1}\} \dots \{\alpha_s\}}_{\{B\}} y \dots), \end{aligned} \quad (\text{B1})$$

and now use the fact that since the  $\mathcal{A}$  are multiquark amplitudes with quark lines of different flavor, these primitives are zero if an  $\{\alpha_c\}$  straddles the  $i$  (due to crossed quark lines). That is, the ordered permutations  $\text{OP}\{A\}\{B\}$  can be split up as follows:

$$\begin{aligned} = \sum_{c=1}^m \left[ \sum_{\text{OP}\{C_c\}\{E\}} \left( \sum_{\text{OP}\{A_c\}\{B\}} \mathcal{A}(x j \underbrace{\{\alpha_{m-1}^T\} \dots \{\alpha_c^T\}}_{\{A_c\}} \underbrace{\{\beta\}}_{\{B\}} \right. \right. \\ \left. \left. \times i \underbrace{\{\alpha_{c-1}^T\} \dots \{\alpha_1^T\}}_{\{C_c\}} \underbrace{\{\alpha_{m+1}\} \dots \{\alpha_s\}}_{\{E\}} y \dots \right) \right]. \end{aligned} \quad (\text{B2})$$

Now consider a term with fixed  $c$  and for one permutation of  $\text{OP}\{A_c\}\{B\}$  in Eq. (B2), define  $\{\gamma_c\} = \{j\delta\}$ ,  $\delta \in \text{OP}\{A_c\}\{B\}$ , and then apply the KK relation to  $\{\gamma_c\}$  in this term,

$$\begin{aligned} \sum_{\text{OP}\{C_c\}\{E\}} \mathcal{A}(x \overbrace{\{\gamma_c\}}^{KK} i \underbrace{\{\alpha_{c-1}^T\} \dots \{\alpha_1^T\}}_{\{C_c\}} \underbrace{\{\alpha_{m+1}\} \dots \{\alpha_s\}}_{\{E\}} y \dots) = - \sum_{\text{OP}\{D_c\}\{F\}} \left( \sum_{\text{OP}\{C_c\}\{E\}} \mathcal{A}(x i \underbrace{\{\gamma_c^T\}}_{\{D_c\}} \underbrace{\{\alpha_{c-1}^T\} \dots \{\alpha_1^T\}}_{\{C_c\}} \underbrace{\{\alpha_{m+1}\} \dots \{\alpha_s\}}_{\{E\}} y \dots) \right) \\ = - \sum_{\text{OP}\{C_c\}\{F\}} \left( \sum_{\text{OP}\{D_c\}\{E\}} \mathcal{A}(x i \underbrace{\{\alpha_{c-1}^T\} \dots \{\alpha_1^T\}}_{\{C_c\}} \underbrace{\{\gamma_c^T\} \{\alpha_{m+1}\} \dots \{\alpha_s\}}_{\{F\}} y \dots) \right) \\ = - \sum_{\text{OP}\{D_c\}\{E\}} \mathcal{A}(x \{\alpha_1\} \dots \{\alpha_{c-1}\} i \underbrace{\{\gamma_c^T\}}_{\{D_c\}} \underbrace{\{\alpha_{m+1}\} \dots \{\alpha_s\}}_{\{E\}} y \dots), \end{aligned} \quad (\text{B3})$$

where the second equality is simply a rewriting of the permutation, and where the last equality follows through the KK relation with  $\{\beta\} = \{\alpha_1\} \dots \{\alpha_{c-1}\}$ . Substituting Eq. (B3) into Eq. (B2) yields the required identity, Eq. (13).

- [1] R. K. Ellis, Z. Kunszt, K. Melnikov, and G. Zanderighi, *Phys. Rep.* **518**, 141 (2012).
- [2] Z. Bern, L. J. Dixon, and D. A. Kosower, *Nucl. Phys.* **B437**, 259 (1995).
- [3] V. Del Duca, L. J. Dixon, and F. Maltoni, *Nucl. Phys.* **B571**, 51 (2000).
- [4] R. Britto, F. Cachazo, B. Feng, and E. Witten, *Phys. Rev. Lett.* **94**, 181602 (2005).
- [5] Z. Bern, L. J. Dixon, D. C. Dunbar, and D. A. Kosower, *Nucl. Phys.* **B425**, 217 (1994).
- [6] Z. Bern, L. J. Dixon, and D. A. Kosower, *Nucl. Phys.* **B513**, 3 (1998).
- [7] R. Britto, F. Cachazo, and B. Feng, *Nucl. Phys.* **B725**, 275 (2005).
- [8] D. Forde, *Phys. Rev. D* **75**, 125019 (2007).
- [9] G. Ossola, C. G. Papadopoulos, and R. Pittau, *Nucl. Phys.* **B763**, 147 (2007).
- [10] R. K. Ellis, W. Giele, and Z. Kunszt, *J. High Energy Phys.* **03** (2008) 003.
- [11] W. T. Giele, Z. Kunszt, and K. Melnikov, *J. High Energy Phys.* **04** (2008) 049.
- [12] C. Berger, Z. Bern, L. Dixon, F. Cordero, D. Forde, H. Ita, D. Kosower, and D. Maître, *Phys. Rev. D* **78**, 036003 (2008).
- [13] L. Dai, K. Melnikov, and F. Caola, *J. High Energy Phys.* **04** (2012) 095.
- [14] F. Caravaglios, M. L. Mangano, M. Moretti, and R. Pittau, *Nucl. Phys.* **B539**, 215 (1999).
- [15] P. Draggiotis, R. H. Kleiss, and C. G. Papadopoulos, *Phys. Lett. B* **439**, 157 (1998).
- [16] M. L. Mangano, M. Moretti, and R. Pittau, *Nucl. Phys.* **B632**, 343 (2002).
- [17] F. Maltoni, K. Paul, T. Stelzer, and S. Willenbrock, *Phys. Rev. D* **67**, 014026 (2003).
- [18] C. Duhr, S. Hoeche, and F. Maltoni, *J. High Energy Phys.* **08** (2006) 062.
- [19] T. Gleisberg and S. Hoeche, *J. High Energy Phys.* **12** (2008) 039.
- [20] W. Giele, Z. Kunszt, and J. Winter, *Nucl. Phys.* **B840**, 214 (2010).
- [21] H. Ita and K. Ozeren, *J. High Energy Phys.* **02** (2012) 118.
- [22] S. Badger, B. Biedermann, P. Uwer, and V. Yundin, *Comput. Phys. Commun.* **184**, 1981 (2013).
- [23] F. A. Berends and W. Giele, *Nucl. Phys.* **B294**, 700 (1987).
- [24] M. L. Mangano, S. J. Parke, and Z. Xu, *Nucl. Phys.* **B298**, 653 (1988).
- [25] R. Kleiss and H. Kuijf, *Nucl. Phys.* **B312**, 616 (1989).
- [26] Z. Bern, J. Carrasco, and H. Johansson, *Phys. Rev. D* **78**, 085011 (2008).
- [27] L. J. Dixon, [arXiv:hep-ph/9601359](https://arxiv.org/abs/hep-ph/9601359).
- [28] M. L. Mangano and S. J. Parke, *Nucl. Phys.* **B299**, 673 (1988).
- [29] D. Kosower, B.-H. Lee, and V. Nair, *Phys. Lett. B* **201**, 85 (1988).
- [30] M. L. Mangano, *Nucl. Phys.* **B309**, 461 (1988).
- [31] T. Melia, K. Melnikov, R. Rontsch, and G. Zanderighi, *J. High Energy Phys.* **12** (2010) 053.