Published for SISSA by Springer

Received: *May 6, 2014*
Accepted: *July 20, 2014*
Published: *August 18, 2014*

The structure of n-point one-loop open superstring amplitudes

Carlos R. Mafra[a] **and Oliver Schlotterer**[b]

[a] *DAMTP, University of Cambridge, Wilberforce Road, Cambridge, CB3 0WA, U.K.*
[b] *Max-Planck-Institut für Gravitationsphysik, Albert-Einstein-Institut, Am Mühlenberg 1, 14476 Golm, Germany*

E-mail: c.r.mafra@damtp.cam.ac.uk, olivers@aei.mpg.de

Abstract: In this article we investigate one-loop amplitudes in maximally supersymmetric superstring theory. The non-anomalous part of the worldsheet integrand is presented for any number of massless open-string states. The polarization dependence is organized into the same BRST-invariant kinematic combinations which also govern the leading string correction to tree-level amplitudes. The dimensions of the bases for both the kinematics and the associated worldsheet integrals is found to be the unsigned Stirling number S_3^{n-1} of first kind. We explain why the same combinatorial structures govern on the one hand finite one-loop amplitudes of equal helicity states in pure Yang-Mills theory and on the other hand the color tensors at order α'^2 of the color-dressed tree amplitude.

Keywords: Scattering Amplitudes, Superstrings and Heterotic Strings

ArXiv ePrint: 1203.6215

doi:10.1007/JHEP08(2014)099

JHEP08(2014)099

Contents

JHEP08(2014)099

1 Introduction

In recent years the pure spinor formalism [4] allowed striking compactness in the computations of scattering amplitudes both in string theory [5, 6, 39–41, 48–50, 57, 60, 61] and directly in its field-theory limit [6, 8, 17, 18, 58, 59]. It has been known since the work of Nilsson [82] and Howe [36] and that the use of a pure spinor simplifies the description of $N = 1$ super-Yang-Mills (SYM). With the advent of the pure spinor formalism this rewarding description was put into the context of the full superstring theory with a underlying BRST symmetry and a new kind of superspace [37].

Using the ideas of [27] for the field theory amplitudes, it was suggested in [17] and proven in [6, 8] that BRST invariance together with the propagator structure of cubic diagrams are sufficient to determine tree-level amplitudes of $D = 10$ SYM to any multiplicity. The recursive BRST cohomology method obtained in [8] leads to compact and elegant supersymmetric answers and makes use of so-called BRST building blocks which can be regarded as superspace representatives of cubic diagrams. The *field-theory* techniques of [8] were subsequentely exploited to also calculate the general color-ordered open *superstring* tree amplitudes in [6, 7]. The punchline is that the n-point string amplitudes are written as a sum of $(n-3)!$ field theory subamplitudes dressed by hypergeometric integrals [7].

The problem of computing one-loop amplitudes in open superstring theory has been dealt with since the 1980's, the first successful result at four-points being [76] in the NS sector and [66] in the R sector. In spite of the technical difficulties caused by the spin structure sums required by the RNS model, [65] provides progress towards higher multiplicity up to seven-points. In the context of heterotic theories, five- and six gluon amplitudes as well as their implications for effective actions were analyzed in [67]. Pure spinor techniques have been applied to one-loop scattering in [5, 9, 48, 63], superspace results up to five-points are available from these references. As for two-loop amplitudes, after an amazing effort by D'Hoker and Phong the four-point amplitude was computed within the RNS formalism in [78] (see also [77]). Two-loop calculations using the pure spinor formalism can be found in [40, 41, 49].

Can this BRST line of reasoning within the pure spinor formalism be extended to loop amplitudes? With this intention in mind, in this paper we apply the technique of BRST-covariant building blocks to address one-loop amplitudes in superstring theory. For any number of massless SYM states, we determine the BRST invariant part of their worldsheet integrand which is unaffected by the hexagon anomaly [21, 22]. The complete kinematic factor turns out to be organized in terms of color-ordered tree-level amplitudes at order α'^2 that are dressed with worldsheet functions in a minimal basis. A beautiful harmony in the combinatorics of both ingredients arises. However, evaluating the (worldsheet- and modular) integrals is left for future work, in particular the extraction of field theory loop integrals as $\alpha' \to 0$ along the lines of [26].

JHEP08(2014)099

Superstring theory has proven to be a fruitful laboratory to learn about hidden structures in the S matrix of its low energy field theories. The open superstring did not only inspire the color organization of gauge theory amplitudes but also provided an elegant proof for Bern-Carrasco-Johansson (BCJ) relations among color-ordered tree amplitudes [23, 24], based on monodromy properties on the worldsheet. Another difficult field theory problem which found a string-inspired answer is the explicit construction of local kinematic numerators for gauge theory tree amplitudes which satisfy all the dual Jacobi identities, see [28]. After these tree-level examples of cross-fertilization between superstring and field theory amplitudes, we hope that this work helps to provide further guidelines to organize multileg one-loop amplitudes in maximally supersymmetric SYM in both ten and four dimensions.[1] Even though the low energy behaviour of the worldsheet integrals is not addressed, our result for the kinematic factor heavily constrains the form of these field theory amplitudes. In particular, the gauge invariant kinematic building blocks $C_{1,\dots}$ to be defined later on appear to be a promising starting point to construct kinematic numerators for higher multiplicity. They could potentially generalize the crossing symmetric factor $s_{12}s_{23}\mathcal{A}^{\rm YM}(1,2,3,4)$ omnipresent in multiloop four-point amplitudes of $N=4$ SYM (where $\mathcal{A}^{\rm YM}(1,2,\dots,n)$ denotes the color-ordered n-point tree amplitude in maximally supersymmetric Yang Mills theory).

This paper is organized as follows. In section two, we review the construction of the n-point SYM tree amplitude from first principles. We start with the massless vertex operators in terms of SYM superfields and sketch how their singularity structure give rise to BRST building blocks representing cubic subdiagrams. As we will argue, BRST invariance forces them to pair up such that color-ordered SYM amplitudes emerge. Section three sets the formal foundation for the computation of one-loop amplitudes using the minimal pure spinor formalism. It motivates the construction of a further family of BRST building blocks which is carried out in section four. The fourth section follows a line of reasoning similar to the tree-level review — the BRST variation of the one-loop specific building blocks allows to a priori determine any BRST invariant to be expected in a one-loop computation. Then in section five, these BRST invariants are derived from an explicit conformal field theory (CFT) computation, in particular the associated worldsheet functions are determined. Section six connects the BRST invariants with α' corrections to tree-level amplitudes and explains why their symmetry properties agree with those of finite one-loop amplitudes in pure Yang-Mills theory. Finally, in the last section, we point out that also the color factors present at the α'^2 order of tree amplitudes align into the same combinatorial patterns. This leads to a duality between the worldsheet integrand of one-loop amplitudes and color-dressed tree amplitudes at $\mathcal{O}(\alpha'^2)$.

To give a brief reference to the main results of this work — the final form for the n-point kinematic factor can be found in equation (5.34) whose notation is explained in subsection 5.4. Subsection 6.2 contains the general conversion rule (6.17) between the BRST invariants $C_{1,\dots}$ and color-stripped $\mathcal{O}(\alpha'^2)$ trees $\mathcal{A}^{F^4}$ as well as low multiplicity

[1] The four-dimensional $N=4$ SYM theory can be obtained by standard dimensional reduction from its ten-dimensional version with $N=1$ supersymmetry [25].

JHEP08(2014)099

examples thereof. According to subsection 7.3, the representation (7.21) of the color-dressed $\mathcal{O}(\alpha'^2)$ tree manifests a duality to the one-loop kinematic factor (5.34).

2 Review of tree-level cohomology building blocks

In this section, we shall review the construction of tree-level amplitudes in ten-dimensional SYM, based on BRST building blocks in pure spinor superspace [8, 17, 18]. Although the problem at hand is of purely field theoretic nature, we shall use the vertex operators and the BRST charge of the pure spinor superstring [4] as the starting point. These ingredients suggest a pure spinor superspace representation for color-ordered tree subdiagrams with one off-shell leg. BRST invariance and the pole structure in the kinematic invariants

$$s_{12\ldots p} \equiv \frac{1}{2}(k_1 + k_2 + \cdots + k_p)^2 \tag{2.1}$$

turn out to be sufficient in order to determine the tree-level SYM amplitude $\mathcal{A}_n^{\rm YM}$ with any number n of external legs [6, 8]. The compactness of the final expression

$$\mathcal{A}^{\rm YM}(1,2,\ldots,n) = \sum_{j=1}^{n-2} \langle M_{12\ldots j}\, M_{j+1\ldots n-1}\, V^n \rangle \tag{2.2}$$

(see sections 2.1 and 2.3 for the definitions of V^n and $M_{12\ldots j}$, respectively) suggests to apply a similar program at loops, we will follow these lines in section 4 and introduce similar superspace variables.

At the level of the full-fledged superstring theory, the main virtue of the BRST building block representation for $\mathcal{A}^{\rm YM}$ is the possibility to identify these SYM constituents within the CFT computation of the superstring disk amplitude. The supersymmetric n-point tree amplitude in superstring theory was shown in [6, 7] to decompose into a sum of $(n-3)!$ color-ordered field theory amplitudes, each one of them being weighted by a separate function of α'. The main result of the current work is a similar decomposition of one-loop supersymmetric amplitudes, based on a new family of BRST building blocks.

2.1 From vertex operators to OPE residues

One of the major tasks in computing the open string tree-level amplitude is the evaluation of the CFT correlation function

$$\langle V^1(z_1)\, V^{n-1}(z_{n-1})\, V^n(z_n)\, U^2(z_2)\, \ldots\, U^{n-2}(z_{n-2}) \rangle \tag{2.3}$$

where V^1 and U^2 denote the vertex operators for the gluon multiplet with conformal-weight zero and one, respectively. They are conformal fields on the worldsheet parametrized by a complex coordinate z. The 8+8 physical degrees of freedom are described by the superfields[2] A_α, A^m, W^α and $\mathcal{F}_{mn}$ of $D = 10$ SYM [19]

$$V^1 = \lambda^\alpha A^1_\alpha, \qquad U^i = \partial\theta^\alpha A^i_\alpha + \Pi^m A^i_m + d_\alpha W^\alpha_i + \frac{1}{2}\mathcal{F}^i_{mn} N^{mn}, \tag{2.4}$$

[2] Throughout this work, SO(1, 9) vector indices are taken from the middle of Latin alphabet $m, n, p, \ldots = 0, 1, \ldots, 9$ whereas Weyl spinor indices $\alpha, \beta, \ldots = 1, 2, \ldots 16$ are taken from the beginning of the Greek alphabet.

JHEP08(2014)099

where λ^α denotes the pure spinor ghost subject to $(\lambda\gamma^m\lambda)=0$ [4]. The remaining ingredients $\partial\theta^\alpha, \Pi^m, d_\alpha$ and N^{mn} of (2.4) are conformal weight-one fields on the worldsheet. The ten-dimensional superfields A_α, A^m, W^α and $\mathcal{F}_{mn}$, depending on the bosonic and fermionic superspace variables x^m and θ_α, obey the following equations of motion [19, 47],

$$\begin{aligned} 2D_{(\alpha}A_{\beta)} &= \gamma^m_{\alpha\beta}A_m & D_\alpha A_m &= (\gamma_m W)_\alpha + k_m A_\alpha \\ D_\alpha \mathcal{F}_{mn} &= 2k_{[m}(\gamma_{n]}W)_\alpha & D_\alpha W^\beta &= \frac{1}{4}(\gamma^{mn})_\alpha{}^\beta \mathcal{F}_{mn}. \end{aligned} \tag{2.5}$$

As shown in [45, 46], their θ expansions can be computed in the gauge $\theta^\alpha A_\alpha = 0$ and read [44]

$$\begin{aligned} A_\alpha(x,\theta) &= \frac{1}{2}a_m(\gamma^m\theta)_\alpha - \frac{1}{3}(\xi\gamma_m\theta)(\gamma^m\theta)_\alpha - \frac{1}{32}F_{mn}(\gamma_p\theta)_\alpha(\theta\gamma^{mnp}\theta) + \cdots \\ A_m(x,\theta) &= a_m - (\xi\gamma_m\theta) - \frac{1}{8}(\theta\gamma_m\gamma^{pq}\theta)F_{pq} + \frac{1}{12}(\theta\gamma_m\gamma^{pq}\theta)(\partial_p\xi\gamma_q\theta) + \cdots \\ W^\alpha(x,\theta) &= \xi^\alpha - \frac{1}{4}(\gamma^{mn}\theta)^\alpha F_{mn} + \frac{1}{4}(\gamma^{mn}\theta)^\alpha(\partial_m\xi\gamma_n\theta) + \frac{1}{48}(\gamma^{mn}\theta)^\alpha(\theta\gamma_n\gamma^{pq}\theta)\partial_m F_{pq} + \cdots \\ \mathcal{F}_{mn}(x,\theta) &= F_{mn} - 2(\partial_{[m}\xi\gamma_{n]}\theta) + \frac{1}{4}(\theta\gamma_{[m}\gamma^{pq}\theta)\partial_{n]}F_{pq} + \frac{1}{6}\partial_{[m}(\theta\gamma_{n]}{}^{pq}\theta)(\xi\gamma_q\theta)\partial_p + \cdots \end{aligned} \tag{2.6}$$

where $a_m(x) = e_m e^{ik\cdot x}$, $\xi^\alpha(x) = \chi^\alpha e^{ik\cdot x}$ are the gluon and gluino polarizations and $F_{mn} = 2\partial_{[m}a_{n]}$ is the linearized field-strength.

The equations of motion (2.5) imply that the vertex operators in (2.4) obey $QV^i = 0$ and $QU^j = \partial V^j$. Since their ingredients V^i and $\int U^j$ are BRST closed, superstring amplitudes (and in particular their field theory limit) should inherit this property.

The correlator (2.3) can be computed by integrating out the conformal worldsheet fields of unit weight within the U^j vertex operator. This amounts to summing over all worldsheet singularities in $z_i \to z_j$ which the fields in question can produce. In any CFT, this information is carried by operator product expansions (OPEs), the first example being

$$V^1(z_1)\,U^2(z_2) \;\to\; \frac{L_{21}}{z_{21}}. \tag{2.7}$$

This defines a composite superfield L_{21} associated with the degrees of freedom of the states with labels 1 and 2, respectively. By iterating this OPE fusion, we define a family of superfields of arbitrary rank [8]

$$L_{21}(z_1)\,U^3(z_3) \;\to\; \frac{L_{2131}}{z_{31}}, \qquad L_{2131\ldots l1}(z_1)\,U^m(z_m) \;\to\; \frac{L_{2131\ldots l1m1}}{z_{m1}} \tag{2.8}$$

which will be referred to as OPE residues.[3] After the fields with conformal weight one have been integrated out using their OPEs, the zero modes of the pure spinor λ^α and θ^α are integrated using the $\langle(\lambda^3\theta^5)\rangle = 1$ prescription reviewed in [6].

[3] It turns out that even if OPE contractions are firstly carried out among $U^i(z_i)U^j(z_j)$ and then merged with V^1, the result is still a combination of $L_{2131\ldots m1}$ permutations. In other words, at tree-level the OPE $U^i(z_i)U^j(z_j)$ does not introduce any independent composite superfields.

JHEP08(2014)099

2.2 From OPE residues to BRST building blocks

A major shortcoming of the OPE residues $L_{2131\ldots m1}$ is their lack of symmetry under exchange of labels $1, 2, 3, \ldots, m$. However, the obstructions to well-defined symmetry properties can be shown to conspire to BRST-exact terms. As a simple example, the symmetric rank-two combination is

$$L_{21} \;+\; L_{12} = -\,Q(A_1 \cdot A_2) \tag{2.9}$$

where $Q = \lambda^\alpha D_\alpha$ denotes the BRST operator of the pure spinor formalism [4] and A_i^m is the vectorial superfield of $D = 10$ SYM. Using the BRST transformation properties of $L_{2131\ldots}$, these BRST-exact admixtures have been identified in [6, 8] up to rank five, and their removal leads to a redefinition of the OPE residues[4]

$$T_{12} \;\equiv\; L_{21} - \frac{1}{2}\,(L_{21} + L_{12}) = L_{[21]}, \qquad T_{123\ldots m} \;\equiv\; L_{2131\ldots m1} - \text{corrections}\,. \tag{2.10}$$

The outcome of (2.10) is an improved family of superfields $T_{123\ldots m}$ which we call BRST building blocks. They are covariant under the action of the BRST charge, e.g.

$$\begin{aligned}
Q\,T_1 &= 0\\
Q\,T_{12} &= s_{12}\,T_1\,T_2\\
Q\,T_{123} &= (s_{123} - s_{12})\,T_{12}\,T_3 - s_{12}\,(T_{23}\,T_1 + T_{31}\,T_2)\\
Q\,T_{1234} &= (s_{1234} - s_{123})\,T_{123}\,T_4 + (s_{123} - s_{12})\,(T_{12}\,T_{34} + T_{124}\,T_3)\\
&\qquad + s_{12}\,(T_{134}\,T_2 + T_{13}\,T_{24} + T_{14}\,T_{23} + T_1\,T_{234})\\
QT_{12\ldots k} &= \sum_{j=2}^{k} \sum_{\alpha\in P(\beta_j)} (s_{12\ldots j} - s_{12\ldots j-1}) T_{12\ldots j-1,\{\alpha\}} T_{j,\{\beta_j\setminus\alpha\}}
\end{aligned} \tag{2.11}$$

where $V_1 \equiv T_1$. The set $\beta_j = \{j+1, j+2, \ldots, k\}$ encompasses the $k-j$ labels to the right of j, and $P(\beta_j)$ denotes its power set. In other words, Q acting on a BRST building block of higher rank yields products of two lower rank analogues together with a Mandelstam variable.

As discussed in [6], at each rank the BRST building blocks obey one new symmetry in its labels while still respecting all the lower-rank symmetries. For example, since the rank-two building block satisfies $T_{(12)} = 0$ all higher-order building blocks also obey $T_{(12)34\ldots} = 0$. At rank-three there is one new symmetry $T_{[123]} = 0$ which is respected by all higher-order ranks, $T_{[123]4\ldots m} = 0$ and so forth. The generalization to rank $m \geq 3$ is given by [6],

$$\begin{aligned}
m = 2p+1: &\quad T_{12\ldots p+1[p+2[\ldots[2p-1[2p,2p+1]]\ldots]]} - 2T_{2p+1\ldots p+2[p+1[\ldots[3[21]]\ldots]]} = 0\\
m = 2p: &\quad\qquad T_{12\ldots p[p+1[\ldots[2p-2[2p-1,2p]]\ldots]]} + T_{2p\ldots p+1[p[\ldots[3[21]]\ldots]]} = 0,
\end{aligned} \tag{2.12}$$

and leaves $(m-1)!$ independent components at rank m. It turns out that the above symmetries are shared by color factors of nonabelian gauge theories formed by contracting structure constants f^{ijk} of the gauge group. At lowest ranks, we have

$$0 = f^{(12)3} \leftrightarrow T_{(12)} = 0, \qquad 0 = f^{[12|a} f^{3]4a} \leftrightarrow T_{[123]} = 0, \tag{2.13}$$

[4] We define (anti-)symmetrization of p indices to include $\frac{1}{p!}$, e.g. $L_{[21]} = \frac{1}{2}(L_{21} - L_{12})$.

JHEP08(2014)099

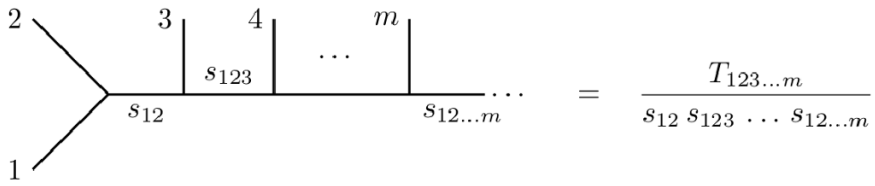

Figure 1. The correspondence of tree graphs with cubic vertices and BRST building blocks.

which states their total antisymmetry and Lie algebraic Jacobi identities, and similarly

$$0 = f^{12a} f^{a[3|b} f^{b|4]c} + f^{34a} f^{a[1|b} f^{b|2]c} \leftrightarrow T_{12[34]} + T_{34[12]} = 0. \tag{2.14}$$

In general, the symmetries of a rank m building block are the same as those of a string of structure constants with $m+1$ labels,

$$f^{12a_2}\, f^{a_2 3 a_3}\, f^{a_3 4 a_4} \ldots f^{a_{m-1} m a_m} \leftrightarrow\ T_{1234\ldots m}\,, \tag{2.15}$$

where the free color index a_m reflects an off-shell leg $m+1$ in the associated cubic diagram.

Therefore the basis of rank m building blocks being $(m-1)!$-dimensional is equivalent to the well-known fact that the basis of contractions of structure constants with p free adjoint indices has dimension $(p-2)!$ after Jacobi identities.

This similarity of building blocks with color factors as well as their BRST variations suggest a diagrammatic interpretation for $T_{123\ldots m}$ in terms of tree subdiagrams with cubic vertices [27] as seen on figure 1. Firstly, the color structure of this diagram is given by (2.15) via Feynman rules and secondly each propagator can be cancelled by one of the Mandelstam variables in the BRST variation $QT_{123\ldots m} \to s_{12}, s_{123}, s_{1234}, \ldots, s_{1234\ldots m}$. In other words, the role of the BRST operator is to cancel propagators.

2.3 From BRST building blocks to Berends-Giele currents

Given the dictionary between cubic tree subdiagrams and BRST building blocks, the next challenge is to combine different diagrams in order to arrive at BRST-invariant SYM amplitudes. The next hierarchy level of building blocks consists of superspace representations $M_{123\ldots m}$ of so-called Berends-Giele currents [43] which can be thought of as color-ordered SYM tree amplitudes with one leg off-shell. They encompass all the cubic diagrams present in the associated SYM tree and consist of kinematic numerators $T_{123\ldots m}$ dressed by their propagators $(s_{12}s_{123}\ldots s_{12\ldots m})^{-1}$, e.g.

$$M_{12} = \frac{T_{12}}{s_{12}}, \qquad M_{123} = \frac{T_{123}}{s_{12}s_{123}} + \frac{T_{321}}{s_{23}s_{123}} \tag{2.16}$$

corresponding to the three- and four-point amplitudes with one leg off-shell. At rank four,

$$M_{1234} = \frac{1}{s_{1234}} \left(\frac{T_{1234}}{s_{12}s_{123}} + \frac{T_{3214}}{s_{23}s_{123}} + \frac{T_{3421}}{s_{34}s_{234}} + \frac{T_{3241}}{s_{23}s_{234}} + \frac{2T_{12[34]}}{s_{12}s_{34}} \right) \tag{2.17}$$

JHEP08(2014)099

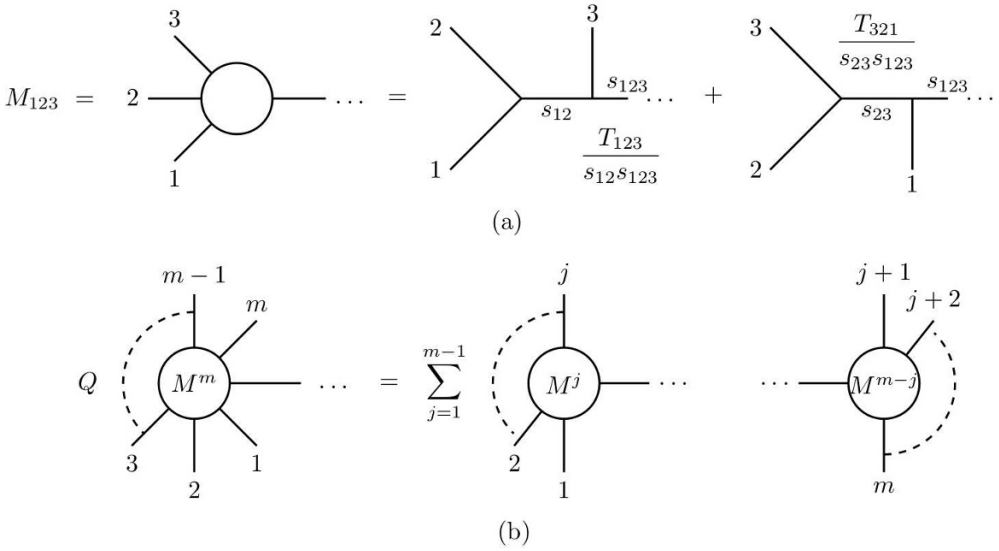

Figure 2. (a) The cubic graphs with one leg off shell which compose the rank three Berends-Giele current M_{123}. (b) The factorization of the current $M_{12\ldots m}$ under the action of the BRST charge. The right-hand side involves the sum over all partitions of m legs which is compatible with the color ordering set by $\{1, 2, \ldots, m\}$.

collects the five cubic diagrams of a color-ordered five-point amplitude. The two diagrams present in M_{123} are shown in figure 2.

The necessity to combine BRST building blocks to full-fledged Berends-Giele currents can be seen from their Q variation: their fine-tuned diagrammatic content makes sure that also the $M_{123\ldots m}$ are covariant under the BRST charge, i.e. (where for rank one, $M_1 \equiv V_1$)

$$\begin{aligned} Q\, M_1 &= 0 \\ Q\, M_{12} &= M_1\, M_2 \\ Q\, M_{123} &= M_{12}\, M_3 + M_1\, M_{23} \\ Q\, M_{1234} &= M_{123}\, M_4 + M_{12}\, M_{34} + M_1\, M_{234}\,. \end{aligned} \tag{2.18}$$

In contrast to $QT_{123\ldots m}$ as given by (2.11), there are no explicit Mandelstam variables in (2.18) because the rank m current already incorporates $m-1$ simultaneous poles. The generalization of (2.18) to higher rank,

$$Q\, M_{12\ldots m} = \sum_{j=1}^{m-1} M_{12\ldots j}\, M_{j+1\ldots m} \tag{2.19}$$

involves all partitions of the m on-shell legs on two Berends-Giele currents which are compatible with the color ordering. The situation is depicted in figure 2b.

2.4 The $D = 10$ SYM amplitude as a pure spinor cohomology problem

Using the Berends-Giele currents reviewed in the previous subsection, a method to recursively compute the ten-dimensional SYM tree-level scattering amplitudes was developed

JHEP08(2014)099

in [8]. It was later shown in [6] that the expressions found in [8] also follow from the field theory limit of tree-level superstring amplitudes computed with the pure spinor formalism.

The method relies on finding an expression in the cohomology of the pure spinor BRST charge, i.e. which is BRST-closed but non-exact,

$$Q\mathcal{A}^{\rm YM}(1,2,\ldots,n)=0, \qquad \mathcal{A}^{\rm YM}(1,2,\ldots,n)\neq\langle Q\mathcal{X}_n\rangle.$$

If we additionally require this cohomology element to reproduce the kinematic poles of a color-ordered SYM subamplitude, the result is uniquely determined to be

$$\mathcal{A}^{\rm YM}(1,2,\ldots,n)=\sum_{j=1}^{n-2}\langle M_{12\ldots j}\,M_{j+1\ldots n-1}\,V^n\rangle. \tag{2.20}$$

In order to show that the right-hand side is in the BRST cohomology first note that $QV_n=0$, whereas

$$Q\sum_{j=1}^{n-2}M_{12\ldots j}\,M_{j+1\ldots n-1}=0 \tag{2.21}$$

follows from (2.19). And secondly, in the momentum phase space of n massless particles where the Mandelstam variable $s_{12\ldots n-1}$ vanishes, $\sum_{j=1}^{n-1}M_{12\ldots j}\,M_{j+1\ldots n-1}$ can not be written as $QM_{12\ldots n-1}$ since $M_{12\ldots n-1}$ contains a divergent propagator $1/s_{12\ldots n-1}$. This rules out BRST-exactness of (2.20).

The number of cubic diagrams in the color-ordered n-point tree amplitude is given by the Catalan number C_{n-2}, see [62], which satisfies the recurrence relation $C_{p+1}=\sum_{i=0}^{p}C_iC_{p-i}$ with $C_0=1$. By its diagrammatic construction, $M_{12\ldots j}$ gathers C_{j-1} pole channels, so the number of poles in the expression (2.20) for the n-point subamplitude is given by $\sum_{i=0}^{n-3}C_iC_{n-3-i}$, which is precisely the recursive definition of C_{n-2}. The expression (2.20) therefore contains the same number of cubic diagrams as the color-ordered n-point amplitude, and the fact that Berends-Giele currents have a notion of color ordering guarantees that the pole channels in (2.20) are precisely those of $\mathcal{A}^{\rm YM}(1,2,\ldots,n)$. The factorization properties of the expression (2.20) are depicted in figure 3, and the reader is referred to [6] for more details.

3 One-loop amplitudes with the minimal pure spinor formalism

This section sketches the prescription towards one-loop amplitudes within the minimal pure spinor formalism. The main goal is to make the one-loop zero mode saturation rule (3.8) for the correlator $\langle V^1\prod_{j=2}^{n}U^j\rangle$ plausible instead of giving an exhaustive review. The reader is referred to [5] for the details omitted in the following discussion.

The prescription to compute n-point one-loop amplitudes for open superstrings is [5]

$$\mathcal{A}_n^{1-\rm loop}=\sum_{\rm top}C_{\rm top}\int_{D_{\rm top}}dt\,\langle(\mu,b)\prod_{P=2}^{10}Z_{B_P}Z_J\prod_{I=1}^{11}Y_{C_I}\,V_1(z_1)\prod_{j=2}^{n}\int dz_jU^j(z_j)\rangle, \tag{3.1}$$

where μ is the Beltrami differential, t is the Teichmüller parameter and b is the b-ghost whose contribution will be discussed below. The sum runs over all one-loop open string

JHEP08(2014)099

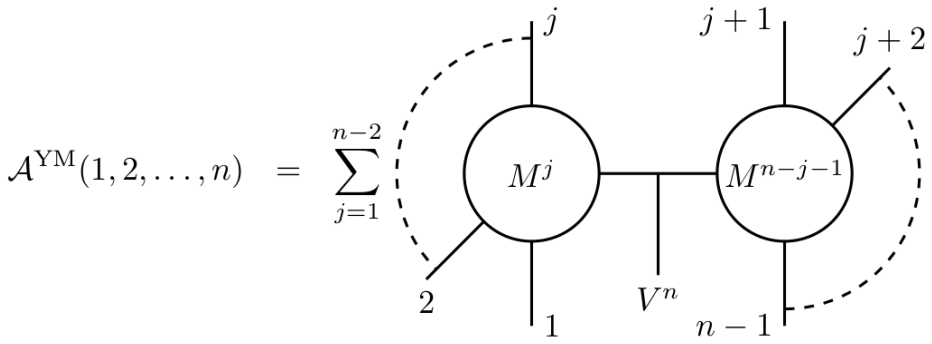

Figure 3. Diagrammatic interpretation of the expression $\sum_{j=1}^{n-2}\langle M_{12\ldots j}M_{j+1\ldots n-1}V^n\rangle$ for the n-point SYM tree amplitudes. The j sum runs over all partitions of the first $n-1$ legs among two Berends-Giele currents.

worldsheet topologies, i.e. over planar and non-planar cylinder diagrams as well as the Moebius strip, see [13]. The integration domain $D_{\rm top}$ for t has to be adjusted accordingly, and the associated color factors $C_{\rm top}$ are single- or double traces over Chan-Paton generators associated with the external states. Both the Chan-Paton traces and the integration region for the z_j must reflect the cyclic ordering of the vertex operators on the boundaries of the genus one worldsheet. The main focus of this work is the simplification of the t integrand in (3.1), so we do not specify further details of $D_{\rm top}$ and $C_{\rm top}$ or comment on the interplay between the topologies.

In order to introduce the remaining elements appearing in (3.1), note that the computation of the CFT correlator at one-loop starts by separating off the zero mode of the conformal weight one variables. The role of the picture-changing operators Z_B, Z_J and Y_C is to ensure that the zero modes of bosonic and fermionic variables are absorbed correctly, see [5]. The angle brackets $\langle\ldots\rangle$ in (3.1) initially denote the path integral over all the worldsheet variables in the pure spinor formalism. The non-zero modes are integrated out using their OPEs as described below and we will follow a procedure where the zero modes of d_α, N^{mn} and the ghost current J are integrated out first, leaving those of λ^α and θ^α for a last step in the computation, e.g. after the superfield expansions of (2.6) are substituted in the expressions of various building blocks. And since general group theory arguments will be used to determine the integrals over zero modes of d_α, N^{mn} and J the precise details of the zero-mode measures of [5] will not be needed.

So unless otherwise stated, every appearance of the pure spinor angle brackets $\langle\ldots\rangle$ in this paper denotes the zero-mode integration of λ^α and θ^α only and will be taken as the definition of pure spinor superspace [37]. This integration can be performed using symmetry arguments alone and follows from the tree-level prescription $\langle(\lambda^3\theta^5)\rangle = 1$ of [4]. Since this tedious process has been mostly automated in [14] we will restrict ourselves to presenting our one-loop results in compact pure spinor superspace form as in the tree-level approach of [6]. Furthermore, the correlation function of the matter variables $x^m(z,\overline{z})$ and $\Pi^m(z)$ is computed as in [10, 20] and will receive no special treatment in the following.

JHEP08(2014)099

The non-zero-modes are integrated out using their OPEs [10]

$$d_\alpha(z_i)\theta^\beta(z_j) \to \eta_{ij}\delta^\beta_\alpha, \qquad \Pi_m(z_i)x^n(z_j,\overline{z}_j) \to -\eta_{ij}\delta^n_m. \tag{3.2}$$

Singularities in colliding worldsheet positions enter through the function η_{ij} which is defined on a given Riemann surface as the derivative of the bosonic Green's function

$$\eta_{ij} := \frac{\partial}{\partial z_i}\,\langle x(z_i,\overline{z}_i)\,x(z_j,\overline{z}_j)\rangle\ .$$

It behaves as z_{ij}^{-1} as the positions approach each other but respects the periodicity properties required by a higher-genus Riemann surface. The representation in terms of Jacobi theta functions will not be needed in the following discussions, only its antisymmetry $\eta_{ij} = -\eta_{ji}$ will play a fundamental role.

In the amplitude prescription (3.1), the b-ghost is a composite operator whose form is given schematically by [5, 30],

$$\begin{aligned} b = {} & (\Pi d + N\partial\theta + J\partial\theta)\,d\,\delta(N) + (w\partial\lambda + J\partial N + N\partial J + N\partial N)\delta(N) \\ & + (N\Pi + J\Pi + \partial\Pi + d^2)(\Pi\delta(N) + d^2\delta'(N)) \\ & + (Nd + Jd)(\partial\theta\delta(N) + d\Pi\delta'(N) + d^3\delta''(N)) \\ & + (N^2 + JN + J^2)(d\partial\theta\delta'(N) + \Pi^2\delta'(N) + \Pi d^2\delta''(N) + d^4\delta'''(N)) \end{aligned} \tag{3.3}$$

where $\delta'(x) = \frac{\partial}{\partial x}\delta(x)$ is defined through integration by parts and the precise index contractions are being omitted. It will be argued in the appendix A that the zero-mode contribution from the b-ghost is unique and given by an expression of the form $d^4\delta'(N)$. Furthermore, the result of the zero-mode integrations in this case is fixed by group theory up to an overall constant, and this is the contribution which will concern us in this paper.

We do not have a constructive proof that the b-ghost does not contribute via OPE contractions (i.e. via nonzero modes), but an indirect argument based on total symmetry of the kinematic factor will follow in subsection 7.4.

In general, the evaluation of the one-loop amplitude (3.1) involves two separate challenges summarized by the formula[5]

$$\mathcal{A}_n^{1-\text{loop}} = \sum_{\text{top}} C_{\text{top}} \int_{D_{\text{top}}} dt \prod_{j=2}^{n} \int dz_j \left\langle \prod_{i=1}^{n} e^{ik_i\cdot x(z_i,\overline{z}_i)} \right\rangle \times K_n\ .$$

Firstly, the computation of the kinematic factor K_n in pure spinor superspace whose generic form is given by

$$K_n = \eta^{n-4}\langle f(\lambda^\alpha,\theta^\alpha;1,2,\ldots,n)\rangle \tag{3.4}$$

(where $1,2,\ldots,n$ denote the physical degrees of freedom of the n external states), and secondly, the evaluation of the integrals over vertex operator positions on the boundary of

[5]Since the Koba Nielsen factor KN= $\left\langle \prod_{i=1}^{n} e^{ik_i\cdot x(z_i,\overline{z}_i)} \right\rangle$ due to the plane wave correlator is a universal prefactor, we define the kinematic factor K_n not to contain KN. Nevertheless, its presence is relevant for integration by parts relating different worldsheet functions, see subsection 5.3.

JHEP08(2014)099

the Riemann surface as well as the modular parameter t. The form of the kinematic part is unique to the pure spinor formalism and will be dealt with in the following sections. Its BRST invariant ingredients will be identified and related to the α'^2 terms in the expansion of the corresponding tree-level amplitudes. The expressions for the integrals over the Riemann surface are exactly like in RNS or Green-Schwarz [11, 12] formalisms and will not play a role in this article. Extracting information on the integrals — in particular their field theory limits — will be left for future work.

3.1 The one-loop prescription for d_α, N^{mn} zero mode saturation

When the number of external states is four, the saturation of d_α zero modes in

$$\mathcal{A}_4^{1-\text{loop}} = \sum_{\text{top}} C_{\text{top}} \int_{D_{\text{top}}} dt \langle (\mu, b) \prod_{P=2}^{10} Z_{B_P} Z_J \prod_{I=1}^{11} Y_{C_I} V_1(z_1) \prod_{j=2}^{4} \int dz_j U^j(z_j) \rangle, \tag{3.5}$$

is unique and determines the amplitude up to an overall coefficient [5, 63]. The picture changing operators, the b-ghost and the external vertices provide ten, four and two d_α zero-modes, respectively, thereby saturating all the sixteen zero modes of d_α. Furthermore, as mentioned after (3.3), the terms with four d_α zero modes from the b-ghost also contain factors which absorb extra zero modes of N^{mn}, either 1, 2 or 3. For the four-point amplitude the only possibility is the absorption of one zero mode of N^{mn} through an overall factor of $\delta'(N)$. Summing it all up, the contribution from the external vertices is proportional to

$$\frac{1}{2} V_1 (dW^2)(dW^3) \mathcal{F}^4_{mn} N^{mn} + \text{cyclic}(234) \tag{3.6}$$

and the remaining zero mode integration is given schematically by

$$\begin{aligned} K_4 = \int [\mathcal{D}\lambda][\mathcal{D}N]\, d^{16}\theta\, d^{16} d\, (\lambda)^{10}\, (d)^{14}\, (\theta)^{11}\, \delta^{11}(\lambda)\, \delta^{10}(N)\, \delta(J)\, \delta'(N) \\ \times\ \frac{1}{2} V_1 (dW^2)(dW^3) \mathcal{F}^4_{mn} N^{mn} + \text{cyc}(234). \end{aligned} \tag{3.7}$$

As one can check in the expressions given in [5], the measure factor $[\mathcal{D}N]$ has ghost-number -8. Therefore the integration of $\int [\mathcal{D}N]\, d^{16} d (\lambda)^{10} (d)^{14} \delta^{10}(N) \delta(J) \delta'(N)$ in (3.7) with ten powers of λ has the net effect of replacing $d_\alpha d_\beta N^{mn}$ from the external vertices by a λ bilinear. The tensor structure is uniquely determined by group theory since the decomposition of $d_\alpha \otimes d_\beta \otimes N^{mn}$ contains only one component in the SO(10) representation (00002) of a chiral pure spinor bilinear:

$$d_\alpha d_\beta N^{mn} \longrightarrow (\lambda \gamma^{[m})_\alpha (\lambda \gamma^{n]})_\beta \tag{3.8}$$

Consequently, (3.6) leads to the following kinematic factor for the four-point one-loop amplitude

$$K_4 = \frac{1}{2} \langle V_1 (\lambda \gamma_m W_2)(\lambda \gamma_n W_3) \mathcal{F}_4^{mn} \rangle + \text{cyclic}(234) \tag{3.9}$$

whose BRST invariance one can easily check using the pure spinor constraint $(\lambda \gamma^m \lambda) = 0$ and elementary corollaries $(\lambda \gamma_m)_\alpha (\lambda \gamma^m)_\beta = 0$ and $(\lambda \gamma^m \gamma_{pq} \lambda) = 0$.

JHEP08(2014)099

According to the arguments in appendix A, the replacement rule (3.8) still applies to one-loop amplitudes with $n \geq 5$ legs. It passes the superspace kinematic factor built from one unintegrated and $n-1$ integrated vertex operators to the tree-level zero mode prescription $\langle \lambda^3 \theta^5 \rangle = 1$:

$$K_n \equiv \langle V^1(z_1) U^2(z_2) U^3(z_3) \dots U^n(z_n) \rangle_{d_\alpha d_\beta N^{mn} \longrightarrow (\lambda \gamma^{[m]})_\alpha (\lambda \gamma^{n]})_\beta}$$

Studying the interplay of (3.8) with the non-zero modes of the conformal fields in U^j is the subject of the next section. Integrating out all but three weight one fields $d_\alpha d_\beta N^{mn}$ obviously requires $n-4$ OPEs, and we will see that they give rise to new families of BRST building blocks.

4 BRST building blocks for loop amplitudes

As reviewed in section 2, tree-level BRST building blocks $T_{12\dots k}$ are defined by a two step procedure. Its starting point have been the residues of the single poles in iterated OPEs of integrated vertex operators $U(z_j)$ with the unintegrated one $V(z_1)$. As a second step, the BRST trivial components of these residues had to be subtracted to obtain symmetry properties suitable for a diagrammatic interpretation. On the genus zero worldsheet governing tree-level amplitudes, conformal fields of weight $+1$ have no zero modes, so all of d_α and N^{mn} are completely integrated out in generating the residues entering BRST building blocks. However, this is no longer the case at one-loop.

As seen in the previous section, the kinematic factor at one-loop comes from the terms in the external vertices which contain two zero modes of d_α and one of N^{mn}. Hence, we have to integrate out weight one fields from the $n-1$ integrated vertex operators until we are left with the combination $(d)^2 N$ which requires a total of $n-4$ OPE contractions. In doing so, one is naturally led to define the composite superfields $\tilde{J}^{mn}_{12}, \tilde{K}^\alpha_{12}$ and higher rank generalizations $\tilde{J}^{mn}_{12\dots k}, \tilde{K}^\alpha_{12\dots k}$ as the remaining single-pole terms $\sim d_\alpha$ or $\sim N_{mn}$ in nested OPEs of multiple integrated vertex operators:

$$\begin{aligned} U_1(z_1) U_2(z_2) &\longrightarrow \frac{\tilde{J}^{mn}_{12} N_{mn}}{z_{21}} + \frac{d_\alpha \tilde{K}^\alpha_{12}}{z_{21}} + \cdots \\ U_1(z_1) U_2(z_2) \dots U_k(z_k) &\longrightarrow \frac{\tilde{J}^{mn}_{12\dots k} N_{mn}}{z_{k,k-1} \dots z_{32} z_{21}} + \frac{d_\alpha \tilde{K}^\alpha_{12\dots k}}{z_{k,k-1} \dots z_{32} z_{21}} + \cdots \end{aligned} \tag{4.1}$$

The ellipsis $\cdots$ indicates terms with Π^m and $\partial\theta^\alpha$ as well as double poles in individual z_{ij}, they do not contribute to the end result for one-loop amplitudes. Given the prescription $d_\alpha d_\beta N^{mn} \mapsto (\lambda \gamma^{[m]})_\alpha (\lambda \gamma^{n]})_\beta$, the quantity of interest built from the $\tilde{K}^\alpha$ superfield is

$$\tilde{K}^m_{12\dots k} \equiv (\lambda \gamma^m)_\alpha \tilde{K}^\alpha_{12\dots k}. \tag{4.2}$$

As a rank $k=2$ example, let us consider the OPE of two integrated vertices. It contains

JHEP08(2014)099

single and double poles

$$U_1(z_1)U_2(z_2) \longrightarrow \frac{1}{z_{21}}\Big[(k_2\cdot A_1)U^2+\frac{1}{2}(W_1\gamma^m W_2)\Pi^m+(k_1\cdot\Pi)(A_1W_2)+\partial\theta^\alpha D_\alpha A_\beta^1 W_2^\beta \\ +\frac{1}{4}(d\gamma^{mn}W_2)\mathcal{F}^1_{mn}+k_m^1(W_1\gamma_n W_2)N^{mn}-\frac{1}{2}\mathcal{F}^1_{mp}\mathcal{F}_n^{2p}N^{mn}-(1\leftrightarrow 2)\Big] \\ +\frac{1+(k_1\cdot k_2)}{z_{21}^2}\Big[(A_1W_2)+(A_2W_1)-(A_1\cdot A_2)\Big] \tag{4.3}$$

with $U^2=\partial\theta^\alpha A^2_\alpha+\Pi^m A^2_m+d_\alpha W_2^\alpha+\frac{1}{2}N_{mn}\mathcal{F}_2^{mn}$, and one can read off

$$\tilde{K}^m_{12}=\frac{1}{4}(\lambda\gamma^m\gamma^{pq}W_2)\mathcal{F}^1_{pq}+(k_2\cdot A_1)(\lambda\gamma^m W_2)\ -\ (1\leftrightarrow 2) \tag{4.4}$$

$$\tilde{J}^{mn}_{12}=\frac{1}{2}\big[(k_2\cdot A_1)\,\mathcal{F}_2^{mn}+\mathcal{F}_{2\,p}^{[m}\,\mathcal{F}_1^{n]p}+k_{12}^{[m}\,(W_1\gamma^{n]}W_2)\big]\ -\ (1\leftrightarrow 2) \tag{4.5}$$

from the superfields contracted with d_α and N_{mn}, respectively.

The definitions in (4.1) lead to the following rank ≤ 3 expressions

$$\tilde{K}^m_1=(\lambda\gamma^m W_1) \tag{4.6}$$

$$\tilde{K}^m_{12}=\frac{1}{4}(\lambda\gamma^m\gamma^{pq}W_2)\mathcal{F}^1_{pq}+(k_2\cdot A_1)(\lambda\gamma^m W_2)\ -\ (1\leftrightarrow 2) \tag{4.7}$$

$$\begin{aligned}\tilde{K}^m_{123}=&-\frac{1}{2}(k_{12}\cdot A_3)\,\tilde{K}^m_{12}-(\lambda\gamma^m W_1)\,k_p^1\,(W_2\gamma^p W_3)+\frac{1}{2}(\lambda\gamma^m W_3)\,k_p^3\,(W_1\gamma^p W_2)\\ &+\frac{1}{4}(k_1\cdot A_2)\Big[(\lambda\gamma^m\gamma^{pq}W_1)\,\mathcal{F}^3_{pq}-(\lambda\gamma^m\gamma^{pq}W_3)\,\mathcal{F}^1_{pq}-4(k_3\cdot A_1)\,(\lambda\gamma^m W_3)\Big]\\ &+\frac{1}{2}(\lambda\gamma^m\gamma^{pq}W_2)\,k_p^1\,(W_1\gamma_q W_3)-\frac{1}{4}(\lambda\gamma^m\gamma^{pq}W_3)\,\mathcal{F}_p^{1\,r}\,\mathcal{F}^2_{qr}\\ &+\frac{1}{16}\,(\lambda\gamma^m\gamma^{pq}\gamma^{rs}W_1)\,\mathcal{F}^3_{pq}\,\mathcal{F}^2_{rs}\ -\ (1\leftrightarrow 2)\end{aligned} \tag{4.8}$$

$$\tilde{J}^{mn}_1=\frac{1}{2}\,\mathcal{F}_1^{mn} \tag{4.9}$$

$$\tilde{J}^{mn}_{12}=\frac{1}{2}\big[\ (k_2\cdot A_1)\,\mathcal{F}_2^{mn}+\mathcal{F}_{2\,p}^{[m}\,\mathcal{F}_1^{n]p}+k_{12}^{[m}\,(W_1\gamma^{n]}W_2)\big]\ -\ (1\leftrightarrow 2) \tag{4.10}$$

$$\begin{aligned}\tilde{J}^{mn}_{123}=&-\frac{1}{2}\big[(k_{12}\cdot A_3)\,\tilde{J}^{mn}_{12}-(k_1\cdot A_2)(k_3\cdot A_1)\,\mathcal{F}_3^{mn}\big]+\mathcal{F}_1^{[m}{}_p\,\mathcal{F}_3{}_q^{\,n]}\,\mathcal{F}_2^{pq}+(k_1\cdot A_2)\,\mathcal{F}_{1\,r}^{[m}\,\mathcal{F}_3^{n]r}\\ &+k_p^1\,\mathcal{F}_2^{p[m}\,(W_1\gamma^{n]}W_3)-k_1^{[m}\,\mathcal{F}_2^{n]p}\,(W_1\gamma_p W_3)+(k_2\cdot A_1)\,k_{23}^{[m}\,(W_2\gamma^{n]}W_3)\\ &+\frac{1}{2}\big[(W_1\gamma^{[m}W_2)\,\mathcal{F}_3^{n]p}\,k_p^{12}-k_{12}^{[m}\mathcal{F}_3^{n]p}\,(W_1\gamma_p W_2)+k_2^p\,(W_1\gamma_p W_3)\,\mathcal{F}_2^{mn}\big]\\ &+\frac{1}{4}\big[(k_{12}^{[m}(W_1\gamma^{n]}\gamma^{pq}W_3)\,\mathcal{F}^2_{pq}+(W_2\gamma^{pq}\gamma^{[m}W_3)\,\mathcal{F}^1_{pq}+k_3^p\,(W_1\gamma_p W_2)\,\mathcal{F}_3^{mn}\big]-(1\leftrightarrow 2)\end{aligned} \tag{4.11}$$

where $k^m_{ij}:=k^m_i+k^m_j$. Expressions for the rank four building blocks $\tilde{K}^m_{1234}$ and $\tilde{J}^{mn}_{1234}$ are available from the authors upon request.

Similar to their tree-level counterparts $T_{12\ldots k}$ [6], the new composite superfields have two essential virtues: on the one hand, they have symmetry properties which reduce the independent rank k components to $(k-1)!$ and thereby suggest an interpretation in terms

JHEP08(2014)099

of tree-level subdiagrams with one off-shell leg. On the other hand, they possess *covariant* BRST variations,

$$\begin{aligned} Q\tilde{K}_1^m &= 0, \\ Q\tilde{K}_{12}^m &= s_{12}\left(T_1\tilde{K}_2^m - T_2\tilde{K}_1^m\right), \\ Q\tilde{K}_{123}^m &= s_{13}\,L_{21}\,\tilde{K}_3^m \;-\; s_{23}\,L_{12}\,\tilde{K}_3^m \;+\; s_{12}\left[\,L_{31}\,\tilde{K}_2^m \;-\; L_{32}\,\tilde{K}_1^m\,\right] \\ &\quad -\; (s_{13}+s_{23})\,T_3\,\tilde{K}_{12}^m \;+\; s_{12}\left[\,T_1\,\tilde{K}_{23}^m \;-\; T_2\,\tilde{K}_{13}^m\,\right]. \end{aligned} \tag{4.12}$$

However, the appearance of the OPE residue L_{21} in the right-hand side of $Q\tilde{K}_{123}^m$ instead of the BRST building block T_{12} signals the need for a redefinition of $\tilde{K}_{123}^m$ analogous to the redefinitions of $L_{2131\ldots}$ to $\tilde{T}_{123\ldots}$ at tree-level, see subsection 2.2. In order to justify this, let us recall the following general lesson from the tree-level analysis: quantities whose Q variation contains BRST exact constituents such as $L_{(21)} = -\frac{1}{2}Q(A_1\cdot A_2)$ combine to BRST trivial parts of the amplitude. It is economic to remove these terms in an early step of the computation, i.e. to study the BRST building block

$$K_{123}^m \equiv \tilde{K}_{123}^m \;+\; \frac{1}{2}\left[\,(s_{13}-s_{23})\,(A_1\cdot A_2)\,K_3^m \;+\; s_{12}\,((A_1\cdot A_3)\,K_2^m \;-\; (A_2\cdot A_3)\,K_1^m)\,\right] \tag{4.13}$$

from now on whose BRST transformation gives rise to T_{12} rather than L_{21}:

$$QK_{123}^m = (s_{13}+s_{23})(T_{12}\,K_3^m - V_3\,K_{12}^m) + s_{12}\big[T_{13}\,K_2^m - V_2\,K_{13}^m - T_{23}\,K_1^m + V_1\,K_{23}^m\big] \tag{4.14}$$

Also the higher rank cases $K_{12\ldots k}^m = \tilde{K}_{12\ldots k}^m + \ldots$ and $J_{12\ldots k}^{mn} = \tilde{J}_{12\ldots k}^{mn} + \ldots$ at $k \geq 4$ require modification to ensure BRST building blocks $T_{12\ldots k}$ rather than the OPE residues $L_{21\ldots k1}$ (with BRST exact components) in their Q transformation. However, in contrast to the tree-level redefinitions $T_{12\ldots k} = L_{21\ldots k1} + \ldots$, the symmetry properties of loop-specific building blocks are already present in OPE residues $\tilde{K}^m$ and $\tilde{J}^{mn}$. For instance, we already have an antisymmetric residue $\tilde{K}_{12}^m = \tilde{K}_{[12]}^m$ at rank two whereas the OPE residue L_{21} has to be projected on its antisymmetric part $T_{12} = L_{21} - L_{(21)}$.

Rank three is the first instance where modifications $Q\tilde{K}_{123}^m = s_{13}L_{21}\tilde{K}_3^m + \ldots$ are necessary to avoid BRST trivial admixtures $L_{(21)} = -\frac{1}{2}Q(A_1\cdot A_2)$ in the Q variation and to instead arrive at $QK_{123}^m = s_{13}T_{12}\tilde{K}_3^m + \ldots$ with $L_{21} \mapsto T_{12}$. Hence, the loop-specific OPE residues $\tilde{K}_{12\ldots k}^m$ are more closely related to their BRST building blocks $K_{12\ldots k}^m$ than the tree-level cousins $L_{21\ldots k1} \leftrightarrow T_{12\ldots k}$.

The BRST variations of OPE residues $\tilde{J}_{12\ldots k}^{mn}$ associated with N^{mn} lead to similar conclusions. Redefinitions $\tilde{J}_{12\ldots k}^{mn} \to J_{12\ldots k}^{mn}$ are needed in order to trade $L_{ji\ldots}$ and $\tilde{J}_{ij\ldots}^{mn}$ present in $Q\tilde{J}_{12\ldots k}^{mn}$ for $T_{ij\ldots}$ and $J_{ij\ldots}^{mn}$ in $QJ_{12\ldots k}^{mn}$. However, when computing their BRST variations one must take into account that the building blocks $J_{12\ldots k}^{mn}$ (or $\tilde{J}_{12\ldots k}^{mn}$) always appear contracted with $(\lambda\gamma^m)_\alpha(\lambda\gamma^n)_\beta$ because of the rule (3.8). So even though one might naively conclude $Q\tilde{J}_1^{mn} = k_1^{[m}(\lambda\gamma^{n]}W_1) \neq 0$, the effective contribution of its BRST variation to an amplitude is $(\lambda\gamma_m)_\alpha(\lambda\gamma_n)_\beta Q\tilde{J}_1^{mn} = 0$. For any $Q\tilde{J}_\ldots^{mn}$ or $QJ_\ldots^{mn}$ displayed in the following, terms that vanish under contraction with $K_\ldots^m K_\ldots^n \sim (\lambda\gamma^m)_\alpha(\lambda\gamma^n)_\beta$ are omitted.

In summary, the Q variations of the BRST building blocks which will appear in loop amplitudes are given by (2.11) and

$$QK_{12}^m = s_{12}\big(T_1K_2^m - T_2K_1^m\big) \tag{4.15}$$

JHEP08(2014)099

$$
\begin{aligned}
QK^m_{123} &= (s_{123}-s_{12})\big(T_{12}K^m_3 - T_3K^m_{12}\big)\\
&\quad + s_{12}\big(T_1K^m_{23} + T_{13}K^m_2 - T_{23}K^m_1 - T_2K^m_{13}\big) \qquad (4.16)\\
QK^m_{1234} &= (s_{1234}-s_{123})\big(T_{123}K^m_4 - T_4K^m_{123}\big)\\
&\quad + (s_{123}-s_{12})\big(T_{12}K^m_{34} + T_{124}K^m_3 - T_{34}K^m_{12} - T_3K^m_{124}\big)\\
&\quad + s_{12}\big(T_{134}K^m_2 + T_{13}K^m_{24} + T_{14}K^m_{23} + T_1K^m_{234} - T_2K^m_{134}\\
&\qquad - T_{24}K^m_{13} - T_{23}K^m_{14} - T_{234}K^m_1\big) \qquad (4.17)\\
QJ^{mn}_{12} &= s_{12}\big(T_1J^{mn}_2 - J^{mn}_1T_2\big) \qquad (4.18)\\
QJ^{mn}_{123} &= (s_{123}-s_{12})\big(T_{12}J^{mn}_3 - J^{mn}_{12}T_3\big)\\
&\quad + s_{12}\big(T_1J^{mn}_{23} + T_{13}J^{mn}_2 - J^{mn}_1T_{23} - J^{mn}_{13}T_2\big) \qquad (4.19)\\
QJ^{mn}_{1234} &= (s_{1234}-s_{123})\big(T_{123}J^{mn}_4 - J^{mn}_{123}T_4\big)\\
&\quad + (s_{123}-s_{12})\big(T_{12}J^{mn}_{34} + T_{124}J^{mn}_3 - J^{mn}_{12}T_{34} - J^{mn}_{124}T_3\big)\\
&\quad + s_{12}\big(T_{134}J^{mn}_2 + T_{13}J^{mn}_{24} + T_{14}J^{mn}_{23} + T_1J^{mn}_{234}\\
&\qquad - J^{mn}_{134}T_2 - J^{mn}_{13}T_{24} - J^{mn}_{14}T_{23} - J^{mn}_1T_{234}\big). \qquad (4.20)
\end{aligned}
$$

The BRST variations $QK^m_{12\ldots k}$ and $QJ^{mn}_{12\ldots k}$ of the new families can be obtained from $QT_{12\ldots k}$ by replacing either the first or the second $T_{ij\ldots}$ on the right hand side by the corresponding $K^m_{ij\ldots}$ or $J^{mn}_{ij\ldots}$. This doubles the number of terms in $QK^m_{12\ldots k}$ and $QJ^{mn}_{12\ldots k}$ compared to $QT_{12\ldots k}$, and the two ways of replacing a $T_{\ldots}$ in the BRST variation by $K^m_{\ldots}$ or $J^{mn}_{\ldots}$ enter with a relative minus sign (where the tree-level building block $T_{\ldots}$ is always understood to be placed on the left of $K^m_{\ldots}$ and $J^{mn}_{\ldots}$).

The above variations generalizes as follows to rank k:

$$
\begin{aligned}
QT_{12\ldots k} &= \sum_{j=2}^{k}\sum_{\alpha\in P(\beta_j)}(s_{12\ldots j}-s_{12\ldots j-1})T_{12\ldots j-1,\{\alpha\}}T_{j,\{\beta_j\setminus\alpha\}}\\
QK^m_{12\ldots k} &= \sum_{j=2}^{k}\sum_{\alpha\in P(\beta_j)}(s_{12\ldots j}-s_{12\ldots j-1})\big(T_{12\ldots j-1,\{\alpha\}}K^m_{j,\{\beta_j\setminus\alpha\}} - T_{j,\{\beta_j\setminus\alpha\}}K^m_{12\ldots j-1,\{\alpha\}}\big)\\
QJ^{mn}_{12\ldots k} &= \sum_{j=2}^{k}\sum_{\alpha\in P(\beta_j)}(s_{12\ldots j}-s_{12\ldots j-1})\big(T_{12\ldots j-1,\{\alpha\}}J^{mn}_{j,\{\beta_j\setminus\alpha\}} - T_{j,\{\beta_j\setminus\alpha\}}J^{mn}_{12\ldots j-1,\{\alpha\}}\big)
\end{aligned}
\qquad (4.21)
$$

where $V_i \equiv T_i$. The set $\beta_j = \{j+1, j+2, \ldots, k\}$ encompasses the $k-j$ labels to the right of j, and $P(\beta_j)$ denotes its power set.

4.1 Unified notation for one-loop BRST building blocks

For each contraction pattern among integrated vertex OPEs, there are three kinematic factors associated with the same $z_i \to z_j$ singularity structure. This corresponds to the three ways of extracting the worldsheet fields $d_\alpha d_\beta N^{mn}$ from three nested U^j OPEs a la (4.1). In other words, we have to sum three different possibilities (d,d,N), (d,N,d) and (N,d,d) to convert the U^j vertices after $n-4$ OPE fusions into building blocks $K^mK^nJ_{mn}$

JHEP08(2014)099

via $d_\alpha d_\beta N^{mn} \mapsto (\lambda\gamma^m)_\alpha(\lambda\gamma^n)_\beta$

$$T^i_{a_1\ldots a_p}\,T^j_{b_1\ldots b_q}\,T^k_{c_1\ldots c_r} \equiv \frac{2}{3}\big(K^m_{a_1\ldots a_p}\,K^n_{b_1\ldots b_q}\,J^{mn}_{c_1\ldots c_r} + K^m_{a_1\ldots a_p}\,J^{mn}_{b_1\ldots b_q}\,K^n_{c_1\ldots c_r} + J^{mn}_{a_1\ldots a_p}\,K^m_{b_1\ldots b_q}\,K^n_{c_1\ldots c_r}\big)\,. \tag{4.22}$$

Note that (4.22) is completely symmetric in i,j,k and under moving the T^i, T^j and T^k (which represent either $K^m_{\ldots}$ or $J^{mn}_{\ldots}$) across each other, i.e. $T^i_{a_1\ldots a_p}T^j_{b_1\ldots b_q} = T^j_{b_1\ldots b_q}T^i_{a_1\ldots a_p}$. As can be seen from the $K^m_{\ldots}K^n_{\ldots} \sim (\lambda\gamma^m)_\alpha(\lambda\gamma^n)_\beta$ in the definition (4.22), the combination $T^iT^jT^k$ has ghost-number two. In combination with the unintegrated vertex V^1 (or OPE contractions thereof with U^j), we arrive at the total ghost number three, as required by the $\langle\lambda^3\theta^5\rangle = 1$ prescription.

In the notation (4.22), the BRST variations $QK^m_{12\ldots k}$ and $QJ^{mn}_{12\ldots k}$ can be written in a unified way as

$$Q\,T^i_{12\ldots k} = \sum_{j=2}^{k}\sum_{\alpha\in P(\beta_j)}(s_{12\ldots j} - s_{12\ldots j-1})\,\big(T_{12\ldots j-1,\{\alpha\}}\,T^i_{j,\{\beta_j\setminus\alpha\}} \;-\; T_{j,\{\beta_j\setminus\alpha\}}\,T^i_{12\ldots j-1,\{\alpha\}}\big)\,.$$

Of course, it has to be kept in mind that only expressions containing a full triplet $T^i_{\ldots}T^j_{\ldots}T^k_{\ldots}$ of loop building blocks are well defined. Recall that the set $\beta_j = j+1, j+2, \ldots, n$ encompasses $n-j$ labels to the right of j, and $P(\beta_j)$ denotes its power set.

4.2 Diagrammatic interpretation of the loop building blocks

According to our discussion above, the $T^i_{\ldots}$ share the symmetry properties and the structure of their Q variation (in particular the Mandelstam variables therein) with the tree-level building blocks $T_{\ldots}$. So we also think of $T^i_{12\ldots k}$ together with the $s_{12}^{-1}, s_{123}^{-1}, \ldots, s_{12\ldots k}^{-1}$ propagators as representing a cubic tree subdiagram.

Since the conformal weight-one fields from U_i can also be contracted with the V_1 vertex, the correlator of (3.1) additionally involves tree-level building blocks $T_{d_1\ldots d_s}$. Hence, every superspace constituent for the open string loop amplitude encompasses four tree-level subdiagrams $T_{\ldots}T^i_{\ldots}T^j_{\ldots}T^k_{\ldots}$, attached to a central vertex with four legs. As a reminder that this is the kinematic factor of a stringy one-loop diagram, we represent this quartic vertex as a box, see figure 4.

We should comment on the shortcoming of the diagrammatic representation figure 4 of $\langle T_{d_1\ldots d_s}\,T^i_{a_1\ldots a_p}\,T^j_{b_1\ldots b_q}\,T^k_{c_1\ldots c_r}\rangle$ that it does not take the asymmetric role of the tree-level BRST building block $T_{d_1\ldots d_s}$ into account, i.e. the lack of $(a_1\ldots a_p) \leftrightarrow (d_1\ldots d_s)$ symmetry. Moving the one-loop building blocks (i.e. the i,j,k superscripts) to different positions amounts to reshuffling contact terms due to the quartic gluon vertex in the SYM action between cubic graphs. For instance, the difference $\langle(T_{12}T^i_3 - T_3T^i_{12})T^j_4T^k_5\rangle$ is proportional to s_{12} when evaluated in components and therefore cancels the propagator present in the common diagram.[6]

[6]In order to see this, consider the two terms on the right hand side of

$$0 = \langle QM^i_{123}T^j_4T^k_5\rangle = \frac{1}{s_{12}}\langle(T_{12}T^i_3 - T_3T^i_{12})T^j_4T^k_5\rangle + \frac{1}{s_{23}}\langle(T_1T^i_{23} - T_{23}T^i_1)T^j_4T^k_5\rangle\,,$$

JHEP08(2014)099

$$\langle T_{d_1\ldots d_s}\, T^i_{a_1\ldots a_p}\, T^j_{b_1\ldots b_q}\, T^k_{c_1\ldots c_r}\rangle \quad = \quad$$

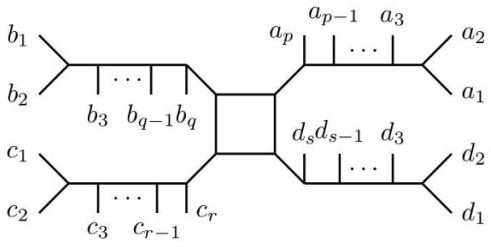

Figure 4. Interpretation of $\langle T_{d_1\ldots d_s}\, T^i_{a_1\ldots a_p}\, T^j_{b_1\ldots b_q}\, T^k_{c_1\ldots c_r}\rangle$ as the kinematic factor of a box diagram. The four tree subdiagrams at the corners are identified with building blocks T and T^i.

A particular motivation for the suggestive box notation comes from the low energy limit of superstring amplitudes. After dimensional reduction to four dimensions, they are supposed to reproduce amplitudes of $N = 4$ SYM — see e.g. [26] for a derivation of the four-point box integral in field theory from a D-dimensional superstring computation in the $\alpha' \to 0$ limit. The fact that only quadruple $T_{\ldots}$ and no triple $T_{\ldots}$ enter the superspace kinematics in the string computation reminds of the "no triangle" property of the underlying field theory [52]. In view of these matching structures in loop diagrams of SYM and kinematic constituents of string amplitudes, we found it natural to represent the central tetravalent vertex gluing together the $T_{\ldots}T^i_{\ldots}T^j_{\ldots}T^k_{\ldots}$ as a box. However, this does not claim a one-to-one correspondence between a particular superspace kinematic factor and a box coefficient in field theory. The systematic reproduction of $N = 4$ SYM amplitudes via $\alpha' \to 0$ limits of the present results is not addressed in this paper and left for future work instead.

4.3 Berends-Giele currents for loop amplitudes

As the next hierarchy level of building blocks we define loop-level Berends-Giele currents, $M^i_{12\ldots p}$ encompassing several tree subdiagrams described by $T^i_{a_1\ldots a_p}$. They are closely related with the field-theory Berends-Giele currents of [43], as thoroughly explained (with examples) in [6]. The collections of subdiagrams $M_{12\ldots p} = \sum T_{a_1\ldots a_p}/s^{p-1}$ which were present in the superspace representations of tree-level amplitudes can be literally carried over to the CFT ingredients of loop amplitudes. In other words, the tree-level formulae (2.16) and (2.17) directly translate into loop-level analogues

$$M^i_{12} = \frac{T^i_{12}}{s_{12}} \tag{4.23}$$

$$M^i_{123} = \frac{T^i_{123}}{s_{12}\, s_{123}} + \frac{T^i_{321}}{s_{23}\, s_{123}} \tag{4.24}$$

$$M^i_{1234} = \frac{1}{s_{1234}}\left(\frac{T^i_{1234}}{s_{12}s_{123}} + \frac{T^i_{3214}}{s_{23}s_{123}} + \frac{T^i_{3421}}{s_{34}s_{234}} + \frac{T^i_{3241}}{s_{23}s_{234}} + \frac{2T^i_{12[34]}}{s_{12}s_{34}}\right). \tag{4.25}$$

see section 4.3 for the definition of M^i_{123}. Cancellations between the first term $\sim s_{12}^{-1}$ and the second one $\sim s_{23}^{-1}$ require that the numerators vanish on the residues of the poles, i.e. $\langle (T_{12}T^i_3 - T_3T^i_{12})T^j_4T^k_5\rangle \sim s_{12}$. This argument can be easily extended to higher multiplicity by virtue of $0 = \langle QM^i_{1234}T^j_5T^k_6\rangle$ and related expressions.

JHEP08(2014)099

It is a necessary condition for BRST invariance that the kinematic factor in loop amplitudes combines the $T^i_{\dots}$ to full-fledged Berends-Giele currents $M^i_{\dots}$. This can be seen from their covariance under Q with no additional Mandelstam factors

$$\begin{aligned} Q\, M^i_{a_1\dots a_p}\, M^j_{b_1\dots b_q}\, M^k_{c_1\dots c_r} = &\sum_{\ell=1}^{p-1} \big(M_{a_1\dots a_\ell}\, M^i_{a_{\ell+1}\dots a_p} - M_{a_{\ell+1}\dots a_p}\, M^i_{a_1\dots a_\ell} \big)\, M^j_{b_1\dots b_q}\, M^k_{c_1\dots c_r} \\ &+ \sum_{\ell=1}^{q-1} \big(M_{b_1\dots b_\ell}\, M^j_{b_{\ell+1}\dots b_q} - M_{b_{\ell+1}\dots b_q}\, M^j_{b_1\dots b_\ell} \big)\, M^i_{a_1\dots a_p}\, M^k_{c_1\dots c_r} \\ &+ \sum_{\ell=1}^{r-1} \big(M_{c_1\dots c_\ell}\, M^k_{c_{\ell+1}\dots c_r} - M_{c_{\ell+1}\dots c_r}\, M^k_{c_1\dots b_\ell} \big)\, M^i_{a_1\dots a_p}\, M^j_{b_1\dots b_q} \end{aligned} \tag{4.26}$$

in close analogy to (2.19) at tree-level. Apart from their simple Q variations (4.26) and their matching with the cubic diagram content of SYM trees, the definitions (4.23) to (4.25) can be motivated by worldsheet integral manipulations (5.15) and (5.16). As detailed in section 5.2, the $M^i_{12\dots p}$ naturally build up once the Green functions are arranged into a form which facilitates integration by parts.

One could also have defined Berends-Giele currents $M^m_{12\dots k}$ and $M^{mn}_{12\dots k}$ for the individual building blocks $K^m_{12\dots k}$ and $J^{mn}_{12\dots k}$ to later define $M^i_{12\dots k}$ by combining them following the pattern seen in (4.22):

$$\begin{aligned} M^i_{a_1\dots a_p}\, M^j_{b_1\dots b_q}\, M^k_{c_1\dots c_r} \equiv \frac{2}{3}\big(&M^m_{a_1\dots a_p}\, M^n_{b_1\dots b_q}\, M^{mn}_{c_1\dots c_r} + M^m_{a_1\dots a_p}\, M^{mn}_{b_1\dots b_q}\, M^n_{c_1\dots c_r} \\ &+ M^{mn}_{a_1\dots a_p}\, M^m_{b_1\dots b_q}\, M^n_{c_1\dots c_r}\big). \end{aligned} \tag{4.27}$$

The combinatorics of zero mode saturation implies that the end result for amplitudes always involves a sum of all the three terms on the right hand side. That is why we will always use the notation on the left hand side of (4.27) in the rest of this work.

4.4 BRST-invariant kinematics for loop amplitudes

Amplitudes computed with the pure spinor formalism give rise to superspace kinematic factors in the cohomology of the BRST operator. We have motivated K and J building blocks from their appearance in the iterated OPEs of integrated vertex operators (along with the d_α and N_{mn} worldsheet fields) and argued that their combinations $M^i_{a_1\dots a_p} M^j_{b_1\dots b_q} M^k_{c_1\dots c_r}$ have covariant BRST variations (4.26) connecting different pole channels. Given the strong constraints which BRST invariance imposes on tree-level SYM amplitudes — see subsection 2.4 — it is natural to explore the Q cohomology using the one-loop building blocks. In this subsection we will write down BRST invariants constructed from the above elements dictated by the minimal formalism. This amounts to anticipating the admissible kinematic structure in the result of the CFT computation of one-loop scattering amplitudes.

As mentioned in subsection 4.2, the one-loop prescription (3.1) containing one unintegrated vertex operator V_1 implies that one tree-level building block $T_{1\dots}$ (combined to a Berends-Giele current $M_{\dots 1\dots}$) has to appear in these BRST invariants, in addition to three one-loop constituents $M^i_{\dots} M^j_{\dots} M^k_{\dots}$. Hence, Q invariant loop kinematics must be built from

JHEP08(2014)099

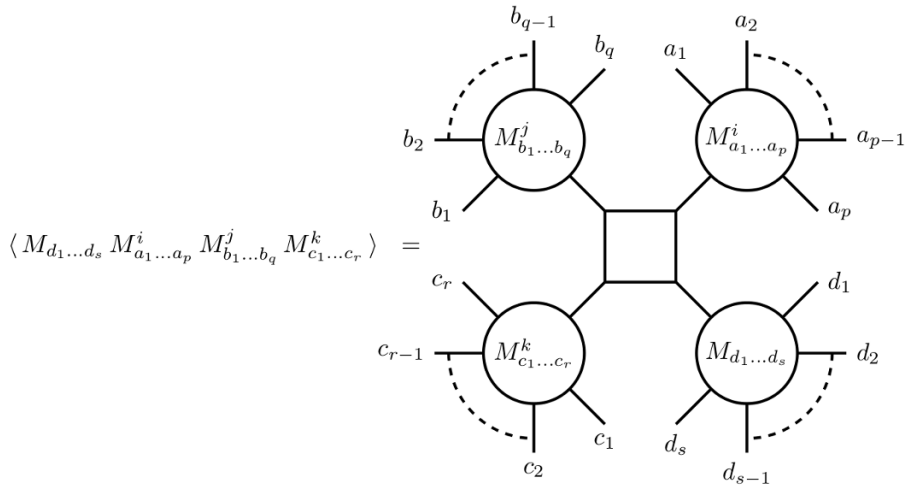

Figure 5. Interpretation of $\langle M_{d_1\ldots d_s} M^i_{a_1\ldots a_p} M^j_{b_1\ldots b_q} M^k_{c_1\ldots c_r} \rangle$ as four Berends-Giele currents (i.e. collections of tree subdiagrams guided by color-ordered tree-level amplitudes), glued together by a central quartic "box"-vertex.

$M_{d_1\ldots d_s} M^i_{a_1\ldots a_p} M^j_{b_1\ldots b_q} M^k_{c_1\ldots c_r}$ with $1 \in \{d_1, \ldots, d_s\}$. The diagrammatic interpretation of such a term follows from the fact that Berends-Giele currents represent color-ordered tree amplitudes with one off-shell leg, see figure 5.

As explained before at the level of $T_{\ldots} T^i_{\ldots} T^j_{\ldots} T^k_{\ldots}$, this diagram does not take the asymmetry in $M_{d_1\ldots d_s} M^i_{a_1\ldots a_p} \leftrightarrow M_{a_1\ldots a_p} M^i_{d_1\ldots d_s}$ into account. The difference between the two $(M_{\ldots}, M^i_{\ldots})$ assignments corresponds to a reshuffling of contact terms in the cubic subdiagrams at the corners of the box.

In the following, we shall give a list of BRST invariants built from $M_{\ldots} M^i_{\ldots} M^j_{\ldots} M^k_{\ldots}$ up to seven-points. They are denoted by $C_{1,a_1\ldots a_p,b_1\ldots b_q,c_1\ldots c_r}$ according to their first term $V^1 M^i_{a_1\ldots a_p} M^j_{b_1\ldots b_q} M^k_{c_1\ldots c_r}$ where the unintegrated vertex is unaffected by OPEs:

$$C_{1,2,3,4} = M_1\, M^i_2\, M^j_3\, M^k_4 \tag{4.28}$$

$$C_{1,23,4,5} = M_1\, M^i_{23}\, M^j_4\, M^k_5 \;+\; M_{12}\, M^i_3\, M^j_4\, M^k_5 \;+\; M_{31}\, M^i_2\, M^j_4\, M^k_5 \tag{4.29}$$

$$\begin{aligned} C_{1,234,5,6} = {} & M_1\, M^i_{234}\, M^j_5\, M^k_6 \;+\; M_{123}\, M^i_4\, M^j_5\, M^k_6 \;+\; M_{412}\, M^i_3\, M^j_5\, M^k_6 \\ & +\; M_{341}\, M^i_2\, M^j_5\, M^k_6 \;+\; M_{12}\, M^i_{34}\, M^j_5\, M^k_6 \;+\; M_{41}\, M^i_{23}\, M^j_5\, M^k_6 \end{aligned} \tag{4.30}$$

$$\begin{aligned} C_{1,23,45,6} = {} & M_1\, M^i_{23}\, M^j_{45}\, M^k_6 \;+\; M_{12}\, M^i_3\, M^j_{45}\, M^k_6 \;-\; M_{13}\, M^i_2\, M^j_{45}\, M^k_6 \\ & +\; M_{14}\, M^i_{23}\, M^j_5\, M^k_6 \;-\; M_{15}\, M^i_{23}\, M^j_4\, M^k_6 \;+\; M_{413}\, M^i_2\, M^j_5\, M^k_6 \\ & +\; M_{512}\, M^i_3\, M^j_4\, M^k_6 \;-\; M_{412}\, M^i_3\, M^j_5\, M^k_6 \;-\; M_{513}\, M^i_2\, M^j_4\, M^k_6 \end{aligned} \tag{4.31}$$

$$\begin{aligned} C_{1,2345,6,7} = {} & M_1\, M^i_{2345}\, M^j_6\, M^k_7 \;+\; M_{512}\, M^i_{34}\, M^j_6\, M^k_7 \;+\; \big[\, M_{12}\, M^i_{345} \;+\; M_{123}\, M^i_{45} \\ & +\; M_{1234}\, M^i_5 \;+\; M_{5123}\, M^i_4 \;-\; (2,3 \leftrightarrow 5,4)\,\big]\, M^j_6\, M^k_7 \end{aligned} \tag{4.32}$$

$$\begin{aligned} C_{1,234,56,7} = {} & M_1\, M^i_{234}\, M^j_{56}\, M^k_7 \;+\; M_{214}\, M^i_3\, M^j_{56}\, M^k_7 \;+\; (M_{15}\, M^i_6 \;-\; M_{16}\, M^i_5)\, M^j_{234}\, M^k_7 \\ & +\; \big[\, M_{12}\, M^i_{34} \;+\; M_{123}\, M^i_4 \;+\; (2 \leftrightarrow 4)\,\big]\, M^j_{56}\, M^k_7 \;+\; \big\{\, \big[\, M_{612}\, M^i_{34}\, M^j_5 \\ & +\; M_{6123}\, M^i_4\, M^j_5 \;+\; M_{5124}\, M^i_3\, M^j_6 \;+\; (2 \leftrightarrow 4)\,\big] \;-\; (5 \leftrightarrow 6)\,\big\}\, M^k_7 \end{aligned} \tag{4.33}$$

JHEP08(2014)099

$$
\begin{aligned}
C_{1,23,45,67} = M_1\, M^i_{23}\, M^j_{45}\, M^k_{67} \;+\; \big[\,(M_{12}\, M^i_3 \;-\; M_{13}\, M^i_2)\, M^j_{45}\, M^k_{67} \;+\; \mathrm{cyc}(23,45,67)\,\big] \\
+\; \big\{\,\big[\,M_{217}\, M^i_3\, M^j_6 \;-\; (2\leftrightarrow 3) \;-\; (6\leftrightarrow 7)\,\big]\, M^k_{45} \;+\; \mathrm{cyc}(23,45,67)\,\big\} \\
+\; \big[\,(M_{7135}+M_{7153})\, M^i_2\, M^j_4\, M^k_6 \;-\; (2\leftrightarrow 3) \;-\; (4\leftrightarrow 5) \;-\; (6\leftrightarrow 7)\,\big]\,.
\end{aligned}
\tag{4.34}
$$

Eight-point amplitudes contain four topologies $C_{1,23456,7,8}$, $C_{1,2345,67,8}$, $C_{1,234,56,78}$ and $C_{1,234,567,8}$ of BRST invariants. They are expanded in appendix C.

4.5 Symmetry properties of the BRST invariants

In this subsection, we examine the symmetry properties of the BRST invariants of the previous subsection and determine the number of independent permutations (at least under linear relations with constant coefficients). In particular, we will argue that the $C_{1,\ldots}$ with label "1" in the first entry form a suitable basis. This ties in with the one-loop prescription (3.1) for string amplitudes: the special role of the unintegrated vertex V_1 implies that only $C_{1,\ldots}$ can appear in the CFT computation, and these ingredients must be able to capture any permutation $C_{i\neq 1,\ldots}$ via linear combination.

In order to see that the reduction to $C_{1,\ldots}$ is possible, first note that the invariants $C_{1,a_1\ldots a_p,b_1\ldots,b_q,c_1\ldots c_r}$ inherit the symmetry properties of the Berends-Giele currents for each of individual three sets of labels $(a_1,\ldots,a_p)$, $(b_1,\ldots,b_q)$ and $(c_1,\ldots,c_r)$, i.e.

$$
M^i_{12\ldots p} = (-1)^{p-1} M_{p\ldots 21}, \qquad M_{\{\beta\}1\{\alpha\}} = (-1)^{n_\beta} \sum_{\sigma\,\in\,\mathrm{OP}(\{\alpha\},\{\beta^T\})} M^i_{1\{\sigma\}}
\tag{4.35}
$$

directly carry over to

$$
\begin{aligned}
C_{1,a_1a_2\ldots a_p,b_1\ldots b_q,c_1\ldots c_r} &= (-1)^{p-1} C_{1,a_p\ldots a_2a_1,b_1\ldots b_q,c_1\ldots c_r} \\
C_{1,\{\beta\}i\{\alpha\},b_1\ldots b_q,c_1\ldots c_r} &= (-1)^{n_\beta} \sum_{\sigma\,\in\,\mathrm{OP}(\{\alpha\},\{\beta^T\})} C_{1,i\{\sigma\},b_1\ldots b_q,c_1\ldots c_r}\,.
\end{aligned}
\tag{4.36}
$$

The notation for the sets α,β,σ is the usual one appearing in the Kleiss-Kuijf relation [42]. The latter implies the subcyclic property (or photon decoupling identity)

$$
\sum_{\sigma\in\mathrm{cyclic}} C_{1,\sigma(a_1a_2\ldots a_p),b_1\ldots b_q,c_1\ldots c_r} = 0.
\tag{4.37}
$$

However, the above symmetries do not relate $C_{i,\ldots}$ to $C_{j\neq i,\ldots}$ (with different labels i,j in the first slot). Equations of that type follow from the BRST cohomology of pure spinor superspace, i.e. from the vanishing of BRST exact terms at ghost number three,

$$
\langle Q\, M^i_{a_1\ldots a_p}\, M^j_{b_1\ldots b_q}\, M^k_{c_1\ldots c_r}\rangle = 0.
\tag{4.38}
$$

The left hand side is always organized into linear combinations of C's, let us illustrate this by examples: the four-point BRST invariant turns out to be totally symmetric,

$$
0 = \langle Q\, M^i_{12}\, M^j_3\, M^k_4\rangle \;\Rightarrow\; C_{1,2,3,4} = C_{2,1,3,4},
\tag{4.39}
$$

JHEP08(2014)099

and five-point invariants can be reduced to $C_{1,ij,k,l} = C_{[1,ij],k,l}$ by means of

$$\begin{aligned} 0 &= \langle Q\, M^i_{123}\, M^j_4\, M^k_5\rangle \;\Rightarrow\; C_{1,23,4,5} = -\,C_{3,21,4,5} \\ 0 &= \langle Q\, M^i_{12}\, M^j_{34}\, M^k_5\rangle \;\Rightarrow\; C_{2,34,1,5} - C_{1,34,2,5} + C_{4,12,3,5} - C_{3,12,4,5} = 0. \end{aligned} \tag{4.40}$$

At six-points, there are three different topologies of BRST exact quantities

$$\begin{aligned} 0 &= \langle Q\, M^i_{1234}\, M^j_5\, M^k_6\rangle \;\Rightarrow\; C_{1,234,5,6} = C_{4,321,5,6} \\ 0 &= \langle Q\, M^i_{123}\, M^j_{45}\, M^k_6\rangle \;\Rightarrow\; C_{1,23,45,6} + C_{3,21,45,6} + C_{4,123,5,6} - C_{5,123,4,6} = 0 \\ 0 &= \langle Q\, M^i_{12}\, M^j_{34}\, M^k_{56}\rangle \;\Rightarrow\; C_{1,34,56,2} - C_{2,34,56,1} + C_{3,12,56,4} - C_{4,12,56,3} \\ &\qquad + C_{5,12,34,6} - C_{6,12,34,5} = 0, \end{aligned} \tag{4.41}$$

and the resulting equations are sufficient to decompose any given $C_{i,jk,lm,n}$ or $C_{i,jkl,m,n}$ into a basis of $C_{1,\ldots}$. A similar recursive argument applies at seven-points due to four types of equations:

$$\begin{aligned} 0 &= \langle Q\, M^i_{12345}\, M^j_6\, M^k_7\rangle \;\Rightarrow\; C_{1,2345,6,7} = -\,C_{5,4321,6,7} \\ 0 &= \langle Q\, M^i_{1234}\, M^j_{56}\, M^k_7\rangle \;\Rightarrow\; C_{5,1234,6,7} - C_{6,1234,5,7} + C_{1,234,56,7} - C_{4,321,56,7} = 0 \\ 0 &= \langle Q\, M^i_{123}\, M^j_{456}\, M^k_7\rangle \;\Rightarrow\; C_{1,456,23,7} + C_{3,456,21,7} + C_{4,123,56,7} + C_{6,123,54,7} = 0 \\ 0 &= \langle Q\, M^i_{123}\, M^j_{45}\, M^k_{67}\rangle \;\Rightarrow\; C_{1,23,45,67} + C_{3,21,45,67} + C_{4,123,67,5} - C_{5,123,67,4} \\ &\qquad + C_{6,123,45,7} - C_{7,123,45,6} = 0 \end{aligned} \tag{4.42}$$

In order to count the number of independent $C_{1,a_1\ldots a_p,b_1\ldots,b_q,c_1\ldots c_r}$, one should keep in mind that there are $(p-1)!$ independent components in $(a_1\ldots a_p)$ due to Berends-Giele symmetries, the same number of cyclically inequivalent configurations. Hence, the number of independent $C_{1,a_1\ldots a_p,b_1\ldots,b_q,c_1\ldots c_r}$ in n-point amplitudes (where $p+q+r=n-1$) is given by the number of ways to distribute $n-1$ elements to three cycles $(a_1\ldots a_p), (b_1\ldots,b_q)$ and $(c_1\ldots c_r)$. This is the defining property of the unsigned Stirling numbers of first kind S_3^{n-1},

$$\begin{aligned} &\#(C_{1,a_1\ldots a_p,b_1\ldots,b_q,c_1\ldots c_r})\Big|_{p+q+r=n-1} = S_3^{n-1} \\ &\quad = 1, 6, 35, 225, 1624, 13132, 118124, 1172700, 12753576, 150917976, \ldots \qquad n \geq 4, \end{aligned} \tag{4.43}$$

and the following table gathers examples of how individual topologies (i.e. different triplets of p, q, r with constant sum) contribute to the Stirling numbers:

5 One-loop amplitudes in pure spinor superspace

The pure spinor BRST cohomology of building blocks will now be used to deduce the form of the n-point one-loop open superstring amplitudes. Apart from the four- and five-point amplitudes which were previously computed without explicit use of building blocks [5, 9, 63], the results for higher-points are strongly guided by their cohomology properties.

From the discussion of section 3, the n-point kinematic factor for one-loop amplitudes is given, up to OPE terms with the b-ghost, by the following correlator

$$K_n = \langle V_1 U_2(z_2) U_3(z_3) \ldots U_n(z_n)\rangle_{ddN} \tag{5.1}$$

JHEP08(2014)099

n	C-topology	# components	
4	$C_{1,2,3,4}$	1	$= 1$
5	$C_{1,23,4,5}$	$\binom{4}{2}$	$= 6$
6	$C_{1,234,5,6}$	$\binom{5}{2} \cdot 2$	$= 20$
6	$C_{1,23,45,6}$	$5 \cdot 3$	$= 15$
7	$C_{1,2345,6,7}$	$\binom{6}{2} \cdot 6$	$= 90$
7	$C_{1,234,56,7}$	$\binom{6}{3} \cdot 3 \cdot 2$	$= 120$
7	$C_{1,23,45,67}$	$5!!$	$= 15$
8	$C_{1,23456,7,8}$	$\binom{7}{2} \cdot 24$	$= 504$
8	$C_{1,2345,67,8}$	$\binom{7}{3} \cdot 3 \cdot 6$	$= 630$
8	$C_{1,234,56,78}$	$\binom{7}{3} \cdot 3 \cdot 2$	$= 210$
8	$C_{1,234,567,8}$	$\frac{1}{2} \cdot 7 \cdot \binom{6}{3} \cdot 2 \cdot 2$	$= 280$

Table 1. The number of independent components of C topologies up to $n = 8$.

where the subscript ddN is a reminder that the substitution rule (3.8) must be applied. It is easy to see that $n-4$ OPE contractions among the vertex operators will have to be performed before the zero-mode combination $d_\alpha d_\beta N^{mn}$ can be extracted. Throughout this section, we will immediately trade all the OPE residues $L_{2131\ldots\ell 1}$ and $\tilde{K}^m_{\ell+1\ldots p}, \tilde{J}^{mn}_{\ell+1\ldots p}$ for the corresponding BRST building blocks $T_{12\ldots\ell}$ and $K^m_{\ell+1\ldots p}, J^{mn}_{\ell+1\ldots p}$. Experience with the tree-level computation [6, 17, 18] shows that their difference can only contribute to BRST trivial kinematics and drops out through total worldsheet derivatives.

The calculation of the kinematic factor will be divided into three steps:

1. Express the correlator (5.1) in terms of BRST building blocks
2. Group these building blocks into Berends-Giele currents
3. Use integration by parts to combine different currents to BRST invariants $C_{1,\ldots}$

Starting from six-points, we will use BRST invariance as an extra input in steps 1 and 2 to fix certain parts of the correlator: this concerns the failure of $\eta_{ij}\eta_{jk}$ products to obey the partial fraction identity $(z_{ij}z_{jk})^{-1} + \mathrm{cyc}(ijk) = 0$ from tree-level. This relation plays an important role for the basis reduction of worldsheet integrals at tree-level, see [7]. After these steps are performed the correlator (5.1) becomes a linear combination of the BRST invariants $C_{1,\ldots}$ constructed in subsection 4.4, which we can regard as the one-loop analogue of the tree-level subamplitudes $\mathcal{A}^{\mathrm{YM}}$. Hence, up to the aforementioned partial fraction subtlety, the one-loop strategy follows the same logical step as the calculation of the n-point tree amplitude in [6].

Imposing BRST invariance from the beginning makes us blind to the hexagon anomaly in $D = 10$ dimensions arising from the boundary of the t integration [21, 22], so in our method we are not able to reproduce the superspace anomaly computed in [74]. In other words, we compute the non-anomalous or BRST-invariant part of the amplitude.

JHEP08(2014)099

Although our final result for K_n won't include leg one on the same footing as all the others, we will prove its hidden total symmetry in subsection 7.4. The basis choice $C_{1,\dots}$ for the kinematic constituents reflects the special role played by leg one entering the computation through the unintegrated V^1 vertex. New cross-connections to color structures at tree-level will be pointed out in section 7 which trivialize the outstanding symmetry proof.

5.1 Step 1: CFT correlator in terms of building blocks

Using the definitions of the building blocks, the CFT correlator (5.1) will encompass all possible combinations of building blocks allowed by its total permutation symmetry in $(234\dots n)$. As mentioned before, $n-4$ OPE contractions must be performed before the ddN zero-modes can be extracted and leave a triplet of building blocks $T^i_{\dots}T^j_{\dots}T^k_{\dots}$ behind.

As a trivial starting example, the four-point kinematic factor does not require any OPE and can be written down immediately using the definitions (4.6), (4.9) and (4.22)

$$K_4 = \langle V_1 U_2 U_3 U_4 \rangle_{ddN} = \langle V_1 T_2^i T_3^j T_4^k \rangle \,. \tag{5.2}$$

The ten possible OPEs in the five-point kinematic factor give rise to two classes of terms, depending on whether the contraction involves the unintegrated vertex or not:

$$\begin{aligned} K_5 &= \langle V_1 U_2 U_3 U_4 U_5 \rangle_{ddN} \\ &= \langle\, V_1 \underbrace{U_2\, U_3}\, U_4\, U_5 \,\rangle \;+\; 5\,\text{permutations}\,(23 \leftrightarrow 24, 25, 34, 35, 45) \\ &\quad + \langle \underbrace{V_1\, U_2}\, U_3\, U_4\, U_5 \rangle \;+\; 3\,\text{permutations}\,(2 \leftrightarrow 3, 4, 5) \end{aligned} \tag{5.3}$$

The resulting BRST building blocks are

$$\langle \underbrace{V_1\, U_2}\, U_3\, U_4\, U_5 \rangle = \eta_{12}\, \langle T_{12}\, T_3^i T_4^j T_5^k \rangle \tag{5.4}$$

$$\langle\, V_1 \underbrace{U_2\, U_3}\, U_4\, U_5 \,\rangle = \eta_{23}\, \langle V_1\, T_{23}^i T_4^j T_5^k \rangle, \tag{5.5}$$

and the validity of the replacement $L_{21} \mapsto T_{12}$ follows from BRST-closedness of $T_3^i T_4^j T_5^k$.

Applying this kind of analysis to the six-point correlator leads to an ambiguity:

$$\begin{aligned} K_6 &= \langle V_1 U_2 U_3 U_4 U_5 U_6 \rangle_{ddN} \\ &= \langle\, \underbrace{V_1\, U_2\, U_3}\, U_4\, U_5\, U_6 \,\rangle + \langle\, \underbrace{V_1\, U_2}\, \underbrace{U_3\, U_4}\, U_5\, U_6 \,\rangle \\ &\quad + \langle\, V_1\, \underbrace{U_2\, U_3\, U_4}\, U_5\, U_6 \,\rangle + \langle\, V_1\, \underbrace{U_2\, U_3}\, \underbrace{U_4\, U_5}\, U_6 \,\rangle \\ &\quad + \text{permutations} \;+\; \eta_{ijk}(\dots). \end{aligned} \tag{5.6}$$

We firstly find those contractions which closely resemble the tree-level procedure (up to $z_{ij}^{-1} \mapsto \eta_{ij}$ and the new building blocks $T^i_{\dots}T^j_{\dots}T^k_{\dots}$)

$$\langle\, \underbrace{V_1\, U_2\, U_3}\, U_4\, U_5\, U_6 \,\rangle = \eta_{12}\eta_{23}\, \langle T_{123}\, T_4^i T_5^j T_6^k \rangle \tag{5.7}$$

$$\langle\, \underbrace{V_1\, U_2}\, \underbrace{U_3\, U_4}\, U_5\, U_6 \,\rangle = \eta_{12}\eta_{34}\, \langle T_{12}\, T_{34}^i T_5^j T_6^k \rangle \tag{5.8}$$

JHEP08(2014)099

$$\langle\, V_1\, \underbrace{U_2\, U_3\, U_4}\, U_5\, U_6\,\rangle = \eta_{23}\eta_{34}\,\langle V_1\, T^i_{234} T^j_5 T^k_6\rangle \tag{5.9}$$

$$\langle\, V_1\, \underbrace{U_2\, U_3}\, \underbrace{U_4\, U_5}\, U_6\,\rangle = \eta_{23}\eta_{45}\,\langle V_1\, T^i_{23} T^j_{45} T^k_6\rangle. \tag{5.10}$$

But in addition to that, the correlator could contain terms with worldsheet functions

$$\eta_{ijk} = \eta_{ij}\,\eta_{ik} + \eta_{ji}\,\eta_{jk} + \eta_{ki}\,\eta_{kj}, \tag{5.11}$$

which are invisible in the $z_i \to z_j$ limit since $(z_{ij}z_{ik})^{-1} + \mathrm{cyc}(ijk) = 0$. These parts of the CFT correlator cannot be fixed on the basis of the leading OPE singularity and symmetry arguments in $(23\ldots n)$. Instead, we will keep them undetermined for the moment and use BRST invariance in the following subsections to argue their absence in the end result. The precise way to combine permutations will be discussed in the next subsection.

Similarly, the seven-point kinematic factor receives contributions from

$$\begin{aligned} K_7 &= \langle V_1 U_2 U_3 U_4 U_5 U_6 U_7\rangle_{ddN} \\ &= \langle\, \underbrace{V_1\, U_2\, U_3\, U_4}\, U_5\, U_6 U_7\,\rangle + \langle\, \underbrace{V_1\, U_2\, U_3}\, \underbrace{U_4\, U_5}\, U_6\, U_7\,\rangle + \langle\, \underbrace{V_1 U_2}\, \underbrace{U_3\, U_4\, U_5}\, U_6\, U_7\,\rangle \\ &\quad + \langle\, \underbrace{V_1 U_2}\, \underbrace{U_3\, U_4}\, \underbrace{U_5\, U_6}\, U_7\,\rangle + \langle\, V_1\, \underbrace{U_2 U_3\, U_4\, U_5}\, U_6\, U_7\,\rangle + \langle\, V_1\, \underbrace{U_2 U_3\, U_4}\, \underbrace{U_5\, U_6}\, U_7\,\rangle \\ &\quad + \langle\, V_1\, \underbrace{U_2 U_3}\, \underbrace{U_4\, U_5}\, \underbrace{U_6\, U_7}\,\rangle + \text{permutations } + \ \eta_{ijk}(\ldots), \end{aligned} \tag{5.12}$$

where the seven different types of OPEs yield

$$\begin{aligned} \langle\, \underbrace{V_1\, U_2\, U_3\, U_4}\, U_5\, U_6 U_7\,\rangle &= \eta_{12}\eta_{23}\eta_{34}\langle T_{1234} T^i_5 T^j_6 T^k_7\rangle \\ \langle\, \underbrace{V_1\, U_2\, U_3}\, \underbrace{U_4\, U_5}\, U_6\, U_7\,\rangle &= \eta_{12}\eta_{23}\eta_{45}\langle T_{123} T^i_{45} T^j_6 T^k_7\rangle \\ \langle\, \underbrace{V_1 U_2}\, \underbrace{U_3\, U_4\, U_5}\, U_6\, U_7\,\rangle &= \eta_{12}\eta_{34}\eta_{45}\langle T_{12} T^i_{345} T^j_6 T^k_7\rangle \\ \langle\, \underbrace{V_1 U_2}\, \underbrace{U_3\, U_4}\, \underbrace{U_5\, U_6}\, U_7\,\rangle &= \eta_{12}\eta_{34}\eta_{56}\langle T_{12} T^i_{34} T^j_{56} T^k_7\rangle \\ \langle\, V_1\, \underbrace{U_2 U_3\, U_4\, U_5}\, U_6\, U_7\,\rangle &= \eta_{23}\eta_{34}\eta_{45}\langle V_1 T^i_{2345} T^j_6 T^k_7\rangle \\ \langle\, V_1\, \underbrace{U_2 U_3\, U_4}\, \underbrace{U_5\, U_6}\, U_7\,\rangle &= \eta_{23}\eta_{34}\eta_{56}\langle V_1 T^i_{234} T^j_{56} T^k_7\rangle \\ \langle\, V_1\, \underbrace{U_2 U_3}\, \underbrace{U_4\, U_5}\, \underbrace{U_6\, U_7}\,\rangle &= \eta_{23}\eta_{45}\eta_{67}\langle V_1 T^i_{23} T^j_{45} T^k_{67}\rangle. \end{aligned} \tag{5.13}$$

These six- and seven-point cases give an idea of the general pattern for the n-point correlator: the kinematic factor K_n encompasses all tree-level building blocks involving the unintegrated vertex $\eta_{12}\eta_{23}\ldots\eta_{\ell-1,\ell}T_{12\ldots\ell}$, multiplied with all the possible topologies of $(\eta^{p-\ell-1}T^i_{\ell+1\ldots p})(\eta^{q-p-1}T^j_{p+1\ldots q})(\eta^{n-q-1}T^k_{q+1\ldots n})$ of the remaining $n-\ell$ legs where zero modes of $d_\alpha d_\beta N^{mn}$ are extracted:

$$\begin{aligned} &\langle V_1(z_1)U_2(z_2)U_3(z_3)\ldots U_n(z_n)\rangle_{ddN} \\ &\quad = \langle \sum_{\ell=1}^{n-3} (\eta_{12}\,\ldots\,\eta_{\ell-1,\ell}\,T_{12\ldots\ell}) \sum_{\ell+1\le p<q<n} (\eta_{\ell+1,\ell+2}\,\ldots\,\eta_{p-1,p}\,T^i_{\ell+1\ldots p}) \\ &\qquad \times\ (\eta_{p+1,p+2}\ldots\eta_{q-1,q}\,T^j_{p+1\ldots q})\ (\eta_{q+1,q+2}\ldots\eta_{n-1,n}\,T^k_{q+1\ldots n}) \\ &\qquad +\ \text{permutations } +\ \eta_{ijk}(\ldots)\,\rangle. \end{aligned} \tag{5.14}$$

JHEP08(2014)099

The next tasks to be addressed in the following subsections are to trade the BRST building blocks for Berends-Giele currents and to resolve the ambiguity about the η_{ijk} terms.

5.2 Step 2: Berends-Giele currents

In the n-point tree amplitude computations of [6] the worldsheet integrands combine the BRST building blocks with z_{ij} poles via $T_{12\ldots p} \leftrightarrow (z_{12}z_{23}\ldots z_{p-1,p})^{-1}$. The essential step for further simplification lies in trading $T_{\ldots}$ for Berends-Giele currents $M_{\ldots}$ using the identity

$$\frac{T_{12\ldots p}}{z_{12}\,z_{23}\ldots z_{p-1,p}} + \mathrm{sym}(23\ldots p) = M_{12\ldots p}\prod_{k=2}^{p}\sum_{m=1}^{k-1}\frac{s_{mk}}{z_{mk}} + \mathrm{sym}(23\ldots p) \tag{5.15}$$

It has already been proven at tree-level [8] that the Berends-Giele currents are the natural objects to describe the SYM amplitudes. The identity (5.15) was the key step in identifying the n-point superstring amplitude as sum of $(n-3)!$ SYM amplitudes [6] dressed by hypergeometric worldsheet integrals [7]. To what extent can the tree-level identity (5.15) and its corollaries be generalized to one-loop?

In order to answer this question note that the tree-level proof of (5.15) required two assumptions: the symmetries of the building blocks and the partial fraction identities $(z_{ij}z_{jk})^{-1} + \mathrm{cyc}(ijk) = 0$. As the loop building blocks $T^{i,j,k}_{\ldots}$ obey the same symmetry identities as their tree-level counterparts, the only obstacle against a direct one-loop generalization of (5.15) comes from the fact that the functions $\eta_{ij}\eta_{ik}$ do not obey a similar partial fraction identity in general. That is why we have defined the totally symmetric function $\eta_{ijk} = \eta_{ij}\eta_{ik} + \eta_{ji}\eta_{jk} + \eta_{ki}\eta_{kj}$ measuring the failure of the tree-level partial fraction identity to hold at higher genus. With this definition at hand, the one-loop generalization of (5.15) is

$$\begin{aligned}\eta_{12}\,\eta_{23}\ldots\eta_{p-1,p}\,T_{12\ldots p} \;+\; \mathrm{sym}(23\ldots p) \;\;=\;\; & M_{12\ldots p}\prod_{k=2}^{p}\sum_{m=1}^{k-1} s_{mk}\,\eta_{mk} \\ & + \mathrm{sym}(23\ldots p) \;+\; \eta_{ijk}(\ldots)\;.\end{aligned} \tag{5.16}$$

Of course, the same identity holds for the loop cousins $(T_{\ldots}, M_{\ldots}) \mapsto (T^i_{\ldots}, M^i_{\ldots})$ since the $T^i_{\ldots}$ enjoy the same symmetry properties as the tree-level building blocks $T_{\ldots}$ and the definition of $M^i_{\ldots}$ in terms of $T^i_{\ldots}$ incorporates the same functional dependence as $M_{\ldots}$ expressed in terms of $T_{\ldots}$.

We will show in the following subsection that discarding η_{ijk} corrections in both (5.16) and (5.14) yields BRST-invariant kinematic factors describing non-anomalous terms in the amplitude.[7] Let us see some examples. For the four-point correlator (5.2) the trading identity is trivial in view of $V_1 = T_1 = M_1$ and $T^i_2 = M^i_2$,

$$K_4 = \langle V_1 T^i_2 T^j_3 T^k_4\rangle = \langle M_1 M^i_2 M^j_3 M^k_4\rangle. \tag{5.17}$$

[7] In fact, η_{ijk} terms are related to the gauge anomaly discussion of [21, 22] in a subtle way. As pointed out by Michael Green [85] and elaborated in future work, the associated BRST non-invariant kinematic factors responsible for the hexagon anomaly might leave gauge-invariant fingerprints in certain subsectors of the amplitude.

JHEP08(2014)099

In order to prevent overcrowding in the formulæ below the following shorthand notation will be used

$$X_{ij} \;\equiv\; s_{ij}\eta_{ij}. \tag{5.18}$$

The five-point correlator (5.3) is also rather trivially converted to Berends-Giele currents $M_{12} = T_{12}/s_{12}$ and $M^i_{23} = T^i_{23}/s_{23}$. The permutations generated by (5.4) and (5.5) combine to ten terms

$$\begin{aligned} K_5 = \langle & X_{12}\, M_{12}\, M_3^i\, M_4^j\, M_5^k + X_{13}\, M_{13}\, M_2^i\, M_4^j\, M_5^k + X_{14}\, M_{14}\, M_2^i\, M_3^j\, M_5^k \\ & + X_{15}\, M_{15}\, M_2^i\, M_3^j\, M_4^k + X_{23}\, M_1\, M_{23}^i\, M_4^j\, M_5^k + X_{24}\, M_1\, M_{24}^i\, M_3^j\, M_5^k \\ & + X_{25}\, M_1\, M_{25}^i\, M_3^j\, M_4^k + X_{34}\, M_1\, M_{34}^i\, M_2^j\, M_5^k + X_{35}\, M_1\, M_{35}^i\, M_2^j\, M_4^k \\ & + X_{45}\, M_1\, M_{45}^i\, M_2^j\, M_3^k \,\rangle. \end{aligned} \tag{5.19}$$

The six-point amplitude is the first instance where the identity (5.16) finds non-trivial application. Dropping the terms proportional to η_{ijk} in lines with the BRST reasoning, the six-point topologies (5.7)—(5.10) give rise to

$$\begin{aligned} K_6 = \langle & \text{ 10 terms } \big[\, M_{123}\, X_{12}(X_{13}+X_{23}) + (2 \leftrightarrow 3) \,\big]\, M_4^i\, M_5^j\, M_6^k \\ & +\; \text{ 30 terms } X_{12}\, M_{12}\, X_{34}\, M_{34}^i\, M_5^j\, M_6^k \\ & +\; \text{ 15 terms } M_1\, X_{23}\, M_{23}^i\, X_{45}\, M_{45}^j\, M_6^k \\ & +\; \text{ 10 terms } M_1 \big[\, M_{234}^i\, X_{23}(X_{24}+X_{34}) + (3 \leftrightarrow 4) \,\big]\, M_5^j\, M_6^k \,\rangle. \end{aligned} \tag{5.20}$$

At this point, we shall be more explicit about the permutations within the correlator. As mentioned before, the correlator must be symmetric in all the legs $(23\ldots n)$ of integrated vertices, but the last term in K_6 only contains 2×10 out of the 60 possible terms $M^i_{pqr}X_{pq}(X_{pr}+X_{qr})$ with $p,q,r\in\{2,3,4,5,6\}$. It turns out that by the symmetry properties of Berends-Giele currents (e.g. $M^i_{234} = M^i_{432}$ and $M^i_{234} + \text{cyc}(234) = 0$ in the rank three case at hand), the expression

$$M^i_{23\ldots p}\left(\prod_{k=3}^{p}\sum_{m=2}^{k-1} X_{mk}\right) \;+\; \text{sym}(34\ldots p) \tag{5.21}$$

is secretly totally symmetric[8] in $(23\ldots p)$ even though only the smaller symmetry in $(34\ldots p)$ is manifest. That is why each of the ten choices to single out three legs from $\{2,3,4,5,6\}$ realizes two out of six possible terms only, without spoiling the overall (23456) symmetry.

[8]The same hidden symmetry occurs in the representation [29]

$$\mathcal{M}_n^{\text{YM}} \;\sim\; \sum_{\sigma\in S_{n-2}} f^{1\sigma(2)a}\, f^{a\sigma(3)b}\, f^{b\sigma(4)c} \ldots f^{y\sigma(n-2)z}\, f^{z\sigma(n-1)n}\, \mathcal{A}^{\text{YM}}\big(1,\sigma(2,3,\ldots,n-1),n\big)$$

of the color-dressed SYM amplitude: the structure constant contractions $(f^{bcd})^{n-2}$ share the symmetry properties of the integrand $\prod_{k=3}^{p}\sum_{m=2}^{k-1} X_{mk}$ and the rank p Berends-Giele current taking the role of a $(p+1)$-point SYM amplitude with one off-shell leg guarantees that the color-ordered $\mathcal{A}_n^{\text{YM}}$ have the same symmetry properties as the $M^i_{\ldots}$. Hence, the total symmetry of $\mathcal{M}_n^{\text{YM}}$ implies that of (5.21) by virtue of the dictionary explained above.

JHEP08(2014)099

It is crucial to note the symmetry properties of the two sides of the $T_{\dots} \leftrightarrow M_{\dots}$ trading identity (5.15). The left-hand side is totally symmetric at tree-level, even in trading leg one for one of the others. But this makes use of partial fraction relations that cause extra terms $\sim \eta_{ijk}$ at loops. The z_i dependence on the right hand side, however, is built from combinations s_{ij}/z_{ij} where it is obvious from the Mandelstam factors that there are no partial fractions at work to see the symmetries. Only the right hand side of (5.15) stays totally symmetric in $(12\dots p)$ under the loop-conversion $s_{ij}/z_{ij} \to s_{ij}\eta_{ij} = X_{ij}$ of worldsheet functions.

For these reasons, the following expression for the seven-point kinematic factor,

$$\begin{aligned} K_7 = \langle\, & 15 \text{ terms } M_1 \left[M^i_{2345}\, X_{23}\, (X_{24}+X_{34})(X_{25}+X_{35}+X_{45}) + \text{sym}(345) \right] M^j_6\, M^k_7 \\ & + \ 60 \text{ terms } M_1 \left[M^i_{234}\, X_{23}\, (X_{24}+X_{34}) + (3 \leftrightarrow 4) \right] X_{56}\, M^j_{56}\, M^k_7 \\ & + \ 15 \text{ terms } M_1\, X_{23}\, M^i_{23}\, X_{45}\, M^j_{45}\, X_{67}\, M^k_{67} \\ & + \ 20 \text{ terms } \left[M_{1234}\, X_{12}\, (X_{13}+X_{23})(X_{14}+X_{24}+X_{34}) + \text{sym}(234) \right] M^i_5\, M^j_6\, M^k_7 \\ & + \ 90 \text{ terms } \left[M_{123}\, X_{12}\, (X_{13}+X_{23}) + (2 \leftrightarrow 3) \right] X_{45}\, M^i_{45}\, M^j_6\, M^k_7 \\ & + \ 60 \text{ terms } X_{12}\, M_{12} \left[M^i_{345}\, X_{34}\, (X_{35}+X_{45}) + (4 \leftrightarrow 5) \right] M^j_6\, M^k_7 \\ & + \ 90 \text{ terms } X_{12}\, M_{12}\, X_{34}\, M^i_{34}\, X_{56}\, M^j_{56}\, M^k_7 \rangle \end{aligned} \tag{5.22}$$

is totally symmetric even though only those six $M^i_{\sigma(2345)}$ permutations $\sigma \in S_4$ with fixed point $\sigma(2) = 2$ occur.

The n-point generalization of the above patterns is given by

$$\begin{aligned} K_n = \langle & \sum_{\ell=1}^{n-3} M_{12\dots\ell} \left(\prod_{k=2}^{\ell} \sum_{m=1}^{k-1} X_{mk} \right) \sum_{\ell+1 \le p < q < n} M^i_{\ell+1\dots p} \left(\prod_{k=\ell+2}^{p} \sum_{m=\ell+1}^{k-1} X_{mk} \right) \\ & \times\ M^j_{p+1\dots q} \left(\prod_{k=p+2}^{q} \sum_{m=p+1}^{k-1} X_{mk} \right) M^k_{q+1\dots n} \left(\prod_{k=q+2}^{n} \sum_{m=q+1}^{k-1} X_{mk} \right) \\ & + \text{ permutations} \rangle\,. \end{aligned} \tag{5.23}$$

5.3 Step 3: integration by parts

In this step the number of one-loop worldsheet integrals will be reduced using partial integration identities. These manipulations have been crucial in the computation of the n-point disk amplitude [6, 7] and had already found appearance in the string inspired rules towards field theory amplitudes [83, 84]. As emphasized in the references, integration by parts allows to eliminate double derivatives of the bosonic Green function.

After this reduction is performed the kinematic factor for the one-loop amplitude becomes a sum over manifestly BRST invariant objects multiplied by $n-4$ powers of X_{ij}; schematically, this means $K_n = \sum X^{n-4} \langle C_{1,\dots} \rangle$.

In order to see how these partial integrations can be performed note that the worldsheet integrands at any loop order contain a universal factor proportional to the correlation function of the plane wave exponential factors, the so-called Koba-Nielsen factor

$$\text{KN} \ \equiv\ \Big\langle \prod_{i=1}^{n} e^{ik_i \cdot x(z_i,\overline{z}_i)} \Big\rangle \ \propto\ \exp\Big(\sum_{i<j}^{n} s_{ij}\, \langle x(z_i,\overline{z}_i)\, x(z_j,\overline{z}_j) \rangle \Big)\,. \tag{5.24}$$

JHEP08(2014)099

The precise form of the bosonic Green's function $\langle x(z_i,\overline{z}_i)x(z_j,\overline{z}_j)\rangle$ in terms of Jacobi theta functions is irrelevant for the analysis in the following. What matters is its appearance in the Koba-Nielsen factor and the antisymmetry of its derivative $\eta_{ij} = \frac{\partial}{\partial z_i}\langle x(z_i,\overline{z}_i)x(z_j,\overline{z}_j)\rangle = -\eta_{ji}$ which can be viewed as the one-loop generalization of the $1/z_{ij}$ single pole at tree-level. The form (5.24) of the Koba-Nielsen factor implies that the combinations $X_{ij} = s_{ij}\eta_{ij}$ can be integrated by parts

$$0 = \int \frac{\partial}{\partial z_i}\,\mathrm{KN} = \int \sum_{j\neq i} s_{ij}\,\eta_{ij}\,\mathrm{KN}. \tag{5.25}$$

Boundary terms at $z_i = z_{i+1}$ do not contribute to (5.25) because the Koba-Nielsen factor vanishes for real and positive $s_{i,i+1}$ as $|z_i - z_{i+1}|^{s_{i,i+1}}$. Analytic continuation in the complex $s_{i,i+1}$ plane allows to extend this argument to generic kinematic regimes.

This identity still holds in presence of further η_{pq} factors in the integrand as long as none of the p,q labels coincides with the differentiation leg i, for instance

$$\begin{aligned}
&\int \mathrm{KN}\; X_{12}\,(X_{13}+X_{23}) = \int \mathrm{KN}\;(X_{34}+X_{35}+\cdots+X_{3n})\,(X_{23}+X_{24}+\cdots+X_{2n})\\
&\int \mathrm{KN}\;\prod_{k=2}^{p}\sum_{m=1}^{k-1} X_{mk} = \int \mathrm{KN}\;\prod_{k=2}^{p}\sum_{m=k+1}^{n} X_{km}.
\end{aligned} \tag{5.26}$$

The ubiquitous $\prod_{k=2}^{p}\sum_{m=1}^{k-1} X_{mk}$ products in equation (5.23) for K_n turn out to be maximally partial-integration-friendly. This has already been exploited in tree-level computations [6].

Once we have removed any appearance of z_1 from X_{ij} via integration by parts (5.26), the remaining terms in the correlator will build up various BRST invariants $C_{1,\ldots}$. This is a trivial statement in the four-point correlator (4.28),

$$K_4 = \langle M_1 M_2^i M_3^j M_4^k\rangle = \langle C_{1,2,3,4}\rangle \tag{5.27}$$

whereas the five-point kinematic factor requires $X_{12} = X_{23}+X_{24}+X_{25}$ and (2345) permutations thereof (which is valid under integration against KN only). After eliminating the X_{1j} at $j = 2,3,4,5$ in (5.19), we find the manifestly BRST-invariant expression

$$\begin{aligned}
K_5 = {}& X_{23}\langle M_1 M_{23}^i M_4^j M_5^k + M_{12}M_3^i M_4^j M_5^k + M_{31}M_2^i M_4^j M_5^k\rangle\\
&+ X_{24}\langle M_1 M_{24}^i M_3^j M_5^k + M_{12}M_4^i M_3^j M_5^k + M_{41}M_2^i M_3^j M_5^k\rangle\\
&+ X_{25}\langle M_1 M_{25}^i M_3^j M_4^k + M_{12}M_5^i M_3^j M_4^k + M_{51}M_2^i M_3^j M_4^k\rangle\\
&+ X_{34}\langle M_1 M_{34}^i M_2^j M_5^k + M_{13}M_4^i M_2^j M_5^k + M_{41}M_2^i M_3^j M_5^k\rangle\\
&+ X_{35}\langle M_1 M_{35}^i M_2^j M_4^k + M_{13}M_5^i M_2^j M_4^k + M_{51}M_2^i M_3^j M_4^k\rangle\\
&+ X_{45}\langle M_1 M_{45}^i M_2^j M_3^k + M_{14}M_5^i M_2^j M_3^k + M_{51}M_2^i M_3^j M_4^k\rangle\\
={}& X_{23}\langle C_{1,23,4,5}\rangle + X_{24}\langle C_{1,24,3,5}\rangle + X_{25}\langle C_{1,25,3,4}\rangle\\
&+ X_{34}\langle C_{1,34,2,5}\rangle + X_{35}\langle C_{1,35,2,4}\rangle + X_{45}\langle C_{1,45,2,3}\rangle
\end{aligned} \tag{5.28}$$

JHEP08(2014)099

which agrees with the expression from [9] when its component expansion is evaluated [14]. The total worldsheet derivatives are suppressed in (5.28) and in all subsequent kinematic factors.

The general lesson to learn from the five-point computation concerns the choice of integral basis and the role of the $M_{12\ldots p}$ terms in (5.23) with leg one attached and $p \geq 2$. Once we eliminate z_1 from every X_{rs} in the integrand, the remaining X^{n-4} polynomials are guaranteed to be minimal under (5.25) and the superfield prefactors must be BRST closed $C_{1,\ldots}$. The superfields associated with the integrands X_{1j} outside the desired basis have in common that leg one is attached to a rank $p \geq 2$ Berends-Giele current $M_{12\ldots p}$. After integration by parts, the worldsheet dependence will be transformed into z_1 independent X_{rs} combinations $(r, s \neq 1)$, so the associated $V_1 M^i_{23\ldots p} M^j_{p+1\ldots q} M^k_{q+1\ldots n}$ permutations will receive corrections containing $M_{12\ldots p}$ at $p \geq 2$. Hence, the job of all the $M_{12\ldots p}$ is to provide the BRST invariant completion of $V_1 M^i_{23\ldots p} M^j_{p+1\ldots q} M^k_{q+1\ldots n}$ to form $C_{1,23\ldots p,p+1\ldots q,q+1\ldots n}$.

Let us consider the six-point amplitude to see these mechanisms in action. The first two lines in (5.20) require integration by parts in the form $X_{12}X_{34} = X_{34}(X_{23}+X_{24}+X_{25}+X_{26})$ and $X_{12}(X_{13}+X_{23}) = (X_{23}+X_{24}+X_{25}+X_{26})(X_{34}+X_{35}+X_{36})$ in order to eliminate all the X_{1j}. The remaining two lines already involve integrands in the z_1 independent basis, and the associated kinematics receive corrections

$$\begin{aligned} X_{23}X_{45}M_1M^i_{23}M^j_{45}M^k_6 \;\;\mapsto\;\; & X_{23}X_{45}\Big(M_1\, M^i_{23}\, M^j_{45}\, M^k_6 \;+\; M_{12}\, M^i_3\, M^j_{45}\, M^k_6 \\ & -\; M_{13}\, M^i_2\, M^j_{45}\, M^k_6 \;+\; M_{14}\, M^i_{23}\, M^j_5\, M^k_6 \;-\; M_{15}\, M^i_{23}\, M^j_4\, M^k_6 \\ & +\; M_{413}\, M^i_2\, M^j_5\, M^k_6 \;+\; M_{512}\, M^i_3\, M^j_4\, M^k_6 \\ & -\; M_{412}\, M^i_3\, M^j_5\, M^k_6 \;-\; M_{513}\, M^i_2\, M^j_4\, M^k_6 \Big) \\ =\;\; & X_{23}X_{45}C_{1,23,45,6} \end{aligned} \tag{5.29}$$

due to the $X_{23}X_{45}$ on the right hand side of integration by parts formulae. By carefully gathering all $X_{23}X_{45}$ corrections, the superfield expressions can be seen to build up the full-fledged $C_{1,23,45,6}$. So the net effect of integrating z_1-dependent X_{pq} by parts is the replacement $M_1M^i_{23}M^j_{45}M^k_6 \mapsto C_{1,23,45,6}$ and $M_1M^i_{234}M^j_5M^k_6 \mapsto C_{1,234,5,6}$:

$$\begin{aligned} K_6 = {} & X_{23}(X_{24}+X_{34})\langle C_{1,234,5,6}\rangle \;+\; X_{24}(X_{23}+X_{43})\langle C_{1,243,5,6}\rangle \\ & +\; X_{23}(X_{25}+X_{35})\langle C_{1,235,4,6}\rangle \;+\; X_{25}(X_{23}+X_{53})\langle C_{1,253,4,6}\rangle \\ & +\; X_{23}(X_{26}+X_{36})\langle C_{1,236,4,5}\rangle \;+\; X_{26}(X_{23}+X_{63})\langle C_{1,263,4,6}\rangle \\ & +\; X_{24}(X_{25}+X_{45})\langle C_{1,245,3,6}\rangle \;+\; X_{25}(X_{24}+X_{54})\langle C_{1,254,3,6}\rangle \\ & +\; X_{24}(X_{26}+X_{46})\langle C_{1,246,3,5}\rangle \;+\; X_{26}(X_{24}+X_{64})\langle C_{1,264,3,5}\rangle \\ & +\; X_{25}(X_{26}+X_{56})\langle C_{1,256,3,4}\rangle \;+\; X_{26}(X_{25}+X_{65})\langle C_{1,265,3,4}\rangle \\ & +\; X_{34}(X_{35}+X_{45})\langle C_{1,345,2,6}\rangle \;+\; X_{35}(X_{34}+X_{54})\langle C_{1,354,2,6}\rangle \\ & +\; X_{34}(X_{36}+X_{46})\langle C_{1,346,2,5}\rangle \;+\; X_{36}(X_{34}+X_{64})\langle C_{1,364,2,5}\rangle \\ & +\; X_{35}(X_{36}+X_{56})\langle C_{1,356,2,4}\rangle \;+\; X_{36}(X_{35}+X_{65})\langle C_{1,365,2,4}\rangle \\ & +\; X_{45}(X_{46}+X_{56})\langle C_{1,456,2,3}\rangle \;+\; X_{46}(X_{45}+X_{65})\langle C_{1,465,2,3}\rangle \\ & +\; X_{23}\,X_{45}\,\langle C_{1,23,45,6}\rangle \;+\; X_{23}\,X_{46}\,\langle C_{1,23,46,5}\rangle \;+\; X_{23}\,X_{56}\,\langle C_{1,23,56,4}\rangle \end{aligned} \tag{5.30}$$

JHEP08(2014)099

$$
\begin{aligned}
&+\ X_{24}\,X_{35}\,\langle C_{1,24,35,6}\rangle\ +\ X_{24}\,X_{36}\,\langle C_{1,24,36,5}\rangle\ +\ X_{24}\,X_{56}\,\langle C_{1,24,56,3}\rangle\\
&+\ X_{25}\,X_{34}\,\langle C_{1,25,34,6}\rangle\ +\ X_{25}\,X_{36}\,\langle C_{1,25,36,4}\rangle\ +\ X_{25}\,X_{46}\,\langle C_{1,25,46,3}\rangle\\
&+\ X_{26}\,X_{34}\,\langle C_{1,26,34,5}\rangle\ +\ X_{26}\,X_{35}\,\langle C_{1,26,35,4}\rangle\ +\ X_{26}\,X_{45}\,\langle C_{1,26,45,3}\rangle\\
&+\ X_{34}\,X_{56}\,\langle C_{1,34,56,2}\rangle\ +\ X_{35}\,X_{46}\,\langle C_{1,35,46,2}\rangle\ +\ X_{36}\,X_{45}\,\langle C_{1,36,45,2}\rangle
\end{aligned}
$$

The above patterns lead to a seven-point kinematic factor given by

$$
\begin{aligned}
K_7 = 15\text{ terms }&\big[\,X_{23}\,(X_{24}+X_{34})\,(X_{25}+X_{35}+X_{45})\,\langle C_{1,2345,6,7}\rangle\ +\ \mathrm{sym}(345)\,\big]\\
+\,60\text{ terms }&\big[\,X_{23}\,(X_{24}+X_{34})\,\langle C_{1,234,56,7}\rangle\ +\ (3\leftrightarrow 4)\,\big]\,X_{56}\\
+\,15\text{ terms }&\big[X_{23}\,X_{45}\,X_{67}\,\langle C_{1,23,45,67}\rangle\,\big].
\end{aligned}
\tag{5.31}
$$

In order to make the permutations in (5.31) more precise and to compactly write down its n-point generalization, we shall introduce some notation that facilitates the bookkeeping of the S_3^{n-1} terms in K_n.

5.4 The closed-form n-point kinematic factor

We have argued in subsection 4.5 that the symmetries (4.36) of the BRST invariants yield a basis with S_3^{n-1} elements under relations with constant coefficients. It became evident from the examples (5.28), (5.30) and (5.31) that each independent $C_{1,\ldots}$ occurs in K_n where (5.21) determines the associated worldsheet function to be

$$
\langle C_{1,23\ldots p,p+1\ldots q,q+1\ldots n}\rangle\ \ \leftrightarrow\ \ \left(\prod_{k=3}^{p}\sum_{m=2}^{k-1}X_{mk}\right)\left(\prod_{k=p+2}^{q}\sum_{m=p+1}^{k-1}X_{mk}\right)\left(\prod_{k=q+2}^{n}\sum_{m=q+1}^{k-1}X_{mk}\right).
$$

Writing down the kinematic factor K_n in a closed form for general multiplicity n is a matter of notation. That is why we shall now introduce a set $\mathcal{S}_3^k$ with S_3^k elements which takes care of the $C_{1,\ldots}$ bookkeeping. It compasses all the partitions of k elements $12\ldots k$ into three indistinguishable cycles, say $(a_1\ldots a_p)$, $(b_1\ldots b_q)$, $(c_1\ldots c_r)$, where $p+q+r=k$ and none of the cycles remains empty, i.e. $p,q,r\neq 0$. For given sets $\{a_1\ldots a_p\}$, $\{b_1\ldots b_q\}$ and $\{c_1\ldots c_r\}$, only cyclically inequivalent configurations are considered as distinct $\mathcal{S}_3^k$ elements. Fixing the first entry a_1,b_1,c_1 of each cycle is one convenient way to implement this, we are then left with permutations

$$
(a_1\sigma(a_2\ldots a_p)),\ (b_1\pi(b_2\ldots b_q)),\ (c_1\rho(c_2\ldots c_r)),\quad \sigma\in S_{p-1},\quad \pi\in S_{q-1},\quad \rho\in S_{r-1}
$$

for a partition characterized by p,q,r. Of course, we have to avoid overcounting due to the indistinguishable cycles, i.e. $(2,3)$, $(4,5)$, $(\ldots)$ is identified with $(4,5)$, $(2,3)$, $(\ldots)$ in $\mathcal{S}_3^k$. A formal way to summarize these properties of $\mathcal{S}_3^k$ is

$$
\begin{aligned}
\mathcal{S}_3^k &= \bigcup_{\substack{p\geq q\geq r\geq 1\\ p+q+r=k}}\frac{\Xi_{p,q,r}(S_k)}{Z_p\times Z_q\times Z_r\times S_{\nu(p,q,r)}}\\
\Xi_{p,q,r}(12\ldots k) &= (12\ldots p)\times(p+1\ldots p+q)\times(p+q+1\ldots k)\\
\nu(p,q,r) &= 1+\delta_{p,q}+\delta_{q,r}\ .
\end{aligned}
\tag{5.32}
$$

JHEP08(2014)099

The map $\Xi_{p,q,r}$ cuts a given S_k permutation of $(12\ldots k)$ into three tuples $(12\ldots p)$, $(p+1\ldots p+q)$ and $(p+q+1\ldots k)$ of cardinality, p, q and r, respectively. Each of them is modded out by the corresponding cyclic group Z_p, Z_q, Z_r, and in case of coinciding cardinalities ($p=q$ or $q=r$ or both), we divide by permutations $S_{\nu(p,q,r)}$ of these tuples of equal size. Indeed, we can check that the number of elements in the individual (p,q,r) contributions to (5.32),

$$\left|\frac{S_k}{Z_p\times Z_q\times Z_r\times S_{\nu(p,q,r)}}\right| = \frac{k!}{pqr\cdot\nu(p,q,r)!}$$

reproduce the entries of table 1.

The structure of the n-point kinematic factor is described by

$$K_n = \sum_{p,q}\Big\{\langle C_{1,23\ldots p,p+1\ldots q,q+1\ldots n}\rangle \left(\prod_{k=3}^{p}\sum_{m=2}^{k-1} X_{mk}\right)\left(\prod_{k=p+2}^{q}\sum_{m=p+1}^{k-1} X_{mk}\right)\left(\prod_{k=q+2}^{n}\sum_{m=q+1}^{k-1} X_{mk}\right) + \mathrm{sym}(34\ldots p)+\mathrm{sym}(p+2,\ldots q)+\mathrm{sym}(q+2,\ldots n)+\text{permutations}\Big\}\,. \tag{5.33}$$

The definition (5.32) of $\mathcal{S}_3^k$ allows to make the permutations involved very precise:

$$\begin{aligned} K_n = & \sum_{\sigma\times\pi\times\rho\in\mathcal{S}_3^{n-1}} \langle C_{1,\sigma(23\ldots p),\pi(p+1\ldots q),\rho(q+1\ldots n)}\rangle \\ & \sigma\left(\prod_{k=3}^{p}\sum_{m=2}^{k-1} X_{mk}\right)\ \pi\left(\prod_{k=p+2}^{q}\sum_{m=p+1}^{k-1} X_{mk}\right)\ \rho\left(\prod_{k=q+2}^{n}\sum_{m=q+1}^{k-1} X_{mk}\right). \end{aligned} \tag{5.34}$$

The variables p,q are related to the cardinality of the permutations σ,π,ρ via $p=|\sigma|+1$ and $q-p=|\pi|$ and should not be confused with the summation variables in (5.32).

We shall conclude this section with a comment on the rigid $s_{ij}\eta_{ij}=X_{ij}$ combinations in the worldsheet integrand (5.34). The $z_i\to z_j$ singularities from $\eta_{ij}=z_{ij}^{-1}+\mathcal{O}(z_{ij})$ in connection with the Koba Nielsen factor (5.24) give rise to kinematic poles in the corresponding Mandelstam variable, at least for some choices of the integration region. The connection between worldsheet poles and massless propagators was thoroughly explored at tree-level [7], and since the $z_i\to z_j$ singularities are local effects on the worldsheet regardless of its global properties, we expect the pole analysis to carry over to higher genus.

The fact that short distance singularities on the worldsheet always occur in the combination $X_{ij}=s_{ij}\eta_{ij}$, any potential kinematic pole is immediately smoothed out by the Mandelstam numerator s_{ij}. That is why the z_i integrals do not introduce any poles in kinematic invariants,[9] i.e. that all massless open string propagators enter through the BRST invariants $C_{1,\ldots}$. However, this does not rule out branch cut singularities in s_{ij} as they are expected from the polylogarithms in field theory loop amplitudes. Systematic study of the non-analytic momentum dependence is a rewarding challenge which we leave for future work.

[9]This does not exclude massless poles from the modular t integration due to closed string exchange in non-planar cylinder diagrams [21, 22].

JHEP08(2014)099

6 One-loop kinematic factors built from tree-level data

In this section, we will show that the BRST invariant constituents $C_{1,\dots}$ of the one-loop kinematic factor K_n can be expanded in terms of SYM tree amplitudes. More precisely, these kinematic building blocks for one-loop amplitudes are local linear combinations of the α'^2 correction $\mathcal{A}^{F^4}$ to color-ordered superstring tree amplitudes, defined by

$$\mathcal{A}^{\text{tree}}(1,2,\dots,n;\alpha') = \mathcal{A}^{\text{YM}}(1,2,\dots,n) + \zeta(2)\alpha'^2\mathcal{A}^{F^4}(1,2,\dots,n) + \mathcal{O}(\alpha'^3)\ . \tag{6.1}$$

The notation $\mathcal{A}^{F^4}$ is motivated by the fact that the first string correction to (6.1) at order[10] α'^2 can be attributed to a supersymmetrized F^4 operator in the low energy effective action [64], see later remarks. Comparing with the central result of [6, 7]

$$\mathcal{A}^{\text{tree}}(1,2,\dots,n;\alpha') = \sum_{\sigma\in S_{n-3}} \mathcal{A}^{\text{YM}}(1,\sigma(2,\dots,n-2),n-1,n)F^{\sigma}(\alpha')$$

for the disk amplitude, one can identify the $\mathcal{O}(\alpha'^2)$ power of the functions F^σ as the expansion coefficients of $\mathcal{A}^{F^4}$ in terms of $(n-3)!$ field theory subamplitudes:

$$\zeta(2)\alpha'^2\mathcal{A}^{F^4}(1,2,\dots,n) = \sum_{\sigma\in S_{n-3}} F^{\sigma}(\alpha')\big|_{\alpha'^2}\mathcal{A}^{\text{YM}}(1,\sigma(2,\dots,n-2),n-1,n) \tag{6.2}$$

The first examples up to multiplicity $n=6$ read

$$\begin{aligned}
\mathcal{A}^{F^4}(1,2,3,4) &= s_{12}s_{23}\mathcal{A}^{\text{YM}}(1,2,3,4)\\
\mathcal{A}^{F^4}(1,2,3,4,5) &= (s_{12}s_{34}-s_{34}s_{45}-s_{12}s_{51})\mathcal{A}^{\text{YM}}(1,2,3,4,5)+s_{13}s_{24}\mathcal{A}^{\text{YM}}(1,3,2,4,5)\\
\mathcal{A}^{F^4}(1,2,3,4,5,6) &= -(s_{45}s_{56}+s_{12}s_{61}-s_{45}s_{123}-s_{12}s_{345}+s_{123}s_{345})\mathcal{A}^{\text{YM}}(1,2,3,4,5,6)\\
&\quad -s_{13}(s_{23}-s_{61}+s_{345})\mathcal{A}^{\text{YM}}(1,3,2,4,5,6)-s_{14}s_{25}\mathcal{A}^{\text{YM}}(1,4,3,2,5,6)\\
&\quad +s_{14}s_{35}\mathcal{A}^{\text{YM}}(1,4,2,3,5,6)-s_{35}(s_{34}-s_{56}+s_{123})\mathcal{A}^{\text{YM}}(1,2,4,3,5,6)\\
&\quad +s_{13}s_{25}\mathcal{A}^{\text{YM}}(1,3,4,2,5,6),
\end{aligned} \tag{6.3}$$

and a $\mathcal{O}(\alpha'^3)$ momentum expansion for the $n=7$ functions F^σ — i.e. the defining data for $\mathcal{A}^{F^4}(1,2,3,4,5,6,7)$ — can be found in the appendix of [7].

As we will show, our BRST invariants governing the one-loop kinematics

$$\langle C_{1,\dots}\rangle = \sum_{\rho}\mathcal{A}^{F^4}(1,\rho(2,3,\dots,n)) = \sum_{\pi} p_\pi(s_{ij})\mathcal{A}^{\text{YM}}(1,\pi(2,3,\dots,n)) \tag{6.4}$$

are linear combinations of SYM trees, accompanied by fine-tuned quadratic polynomials $p_\pi(s_{ij})$ in Mandelstam variables. The summation ranges for the S_{n-1} permutations ρ,π will be made precise soon.

Since the three-point tree does not receive any α' corrections, higher-point disk amplitudes do not factorize on exclusively cubic vertices. Hence, the role of the Mandelstam bilinears $p_\pi(s_{ij})$ lies in avoiding $n-3$ simultaneous poles in any $\mathcal{A}^{F^4}$. One can attribute these α'^2

[10] Higher dimensional operators such as $D^{2n}F^4$ and F^{4+n} with $n\geq 1$ contribute to (6.1) at orders α'^{2+n} and are not reflected in $\langle C_{1,\dots}\rangle$ which carries the same mass dimension as $\mathcal{A}^{F^4}$.

JHEP08(2014)099

corrections to a quartic contact interaction $\sim \mathrm{Tr}\{F^4\}$ (formed by the non-linearized gluon field strength F) in the low energy effective action [64] (which explains the terminology $\mathcal{A}^{F^4}$). The diagrams of $\mathcal{A}^{F^4}$ having one quartic vertex and cubic SYM vertices otherwise require $n-4$ propagators (instead of the $n-3$ propagators in cubic $\mathcal{A}^{\rm YM}$ diagrams).

In fact, the appearance of the tree-level kinematics $\mathcal{A}^{F^4}$ due to the (supersymmetric completion of the) operator $\sim \mathrm{Tr}\{F^4\}$ in one-loop amplitudes can be explained by supersymmetry: naive power counting shows that BRST invariants $C_{1,\ldots}$ are generated by a term of mass dimension eight in the effective action. The vertex $\sim \mathrm{Tr}\{F^4\}$ is the unique mass dimension eight operator compatible with 16 supercharges, i.e. $N=1$ supersymmetry in ten spacetime dimensions or $N=4$ supersymmetry in four dimensions. Cubic operators of type $\sim \mathrm{Tr}\{D^{2k}F^3\}$ can be ruled out since none of them is supersymmetrizable. That is why one-loop kinematics in maximally supersymmetric theories have no other choice than reproducing the $\mathcal{A}^{F^4}$ which have firstly been observed at tree-level.

The organization of this section proceeds as follows: we will first develop a pure spinor superspace representation for $\mathcal{A}^{F^4}$ in terms of quadruple Berends-Giele currents $M_{d_1\ldots d_s}M^i_{a_1\ldots a_p}M^j_{b_1\ldots b_q}M^k_{c_1\ldots c_r}$ using their diagrammatic interpretation from figure figure 5. The central box in these diagrams is then identified with the aforementioned quartic contact interaction vertex $\sim \mathrm{Tr}\{F^4\}$. We exploit the Berends-Giele representation to identify the $\mathcal{A}^{F^4}$ as linear combinations of the one-loop BRST invariants $C_{1,\ldots}$. Finally, the S_3^{n-1} basis for $C_{1,\ldots}$ can be used to explain amplitude relations between $\mathcal{A}^{F^4}$ permutations and (closely related) finite one-loop amplitudes in pure (non-supersymmetric) Yang-Mills theory.

6.1 Diagrammatic expansion of tree-level α'^2 corrections

Following the ideas of [27], a method which associates pure spinor building blocks to cubic tree diagrams of SYM amplitudes in $D=10$ was reviewed in section 2 on the basis of [8, 17, 18]. The pure spinor superfield method of [8] rests on two basic assumptions:

1. the kinematic numerator of a cubic graph can only contain BRST building blocks whose Q variation cancels one of the kinematic poles

2. the sum of the expressions associated to all cubic graphs must be in the pure spinor BRST cohomology.

Now we are interested in an analogous diagrammatic method for constructing the tree-level α'^2 corrections and relating them to one-loop kinematic structures.

At n-points, $\mathcal{A}^{F^4}(1,2,\ldots,n)$ has $n-4$ simultaneous poles corresponding to diagrams with $n-4$ cubic vertices and one quartic vertex. Since we are using the same superspace ingredients $M_{d_1\ldots d_s}M^i_{a_1\ldots a_p}M^j_{b_1\ldots b_q}M^k_{c_1\ldots c_r}$ present in one-loop BRST invariants, the box notation introduced in subsections 4.2 and 4.4 will be kept and can be identified with the tree-level quartic vertex[11] $\sim \mathrm{Tr}\{F^4\}$. The unified diagrammatic language for both α'^2

[11] Even though the diagrammatic rules in this subsection might suggest an association of the kinematic factors $\mathcal{A}^{F^4}$ with box integrals in the field theory limit, they will also find appearance along with pentagons and higher n-gons. The $\alpha'\to 0$ limit of the worldsheet integrals (to be analyzed in later work) will determine the kinematic coefficient of higher n-gons in terms of $\mathcal{A}^{F^4}$ — at least up to anomalous terms.

JHEP08(2014)099

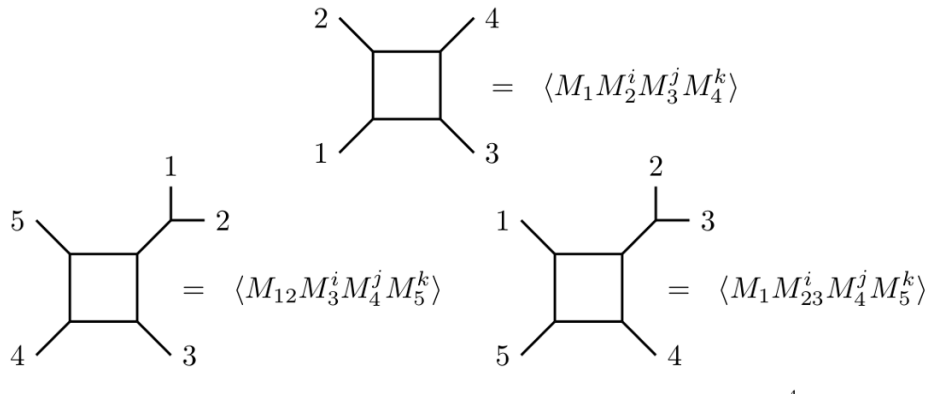

Figure 6. The building block prescription for the four- and five-point $\mathcal{A}^{F^4}$ diagrams. The rule is that the Berends-Giele current with leg one is always to the left, carries no i, j, k labels and the combination of superfields must contain the same kinematic poles of the graph.

corrected trees and loop-level kinematic factors emphasizes that they can be represented by the same class of subamplitudes $\mathcal{A}^{F^4}$. As mentioned before, this can be traced back to the uniqueness of $N = 1$ supersymmetric dimension eight operators in $D = 10$.

The four- and five-point diagrams associated with the tree-level α'^2 correction are depicted in figure 6, together with their pure spinor superfield mapping. The expression

$$\mathcal{A}^{F^4}(1,2,3,4) = \langle M_1 M_2^i M_3^j M_4^k \rangle \tag{6.5}$$

correctly reflects the absence of poles in $\mathcal{A}^{F^4}(1,2,3,4)$ and is BRST closed.

The five-point $\mathcal{A}^{F^4}(1,2,3,4,5)$ has two kinds of Berends-Giele-constituents. They are characterized by the position of the leg with label one — it can either enter through the cubic vertex ($\to M_{1j}$) or as a standalone corner ($\to M_1$) of the box. The superfield mapping is slightly different for each possibility, and the rule is that leg number one is never associated with loop-specific Berends-Giele currents $M^{i,j,k}_{\cdots}$. The dictionary of figure 6 leads to the following Q closed expression

$$\begin{aligned}\mathcal{A}^{F^4}(1,2,3,4,5) = {} & \langle M_{12} M_3^i M_4^j M_5^k \rangle + \langle M_1 M_{23}^i M_4^j M_5^k \rangle \\ & + \langle M_1 M_2^i M_{34}^j M_5^k \rangle + \langle M_1 M_2^i M_3^j M_{45}^k \rangle + \langle M_{51} M_2^i M_3^j M_4^k \rangle \,.\end{aligned} \tag{6.6}$$

In the four-point case, it was shown in [57] on superfield level that $\langle M_1 M_2^i M_3^j M_4^k \rangle$ agrees with the SYM tree representation $\mathcal{A}^{F^4}(1,2,3,4) = s_{12}s_{23}\mathcal{A}^{\rm YM}(1,2,3,4)$. This requires the pure spinor superspace expression (2.20) for the latter,

$$\langle M_1 M_2^i M_3^j M_4^k \rangle = s_{23}\langle T_{12} V_3 V_4 \rangle + s_{12}\langle V_1 T_{23} V_4 \rangle = s_{12}s_{23}\mathcal{A}^{\rm YM}(1,2,3,4). \tag{6.7}$$

However, we could not find a superspace proof for (6.6) to agree with the $\mathcal{A}^{\rm YM}$ combination $(s_{12}s_{34} - s_{34}s_{45} - s_{12}s_{51})\mathcal{A}^{\rm YM}(1,2,3,4,5) + s_{13}s_{24}\mathcal{A}^{\rm YM}(1,3,2,4,5)$ required by (6.3). Instead, we have checked that this combination of five gluon trees matches the bosonic terms of the component expansion of (6.6). The agreement of the gluonic components extends to the full supermultiplet because the $\langle \lambda^3 \theta^5 \rangle = 1$ prescription respects supersymmetry.

JHEP08(2014)099

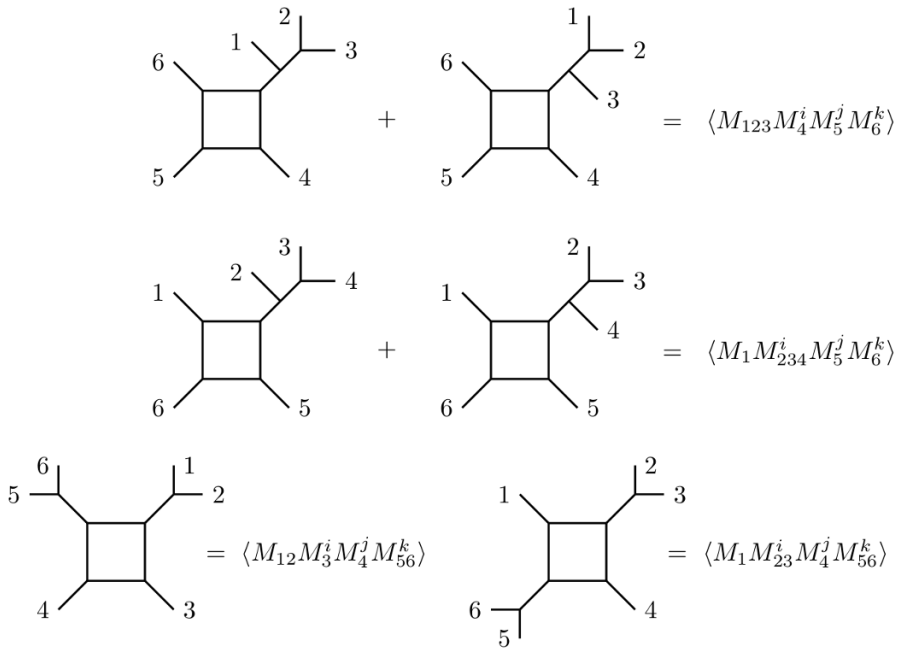

Figure 7. Pure spinor diagrammatic rules for the six-point α'^2 correction $\mathcal{A}^{F^4}$. The leg number one is never associated with a loop-specific Berends-Giele current $M^{i,j,k}$ and the labels in the superfields are arranged according the order in which they appear in the diagrams.

For six-points the story is the same, and the mappings between diagrams and superfields depend on the position of leg number one.

Following the mapping rules depicted in figure 7 the 21 graphs which compose the six-point $\mathcal{A}^{F^4}$ are represented by the following 15 terms in pure spinor superspace,

$$\begin{aligned}\mathcal{A}^{F^4}(1,2,\ldots,6) = {} & \langle M_{123}M_4^iM_5^jM_6^k\rangle + \langle M_{612}M_3^iM_4^jM_5^k\rangle + \langle M_{561}M_2^iM_3^jM_4^k\rangle \\ & + \langle M_1M_{234}^iM_5^jM_6^k\rangle + \langle M_1M_2^iM_{345}^jM_6^k\rangle + \langle M_1M_2^iM_3^jM_{456}^k\rangle \\ & + \langle M_{12}M_{34}^iM_5^jM_6^k\rangle + \langle M_{61}M_{23}^iM_4^jM_5^k\rangle + \langle M_{12}M_3^iM_4^jM_{56}^k\rangle \\ & + \langle M_{61}M_2^iM_3^jM_{45}^k\rangle + \langle M_1M_2^iM_{34}^jM_{56}^k\rangle + \langle M_1M_{23}^iM_{45}^jM_6^k\rangle \\ & + \langle M_{12}M_3^iM_{45}^jM_6^k\rangle + \langle M_{61}M_2^iM_{34}^jM_5^k\rangle + \langle M_1M_{23}^iM_4^jM_{56}^k\rangle,\end{aligned} \tag{6.8}$$

which one can check to be BRST-closed using the formulæ given in the previous section. The first six terms in (6.8) altogether describe twelve graphs whereas each of the last nine terms describe a single graph.

Let us state the general rule for proceeding beyond six-points: our superspace proposal for $\mathcal{A}^{F^4}(\ldots)$ encompasses all distinct planar diagrams (with unit relative weight) which are

1. made of one totally symmetric quartic vertex and otherwise antisymmetric cubic vertices from SYM

2. compatible with the cyclic ordering of the external legs in $\mathcal{A}^{F^4}(\ldots)$.

JHEP08(2014)099

The sum of their $M_{\dots}M^i_{\dots}M^j_{\dots}M^k_{\dots}$ superspace representatives can be checked to enjoy BRST invariance. Up to multiplicity nine, we have

$$\begin{aligned}\mathcal{A}^{F^4}(1,2,\dots,7) &= \langle M_{1234}M^i_5M^j_6M^k_7\rangle + \langle M_{123}M^i_{45}M^j_6M^k_7\rangle + \langle M_{123}M^i_4M^j_{56}M^k_7\rangle \\ &\quad + \langle M_{123}M^i_4M^j_5M^k_{67}\rangle + \langle M_{12}M^i_{34}M^j_{56}M^k_7\rangle + \text{cyc}(1234567) \\ &= 35 \text{ terms and } 84 \text{ diagrams} \end{aligned} \tag{6.9}$$

$$\begin{aligned}\mathcal{A}^{F^4}(1,2,\dots,8) &= \langle M_{12345}M^i_6M^j_7M^k_8\rangle + \langle M_{123}M^i_{456}M^j_7M^k_8\rangle \\ &\quad + \langle M_{1234}\big(M^i_{56}M^j_7M^k_8 + M^i_5M^j_{67}M^k_8 + M^i_5M^j_6M^k_{78}\big)\rangle \\ &\quad + \langle M_{123}\big(M^i_{45}M^j_{67}M^k_8 + M^i_{45}M^j_6M^k_{78} + M^i_4M^j_{56}M^k_{78}\big)\rangle + \text{cyc}(12345678) \\ &\quad + \langle M_{123}M^i_4M^j_{567}M^k_8\rangle + \text{cyc}(1234) \\ &\quad + \langle M_{12}M^i_{34}M^j_{56}M^k_{78}\rangle + \langle M_{81}M^i_{23}M^j_{45}M^k_{67}\rangle \\ &= 70 \text{ terms and } 330 \text{ diagrams} \end{aligned} \tag{6.10}$$

$$\begin{aligned}\mathcal{A}^{F^4}(1,2,\dots,9) &= \langle M_{123456}M^i_7M^j_8M^k_9\rangle + \langle M_{123}M^i_{45}M^j_{67}M^k_{89}\rangle \\ &\quad + \langle M_{12345}\big(M^i_{67}M^j_8M^k_9 + M^i_6M^j_{78}M^k_9 + M^i_6M^j_7M^k_{89}\big)\rangle \\ &\quad + \langle M_{1234}\big(M^i_{56}M^j_{78}M^k_9 + M^i_{56}M^j_7M^k_{89} + M^i_5M^j_{67}M^k_{89}\big)\rangle \\ &\quad + \langle M_{1234}\big(M^i_{567}M^j_8M^k_9 + M^i_5M^j_{678}M^k_9 + M^i_5M^j_6M^k_{789}\big)\rangle \\ &\quad + \langle M_{123}\big(M^i_{456}M^j_{78}M^k_9 + M^i_{45}M^j_{678}M^k_9 + M^i_{456}M^j_7M^k_{89}\big)\rangle \\ &\quad + \text{cyc}(123456789) \\ &= 126 \text{ terms and } 1287 \text{ diagrams} \end{aligned} \tag{6.11}$$

Summing over cyclic permutations in the specified labels slightly abuses notation in view of the rule that leg number one is always attached to the tree-level current $M_{\dots}$ rather than to $M^{i,j,k}_{\dots}$. For example, the cyclic orbit of $M_{123}M^i_4M^j_5M^k_6$ shall be understood as

$$\begin{aligned}M_{123}M^i_4M^j_5M^k_6 + \text{cyc}(123456) = M_{123}M^i_4M^j_5M^k_6 + M_1M^i_{234}M^j_5M^k_6 + M_1M^i_2M^j_{345}M^k_6 \\ + M_1M^i_2M^j_3M^k_{456} + M_{612}M^i_3M^j_4M^k_5 + M_{561}M^i_2M^j_3M^k_4.\end{aligned}$$

Using this refined cyclic summation, one can verify BRST closure of the above expressions as well as the correct diagrammatic content to represent the α'^2 corrections $\mathcal{A}^{F^4}$ to tree amplitudes. Moreover, as a sufficient condition, we have explicitly checked their agreement up to $n = 6$ by computing the bosonic component expansions [14] and comparing with (6.3). It is highly plausible that the (well-tested) experimental rule

BRST-closed objects with the same kinematic pole structure are proportional

persists for $n \geq 7$ legs.

The above expressions for $\mathcal{A}^{F^4}(1,2,\dots,n)$ are not manifestly cyclic invariant in $(1,2,\dots,n)$ because the leg number one is treated differently. This is an artifact of the one-loop prescription from section 3 which associates only leg number one with the unintegrated vertex operator V^1. But it can be shown that the difference to another choice of $V^{i\neq 1}$ is BRST-exact and therefore zero,

$$\mathcal{A}^{F^4}(1,2,\dots,n) - \mathcal{A}^{F^4}(2,3,\dots,n,1) = \langle Q\,\mathcal{X}_n\rangle = 0, \tag{6.12}$$

JHEP08(2014)099

for example,[12]

$$\begin{aligned}
\mathcal{X}_4 &= M^i_{12}M^j_3M^k_4\\
\mathcal{X}_5 &= M^i_{123}M^j_4M^k_5 + M^i_{512}M^j_3M^k_4 + M^i_{12}(M^j_3M^k_{45} + M^j_{34}M^k_5) \qquad (6.15)\\
\mathcal{X}_6 &= M^i_{1234}M^j_5M^k_6 + M^i_{6123}M^j_4M^k_5 + M^i_{5612}M^j_3M^k_4 + M^i_{123}(M^j_4M^k_{56} + M^j_{45}M^k_6)\\
&\quad + M^i_{612}(M^j_3M^k_{45} + M^j_{34}M^k_5) + M^i_{12}(M^j_3M^k_{456} + M^j_{34}M^k_{56} + M^j_{345}M^k_6). \qquad (6.16)
\end{aligned}$$

6.2 Tree-level α'^2 corrections versus one-loop kinematics

This subsection builds a bridge between tree-level α'^2 corrections $\mathcal{A}^{F^4}$ and the one-loop kinematics $C_{1,\ldots}$. Both of them have a superspace representation in terms of $M_{d_1\ldots d_s}M^i_{a_1\ldots a_p}M^j_{b_1\ldots b_q}M^k_{c_1\ldots c_r}$ — see the previous subsection for $\mathcal{A}^{F^4}$ and (4.28)–(4.34) for $C_{1,\ldots}$. Using the symmetry properties (4.35) of Berends-Giele currents, we find

$$\mathcal{A}^{F^4}(1,2,\ldots,n) = \sum_{2\le p<q\le n-1} \langle C_{1,23\ldots p,p+1\ldots q,q+1\ldots n}\rangle, \qquad (6.17)$$

where legs $23\ldots n$ are distributed in all possible ways among the three slots of the BRST invariants $C_{1,\ldots}$ which preserve their order. This leads to $(n-2)(n-3)/2$ terms in the $C_{1,\ldots}$ expansion of the color-ordered $\mathcal{A}^{F^4}(1,2,\ldots,n)$, let us display examples up to multiplicity $n=8$:

$$\begin{aligned}
\mathcal{A}^{F^4}(1,2,\ldots,4) &= \langle C_{1,2,3,4}\rangle \qquad (6.18)\\
\mathcal{A}^{F^4}(1,2,\ldots,5) &= \langle C_{1,23,4,5}\rangle + \langle C_{1,34,2,5}\rangle + \langle C_{1,45,2,3}\rangle \qquad (6.19)\\
\mathcal{A}^{F^4}(1,2,\ldots,6) &= \langle C_{1,234,5,6}\rangle + \langle C_{1,345,2,6}\rangle + \langle C_{1,456,2,3}\rangle + \langle C_{1,23,45,6}\rangle\\
&\quad + \langle C_{1,23,56,4}\rangle + \langle C_{1,34,56,2}\rangle \qquad (6.20)\\
\mathcal{A}^{F^4}(1,2,\ldots,7) &= \langle C_{1,2345,6,7}\rangle + \langle C_{1,3456,2,7}\rangle + \langle C_{1,4567,2,3}\rangle + \langle C_{1,234,56,7}\rangle\\
&\quad + \langle C_{1,234,67,5}\rangle + \langle C_{1,23,456,7}\rangle + \langle C_{1,23,567,4}\rangle + \langle C_{1,345,67,2}\rangle\\
&\quad + \langle C_{1,34,567,2}\rangle + \langle C_{1,23,45,67}\rangle \qquad (6.21)\\
\mathcal{A}^{F^4}(1,2,\ldots,8) &= \langle C_{1,23456,7,8}\rangle + \langle C_{1,34567,2,8}\rangle + \langle C_{1,45678,2,3}\rangle + \langle C_{1,2345,67,8}\rangle\\
&\quad + \langle C_{1,2345,78,6}\rangle + \langle C_{1,234,567,8}\rangle + \langle C_{1,234,678,5}\rangle + \langle C_{1,23,4567,8}\rangle\\
&\quad + \langle C_{1,23,5678,4}\rangle + \langle C_{1,3456,78,2}\rangle + \langle C_{1,345,678,2}\rangle + \langle C_{1,34,5678,2}\rangle\\
&\quad + \langle C_{1,23,45,678}\rangle + \langle C_{1,23,456,78}\rangle + \langle C_{1,234,56,78}\rangle \qquad (6.22)
\end{aligned}$$

[12]The general formula for $\mathcal{X}_n$ can be conveniently written using the definition

$$E^{ij}_{12\ldots p} = \sum_{k=1}^{p-1} M^i_{12\ldots k}M^j_{k+1\ldots p} \qquad (6.13)$$

as

$$\mathcal{X}_n = \sum_{p=2}^{n-2} M^i_{12\ldots p}E^{jk}_{p+1\ldots n} + \mathrm{tcyc}(12\ldots n) - \sum_{p=3}^{n-1} E^{ij}_{23\ldots p}M^k_{p+1,\ldots n,1} \qquad (6.14)$$

and $\mathrm{tcyc}(12\ldots n)$ means the *truncated* cyclic permutations of the enclosed indices. It is defined such that the permutations which lead to the leg number one not being in the "first" M^i are dropped. For example, $M^i_{123}M^j_4M^k_5 + \mathrm{tcyc}(12345) = M^i_{123}M^j_4M^k_5 + M^i_{512}M^j_3M^k_4 + M^i_{451}M^j_2M^k_3$.

JHEP08(2014)099

We will argue in the next subsection that the representation (6.17) of $\mathcal{A}^{F^4}$ in terms of $C_{1,\ldots}$ is invertible, i.e. that one can express any individual BRST invariant $C_{1,\ldots}$ in terms of $\mathcal{A}^{F^4}$ permutations with rational coefficients. As promised in (6.4), this implies that all kinematic ingredients $C_{1,\ldots}$ of one-loop amplitudes can be written in terms of tree-level kinematics $\mathcal{A}^{F^4}$, and the latter can in principle be expressed through $\mathcal{A}^{\rm YM}$ permutations (with bilinear Mandelstam coefficients). The reduction of five-point $C_{1,ij,k,l}$ to SYM trees proceeds as follows,

$$\begin{aligned}\langle C_{1,24,3,5}\rangle =& \mathcal{A}^{F^4}(1,3,2,4,5) - \mathcal{A}^{F^4}(1,4,3,2,5) - \mathcal{A}^{F^4}(1,3,4,5,2).\\ =& \mathcal{A}^{\rm YM}(1,2,3,4,5)s_{34}\left[s_{12} - s_{45} - \frac{s_{12}s_{23}}{s_{14}}\right]\\ &+ \mathcal{A}^{\rm YM}(1,3,2,4,5)\left[(s_{23}+s_{34})(s_{12}-s_{45})+\frac{s_{23}}{s_{14}}(s_{12}s_{24}-s_{51}(s_{34}+s_{45}))\right],\end{aligned} \tag{6.23}$$

and we shall finally give a six-point example:

$$\begin{aligned}3\langle C_{1,345,2,6}\rangle =& \mathcal{A}^{F^4}(1,2,3,6,4,5) + \mathcal{A}^{F^4}(1,2,5,4,3,6) + \mathcal{A}^{F^4}(1,2,5,4,6,3)\\ &+ \mathcal{A}^{F^4}(1,2,6,3,4,5) + \mathcal{A}^{F^4}(1,3,2,4,5,6) + \mathcal{A}^{F^4}(1,3,2,4,6,5)\\ &- \mathcal{A}^{F^4}(1,3,2,5,4,6) + \mathcal{A}^{F^4}(1,3,4,2,5,6) + \mathcal{A}^{F^4}(1,3,4,2,6,5)\\ &+ \mathcal{A}^{F^4}(1,3,4,6,2,5) + \mathcal{A}^{F^4}(1,6,2,3,4,5) + \mathcal{A}^{F^4}(1,6,4,3,2,5)\end{aligned} \tag{6.24}$$

$$\begin{aligned}3\langle C_{1,23,45,6}\rangle =& -\mathcal{A}^{F^4}(1,2,3,4,5,6) + \mathcal{A}^{F^4}(1,2,3,5,4,6) + \mathcal{A}^{F^4}(1,2,3,6,4,5)\\ &- \mathcal{A}^{F^4}(1,2,3,6,5,4) + \mathcal{A}^{F^4}(1,2,5,3,4,6) - \mathcal{A}^{F^4}(1,2,5,6,4,3)\\ &+ \mathcal{A}^{F^4}(1,2,6,3,4,5) - \mathcal{A}^{F^4}(1,2,6,3,5,4) + \mathcal{A}^{F^4}(1,2,6,4,3,5)\\ &+ \mathcal{A}^{F^4}(1,2,6,4,5,3) - \mathcal{A}^{F^4}(1,2,6,5,3,4) - \mathcal{A}^{F^4}(1,3,2,5,4,6)\\ &+ \mathcal{A}^{F^4}(1,3,2,5,6,4) - 2\mathcal{A}^{F^4}(1,3,2,6,4,5) + 2\mathcal{A}^{F^4}(1,3,2,6,5,4)\\ &+ \mathcal{A}^{F^4}(1,3,4,5,2,6) + \mathcal{A}^{F^4}(1,6,5,2,3,4)\,.\end{aligned} \tag{6.25}$$

It is not difficult to verify the above relations using the explicit expressions (6.19) and (6.20) together with the symmetries obeyed by the invariants $C_{1,\ldots}$.

6.3 KK-like identities for $\mathcal{A}^{F^4}$ and finite QCD amplitudes

We have argued in subsection 4.5 that the symmetries (4.36) of the $C_{1,\ldots}$ align them into a S_3^{n-1}-dimensional basis under relations with rational coefficients. This subsection focuses on amplitude relations between $\mathcal{A}^{F^4}$ following from their expansion (6.17) in terms of $C_{1,\ldots}$. Cyclic symmetry and $(-1)^n$ parity leave $(n-1)!/2$ potentially independent $\mathcal{A}^{F^4}(1,2,\ldots,n)$ permutations, but since they are all linear combinations of S_3^{n-1} independent $C_{1,\ldots}$, there must be lots of identities with rational coefficients among them. Following the terminology of [31], we will refer to these relations as "Kleiss Kuijf-like" (KK-like). The first example is $\mathcal{A}^{F^4}(1,2,3,4) = \mathcal{A}^{F^4}(1,3,2,4)$ being totally symmetric. Examples at five-points are

$$0 = \mathcal{A}^{F^4}(1,2,3,4,5) - \mathcal{A}^{F^4}(1,4,2,3,5) - \mathcal{A}^{F^4}(1,2,4,5,3)$$

JHEP08(2014)099

$$+\mathcal{A}^{F^4}(1,2,4,3,5)-\mathcal{A}^{F^4}(1,3,2,4,5)-\mathcal{A}^{F^4}(1,2,3,5,4), \tag{6.26}$$

$$0=\mathcal{A}^{F^4}(1,2,3,4,5)+\mathrm{sym}(2,4,5), \tag{6.27}$$

they can be easily checked using $\mathcal{A}^{F^4}(1,2,3,4,5)=\langle C_{1,23,4,5}+C_{1,2,34,5}+C_{1,2,3,45}\rangle$.

The basis dimension S_3^{n-1} for $C_{1,\ldots}$ furnishes an upper bound on the number of independent $\mathcal{A}^{F^4}$ under KK-like relations (e.g. one has at most six independent $\mathcal{A}^{F^4}(1,\sigma(2,3,4,5))$ under (6.26) and (6.27)). If this bound is saturated, then the equation system

$$\mathcal{A}^{F^4}(1,\sigma(2,\ldots,n))=\sum_{2\leq\sigma(p)<\sigma(q)\leq n-1}\langle C_{1,\sigma(23\ldots p),\sigma(p+1\ldots q),\sigma(q+1\ldots n)}\rangle, \tag{6.28}$$

is invertible and we can solve it for $C_{1,\ldots}$ in terms of $\mathcal{A}^{F^4}$ permutations. We will now give an indirect argument that this is indeed the case.

Relations of type (6.26) and (6.27) have already been observed in [31] for finite one-loop amplitudes in four-dimensional pure Yang Mills theory involving gluons of positive helicity only. Using the all multiplicity formula from [55][13]

$$A^{(1)}_{n;1}(1^+,2^+,\ldots,n^+)=-\frac{i}{48\pi^2}\frac{\sum_{i<j<k<l}\langle ij\rangle[jk]\langle kl\rangle[li]}{\langle 12\rangle\langle 23\rangle\ldots\langle n1\rangle} \tag{6.29}$$

(with $\langle ij\rangle$ and $[ij]$ denoting products of the momentum spinors of gluons i and j) the authors of [31] derive amplitude relations between different $A^{(1)}_{n;1}$ permutations and also find the basis dimension S_3^{n-1} under KK-like relations. Moreover, they develop a diagrammatic method to handle the symmetries using graphs with one quartic vertex and otherwise cubic vertices. This strongly resembles our diagrammatic interpretation of one-loop building blocks $\langle T_{d_1\ldots d_s}T^i_{a_1\ldots a_p}T^j_{b_1\ldots b_q}T^k_{c_1\ldots c_r}\rangle$. Reference [73] puts the idea to derive relations between box coefficients from quartic expressions in Berends-Giele currents into a more general context.

The coincidence between α'^2 corrections to superstring tree amplitudes and four-dimensional pure Yang Mills amplitudes was firstly noticed in [53, 54]. The authors point out that the four-dimensional reduction of gluonic $\mathcal{A}^{F^4}$ amplitudes[14] in MHV helicity configurations are proportional to (6.29),

$$\mathcal{A}^{F^4}(1^-,2^-,3^+,4^+\ldots,n^+)\sim\langle 12\rangle^4\;A^{(1)}_{n;1}(1^+,2^+,\ldots,n^+) \tag{6.30}$$

up to the permutation-insensitive "MHV-dressing" $\langle 12\rangle^4$. This explains why four-dimensional MHV representatives of $\mathcal{A}^{F^4}$ fall into the same S_3^{n-1}-dimensional basis found in [31] under KK-like relations. In other words, the MHV components of the $\mathcal{A}^{F^4}$ saturate the upper bound of S_3^{n-1} basis amplitudes found through our reasoning above based on $C_{1,\ldots}$ expansions. Generalizations of four-dimensional MHV $\mathcal{A}^{F^4}$ to other helicities, to other

[13]The expression (6.29) for pure Yang Mills amplitudes $A^{(1)}_{n;1}$ was observed in [79] to agree with dimension-shifted one-loop amplitudes of $N=4$ SYM in $D\mapsto D+4$ dimensions.

[14]The $\mathcal{A}^{\mathrm{YM}}$ representation (6.2) of $\mathcal{A}^{F^4}$ is dimension-agnostic — the functional dependence of SYM trees on gluon polarization vectors is the same in any number of dimensions, and one can use spinor helicity variables and the Parke Taylor formula [56] in the four-dimensional MHV situation.

JHEP08(2014)099

supermultiplet members and to higher dimensions can only require a larger basis than the MHV specialization, but exceeding the S_3^{n-1} is incompatible with the upper bound found for ten-dimensional superamplitudes.[15] This completes our argument why (6.28) admits to express any BRST invariant $C_{1,\dots}$ as a linear combination of $\mathcal{A}^{F^4}$.

To conclude this section, let us display higher order examples for KK-like identities between $\mathcal{A}^{F^4}$. At six-points, for instance, one can check

$$\begin{aligned}0 = {}& \mathcal{A}^{F^4}(1,5,4,3,6,2) - \mathcal{A}^{F^4}(1,5,4,2,6,3) - \mathcal{A}^{F^4}(1,5,4,6,2,3) + \mathcal{A}^{F^4}(1,5,4,6,3,2)\\ & + \mathcal{A}^{F^4}(1,5,6,4,2,3) - \mathcal{A}^{F^4}(1,5,6,4,3,2) + \mathcal{A}^{F^4}(1,6,2,3,4,5) - \mathcal{A}^{F^4}(1,6,3,2,4,5)\\ & - \mathcal{A}^{F^4}(1,6,4,2,3,5) - \mathcal{A}^{F^4}(1,6,4,2,5,3) + \mathcal{A}^{F^4}(1,6,4,3,2,5) + \mathcal{A}^{F^4}(1,6,4,3,5,2)\\ & - \mathcal{A}^{F^4}(1,6,4,5,2,3) + \mathcal{A}^{F^4}(1,6,4,5,3,2) + \mathcal{A}^{F^4}(1,6,5,4,2,3) - \mathcal{A}^{F^4}(1,6,5,4,3,2).\end{aligned}$$

using (6.20), and a neat form for all-multiplicity relations is given in [31]:

$$0 = 2\mathcal{A}^{F^4}(1,2,\dots,n) - (-1)^n \sum_{\sigma\in \mathrm{OP}(\{4\}\cup\{\beta\})} \mathcal{A}^{F^4}(3,\{\sigma\},5) + \mathrm{sym}(45\dots n).$$

Similar to KK relations, the notation $\mathrm{OP}(\{4\}\cup\{\beta\})$ includes all ways to shuffle leg four into the set $\{\beta\} = \{2,1,n,n-1,\dots,6\}$ while preserving the order of the latter.

6.4 BCJ-like identities for $\mathcal{A}^{F^4}$

We always pointed out that the basis dimensions S_3^{n-1} for both $C_{1,\dots}$ and $\mathcal{A}^{F^4}$ only take relations with constant, rational coefficients into account which we call KK-like. So far, we completely neglected the fact that $\mathcal{A}^{F^4}$ decompose into $\mathcal{A}^{\mathrm{YM}}$ permutations (weighted by bilinears in s_{ij}) which are well-known to have a $(n-3)!$ basis under KK- and BCJ relations [27]. Starting from $(n=5)$-points, the $\mathcal{A}^{\mathrm{YM}}$ basis contains strictly less elements than the "KK-like" basis of $\mathcal{A}^{F^4}$ since $(n-3)! < S_3^{n-1}$ for $n \geq 5$. Hence, there must be extra relations with Mandelstam coefficients between $\mathcal{A}^{F^4}$ that are independent under KK-like relations.

At five-points, extra identities with bilinear coefficients in Mandelstam variables reduce the $\mathcal{A}^{F^4}$ or pure Yang-Mills amplitudes $A^{(1)}_{n;1}$ to two independent ones (in agreement with the $(n-3)!$ basis of $\mathcal{A}^{\mathrm{YM}}$). Examples on the $A^{(1)}_{n;1}$ side are shown in equation (5.2) of [31], we have checked that they are also satisfied by $\mathcal{A}^{F^4}$. However, the most compact relations we could find between five-point BRST invariants involve the $C_{1,\dots}$ rather than $\mathcal{A}^{F^4}$. Let $P_i = \sum_{j=1}^5 x_{ij}s_j$ denote linear polynomials in Mandelstam variables $s_i = s_{i,i+1}$ with constants x_{ij}, then the ansatz

$$\begin{aligned}&s_{23}P_1C_{1,23,4,5} + s_{24}P_2C_{1,24,3,5} + s_{25}P_3C_{1,25,3,4}\\ &+s_{34}P_4C_{1,34,2,5} + s_{35}P_5C_{1,35,2,4} + s_{45}P_6C_{1,45,2,3} = 0\end{aligned} \tag{6.31}$$

is sufficient to find a two-dimensional basis of BRST invariants. The ansatz (6.31) is motivated by the fact that the $\frac{1}{s_{23}}$ pole in $C_{1,23,4,5}$ does not appear in any other $C_{1,\dots}$, so

[15] This is not a strict proof that the non-MHV $\mathcal{A}^{F^4}$ obey the *same* $(n-1)!/2 - S_3^{n-1}$ KK-like relations as their MHV cousins and the pure Yang Mills amplitudes (6.29), but we take strong confidence from the fact that our $\mathcal{A}^{F^4}$ in their helicity-agnostic $C_{1,\dots}$ representation obey all the $A^{(1)}_{n;1}$ amplitude relations explicitly written in [31].

JHEP08(2014)099

it must be cancelled by a s_{23} prefactor for $C_{1,23,4,5}$. Plugging in the polynomials $P_i = \sum_{j=1}^{5} x_{ij} s_j$ and solving the system of equations which follow from a component evaluation of (6.31) using [14] lead to four independent quadratic relations between $C_{1,\ldots}$. As a result, we can express $\{C_{1,24,3,5}, C_{1,25,3,4}, C_{1,34,2,5}, C_{1,35,2,4}\}$ in terms of a $\{C_{1,23,4,5}, C_{1,45,2,3}\}$ basis

$$C_{1,ij,k,l} = \pm \frac{s_{1k}}{s_{ij}} \left(s_{1i}s_{1l} - s_{jl}s_{ij}\right) \frac{C_{1,23,4,5}}{(s_{23} - s_{45})s_{45}} \pm \frac{s_{1l}}{s_{ij}} \left(s_{1j}s_{1k} - s_{ik}s_{ij}\right) \frac{C_{1,45,2,3}}{(s_{45} - s_{23})s_{23}} \quad (6.32)$$

where the signs are given by $\{(+-), (++), (--), (-+)\}$, respectively.

7 Harmony between color, kinematics and worldsheet integrands

In this section, we will explore the common combinatorial structures that govern on the one hand the kinematic building blocks $C_{1,\ldots}$ of one-loop amplitudes and the corresponding worldsheet integrands $X_{ij} = s_{ij}\eta_{ij}$, on the other hand also the color factors from the α'^2 corrections to tree amplitudes. In all the three cases, the basis dimensions are given by the unsigned Stirling number S_3^{n-1}. It can be viewed as the one-loop analogue of the magic number $(n-3)!$ omnipresent in tree-level bases of worldsheet integrals as well as color-ordered string- and SYM amplitudes. This coincidence has led us to a harmonious duality between color-dressed tree amplitudes at order α'^2 and the integrand of one-loop amplitudes in open superstring theory.

In the open string sector, the color-dressed tree amplitude is given by

$$\mathcal{M}_n^{\text{tree}}(\alpha') = \sum_{\sigma \in S_{n-1}/Z_2} \overset{\leftrightarrow}{\text{Tr}}\left[T^{a_{\sigma(1)}} T^{a_{\sigma(2)}} \cdots T^{a_{\sigma(n-1)}} T^{a_n}\right] \mathcal{A}^{\text{tree}}(\sigma(1,2,\ldots,n-1), n; \alpha') \quad (7.1)$$

where the summation includes all cyclically inequivalent permutations of the labels modded out by the $(-1)^n$ parity of color-stripped n-point amplitudes. The T^{a_i} denote the Chan-Paton factors[16] in the fundamental representation of the gauge group, and parity weighting is represented as

$$\overset{\leftrightarrow}{\text{Tr}}\left[T^{a_1} T^{a_2} \cdots T^{a_n}\right] := \text{Tr}\left[T^{a_1} T^{a_2} \ldots T^{a_n} + (-1)^n T^{a_n} T^{a_{n-1}} \cdots T^{a_1}\right]. \quad (7.2)$$

A convenient basis for these parity weighted traces involves structure constants f^{abc} and symmetrized traces $d^{a_1 a_2 a_3 \ldots a_{2n}}$ of even rank only, the latter being defined as [32],

$$d^{a_1 a_2 \ldots a_{2n}} := \frac{1}{(2n-1)!} \sum_{\sigma \in S_{2n-1}} \text{Tr}\left[T^{a_{\sigma(1)}} \cdots T^{a_{\sigma(2n-1)}} T^{a_{2n}}\right]. \quad (7.3)$$

We will use shorthands $f^{123} \equiv f^{a_1 a_2 a_3}$ and $d^{12\ldots k} \equiv d^{a_1 a_2 \ldots a_k}$ for the (adjoint) color degrees of freedom.

As mentioned in [32], the explicit computation of symmetrized traces is tedious to perform by hand but it is also well-suited for a computer implementation. The first non-trivial relations are relatively compact [33, 35]

$$\overset{\leftrightarrow}{\text{Tr}}\left(T^1 T^2 T^3\right) = \frac{i}{2} f^{123} \quad (7.4)$$

[16] Our normalization conventions are fixed by $\text{Tr}[T^a T^b] = \delta^{ab}/2$ and $[T^a, T^b] = i f^{abc} T^c$.

JHEP08(2014)099

$$\overset{\leftrightarrow}{\mathrm{Tr}}\,(T^1\,T^2\,T^3\,T^4) = 2d^{1234} + \frac{1}{6}\big(f^{23n}\,f^{14n} - f^{12n}\,f^{34n}\big), \tag{7.5}$$

$$\begin{aligned}\overset{\leftrightarrow}{\mathrm{Tr}}\,(T^1T^2T^3T^4T^5) = \frac{i}{12}\big[\, & f^{12a}f^{a4b}f^{b35} - 3f^{12a}f^{a3b}f^{b45} + f^{13a}f^{a2b}f^{b45} \\ & + f^{13a}f^{a4b}f^{b25} + f^{14a}f^{a2b}f^{b35} - f^{14a}f^{a3b}f^{b25}\big] \\ & + i\big[\, f^{23a}d^{a145} + f^{24a}d^{a135} + f^{25a}d^{a134} \\ & + f^{34a}d^{a125} + f^{35a}d^{a124} + f^{45a}d^{a123}\big],\end{aligned} \tag{7.6}$$

but the lenghty relations for $n = 6$ and 7 were computed using the `color` package of `FORM` [15, 16] and the $n = 6$ case can be found in the appendix B. Note that (7.5) and (7.6) have been cast into a minimal form in the sense that all the generalized Jacobi relations

$$f^{a[ij}f^{k]ab} = 0 \tag{7.7}$$

$$d^{a(ijk}f^{l)ab} = 0 \tag{7.8}$$

are taken into account. For the color structures involving a symmetrized trace, this amounts to placing leg number one to the $d^{1\cdots}$.

Once the color-dressed disk amplitude (7.1) is rewritten in this color basis, the subamplitude relations at various orders in α' impose selection rules on what kinds of tensor contribute to $\mathcal{M}_n^{\text{tree}}(\alpha')$ at the order in question. Keeping the first two terms in (6.1) $\sim \alpha'^0, \alpha'^2$, the KK identities [42] between $\mathcal{A}^{\text{YM}}$ select those color tensors with $n-2$ powers of structure constants and project out any symmetrized trace (7.3). The subamplitudes $\mathcal{A}^{F^4}$ associated with the first string corrections, on the other hand, satisfy another set of relations which we called KK-like in the discussion of subsection 6.3. They select those color tensors made of $n-4$ structure constants f^{abc} and one symmetrized four-trace $d^{1234} := \frac{1}{6}\sum_{\sigma\in S_3}\mathrm{Tr}\big[T^{\sigma(1)}T^{\sigma(2)}T^{\sigma(3)}T^4\big]$. This ties in with the symbolic vertices D^{abcd} and F^{abc} used in [31] to gain intuition for the amplitude relation between finite one-loop pure Yang Mills amplitudes $A^{(1)}_{n;1}$. The color tensors $d^{a_1a_2a_3a_4}(f^{bcd})^{n-4}$ selected by $\mathcal{A}^{F^4}$ are another manifestation of their intimate connection to the $A^{(1)}_{n;1}$.

As a first example, let us consider the four-point color-dressed amplitude up to $\mathcal{O}(\alpha'^2)$,

$$\begin{aligned}\mathcal{M}_4^{\text{tree}}(\alpha') = -\frac{1}{2}\,\big(\, & f^{12a}\,f^{a34}\,\mathcal{A}^{\text{YM}}(1,2,3,4) \;+\; f^{13a}\,f^{b24}\,\mathcal{A}^{\text{YM}}(1,3,2,4)\,\big) \\ & +\; 6\zeta(2)\alpha'^2\,d^{1234}\,\mathcal{A}^{F^4}(1,2,3,4) \;+\; \mathcal{O}(\alpha'^3),\end{aligned} \tag{7.9}$$

see [33] for the color structure at higher order in α'. The notation for higher multiplicity versions of (7.1) shall be lightened using

$$\mathcal{M}_n^{\text{tree}}(\alpha') \;\equiv\; \mathcal{M}_n^{\text{YM}} \;+\; \zeta(2)\,\alpha'^2\mathcal{M}_n^{F^4} \;+\; \mathcal{O}(\alpha'^3), \tag{7.10}$$

and the α'^2 correction $\mathcal{M}_n^{F^4}$ will be the object of main interest in this section where we show its tight connection to the one-loop integrand (5.34).

Before looking at the color tensor structure at order α'^2 and their interplay with $\mathcal{A}^{F^4}$ symmetries, let us review the color organization at the SYM level α'^0. At five-points, the KK relations for the field theory subamplitudes yield

$$\mathcal{M}_5^{\text{YM}} = -\frac{i}{2}\mathcal{A}^{\text{YM}}(1,2,3,4,5)f^{12a}f^{a3b}f^{b45} + \text{sym}(234)$$

JHEP08(2014)099

in agreement with the color-decomposition proposed by [29]. More generally, this reference suggests the following $(n-2)!$ element Kleiss-Kuijf bases

$$\big\{f^{1\sigma(2)a}\,f^{a\sigma(3)b}\,\cdots\,f^{z\sigma(n-1)n},\ \sigma\in S_{n-2}\big\},\quad \big\{\mathcal{A}^{\rm YM}\big(1,\sigma(2,3,\ldots,n-1),n\big),\ \sigma\in S_{n-2}\big\} \tag{7.11}$$

for the color factors $(f^{bcd})^{n-2}$ and the SYM subamplitudes (using Jacobi identities for the former and KK relations for the latter). In this setting, one can reproduce the $(n-2)!$ color-decomposition proven in [29]

$$\mathcal{M}_n^{\rm YM}=\frac{i^{n-2}}{2}\sum_{\sigma\in S_{n-2}}f^{1\sigma(2)a}\,f^{a\sigma(3)b}\,\cdots\,f^{z\sigma(n-1)n}\,\mathcal{A}^{\rm YM}\big(1,\sigma(2,3,\ldots,n-1),n\big)\,, \tag{7.12}$$

starting from (7.1), and the cancellation of $d^{12\ldots 2k}$ contributions at order α'^0 becomes manifest due to KK relations. In the remainder of this section, we will find remnants of (7.12) in $\mathcal{M}_n^{F^4}$, in particular the basis choice (7.11) for $(f^{bcd})^{n-2}$ color factors is path-breaking for the organization of the color tensors $d^{a_1a_2a_3a_4}(f^{bcd})^{n-4}$ relevant at α'^2 order.

7.1 The color-dressed $(n\leq 7)$-point disk amplitude at order α'^2

Keeping the dual bases (7.11) for $\mathcal{M}_n^{\rm YM}$ in mind, we shall next give $n=5,6,7$-point results for $\mathcal{M}_n^{F^4}$. According to (7.6), the five-point color tensors $d^{aijk}f^{lab}$ are brought into a (six element) basis of $d^{aij1}f^{lab}$ (with leg one attached to the d tensor) via generalized Jacobi identity $d^{a(ijk}f^{l)ab}=0$. This leads to a compact result for $\mathcal{M}_5^{F^4}$:

$$\begin{aligned}\mathcal{M}_5^{F^4}&=\sum_{\sigma\in S_4/Z_2}\overset{\leftrightarrow}{\rm Tr}\left[T^{a_{\sigma(1)}}T^{a_{\sigma(2)}}T^{a_{\sigma(3)}}T^{a_{\sigma(4)}}T^{a_5}\right]\mathcal{A}^{F^4}(\sigma(1,2,3,4),5)\\&=6i\langle C_{1,23,4,5}\,f^{23a}d^{a145}+C_{1,24,3,5}\,f^{24a}d^{a135}+C_{1,25,3,4}\,f^{25a}d^{a134}\\&\qquad+C_{1,34,2,5}\,f^{34a}d^{a125}+C_{1,35,2,4}\,f^{35a}d^{a124}+C_{1,45,2,3}\,f^{45a}d^{a123}\rangle\end{aligned} \tag{7.13}$$

First of all, the symmetries of $\mathcal{A}^{F^4}$ imply that color factors of type $f^{1\sigma(2)a}f^{a\sigma(3)b}f^{b\sigma(4)5}$ drop out, see the first two lines of (7.6). Secondly, the expansion $\mathcal{A}^{F^4}(1,2,3,4,5)=\langle C_{1,23,4,5}+C_{1,2,34,5}+C_{1,2,3,45}\rangle$ makes all contributions to a fixed (basis) color tensor $f^{ija}d^{a1kl}$ collapse to one single term $C_{1,ij,k,l}$.[17] The precise correspondence $C_{1,ij,k,l}\leftrightarrow f^{ija}d^{a1kl}$ between kinematic and color basis elements is a non-trivial reorganization when looked from the perspective of the composing $\mathcal{A}^{F^4}$ terms, especially at higher orders.

Even more striking cancellations occur when the symmetrized trace decompositions of the appendix B are used to evaluate the six- and seven-point color-dressed amplitudes at

[17] Going through the calculation reveals that the terms proportional to $f^{23a}d^{a145}$ are

$$\begin{aligned}&\mathcal{A}^{F^4}(1,2,3,4,5)+\mathcal{A}^{F^4}(1,2,3,5,4)+\mathcal{A}^{F^4}(1,2,4,3,5)+\mathcal{A}^{F^4}(1,2,4,5,3)\\&+\mathcal{A}^{F^4}(1,2,5,3,4)+\mathcal{A}^{F^4}(1,2,5,4,3)-\mathcal{A}^{F^4}(1,3,2,4,5)-\mathcal{A}^{F^4}(1,3,2,5,4)\\&-\mathcal{A}^{F^4}(1,3,4,2,5)-\mathcal{A}^{F^4}(1,3,5,2,4)+\mathcal{A}^{F^4}(1,4,2,3,5)-\mathcal{A}^{F^4}(1,4,3,2,5)\end{aligned}$$

and yet they collapse to a single term $6C_{1,23,4,5}$ once the relation (6.19) and the symmetries of the one-loop BRST invariants are used.

JHEP08(2014)099

order α'^2: the $35 = S_3^5$ term sum in the six-point case

$$
\begin{aligned}
-\frac{1}{6}\,\mathcal{M}_6^{F^4} = \langle & C_{1,23,45,6} f^{23a} d^{a16b} f^{b45} \;+\; C_{1,23,46,5} f^{23a} d^{a15b} f^{b46} \;+\; C_{1,23,56,4} f^{23a} d^{a14b} f^{b56} \\
& +\; C_{1,24,35,6} f^{24a} d^{a16b} f^{b35} \;+\; C_{1,24,36,5} f^{24a} d^{a15b} f^{b36} \;+\; C_{1,24,56,3} f^{24a} d^{a13b} f^{b56} \\
& +\; C_{1,25,34,6} f^{25a} d^{a16b} f^{b34} \;+\; C_{1,25,36,4} f^{25a} d^{a14b} f^{b36} \;+\; C_{1,25,46,3} f^{25a} d^{a13b} f^{b46} \\
& +\; C_{1,26,34,5} f^{26a} d^{a15b} f^{b34} \;+\; C_{1,26,35,4} f^{26a} d^{a14b} f^{b35} \;+\; C_{1,26,45,3} f^{26a} d^{a13b} f^{b45} \\
& +\; C_{1,34,56,2} f^{34a} d^{a12b} f^{b56} \;+\; C_{1,35,46,2} f^{35a} d^{a12b} f^{b46} \;+\; C_{1,36,45,2} f^{36a} d^{a12b} f^{b45} \\
& +\; \big[\, C_{1,234,5,6} f^{23a} f^{a4b} \;+\; C_{1,243,5,6} f^{24a} f^{a3b} \,\big] d^{b156} \\
& +\; \big[\, C_{1,235,4,6} f^{23a} f^{a5b} \;+\; C_{1,253,4,6} f^{25a} f^{a3b} \,\big] d^{b146} \\
& +\; \big[\, C_{1,245,3,6} f^{24a} f^{a5b} \;+\; C_{1,254,3,6} f^{25a} f^{a4b} \,\big] d^{b136} \\
& +\; \big[\, C_{1,236,4,5} f^{23a} f^{a6b} \;+\; C_{1,263,4,5} f^{26a} f^{a3b} \,\big] d^{b145} \\
& +\; \big[\, C_{1,246,3,5} f^{24a} f^{a6b} \;+\; C_{1,264,3,5} f^{26a} f^{a4b} \,\big] d^{b135} \\
& +\; \big[\, C_{1,256,3,4} f^{25a} f^{a6b} \;+\; C_{1,265,3,4} f^{26a} f^{a5b} \,\big] d^{b134} \\
& +\; \big[\, C_{1,345,2,6} f^{34a} f^{a5b} \;+\; C_{1,354,2,6} f^{35a} f^{a4b} \,\big] d^{b126} \\
& +\; \big[\, C_{1,346,2,5} f^{34a} f^{a6b} \;+\; C_{1,364,2,5} f^{36a} f^{a4b} \,\big] d^{b125} \\
& +\; \big[\, C_{1,356,2,4} f^{35a} f^{a6b} \;+\; C_{1,365,2,4} f^{36a} f^{a5b} \,\big] d^{b124} \\
& +\; \big[\, C_{1,456,2,3} f^{45a} f^{a6b} \;+\; C_{1,465,2,3} f^{46a} f^{a5b} \,\big] d^{b123} \rangle
\end{aligned}
\tag{7.14}
$$

exhibits color-kinematic-correspondence

$$
C_{1,23,45,6} \leftrightarrow f^{23a} d^{a16b} f^{b45}, \quad C_{1,234,5,6} \leftrightarrow f^{23a} f^{a4b} d^{b156}.
$$

Likewise, the $225 = S_3^6$ terms in

$$
\begin{aligned}
\frac{i}{6}\,\mathcal{M}_7^{F^4} = \langle & 15 \text{ terms } \big[\, \langle C_{1,2345,6,7} f^{23a} f^{a4b} f^{b5c} + \text{sym}(345) \,\big] d^{c167} \\
& +\; 60 \text{ terms } \big[\, C_{1,234,56,7} f^{23a} f^{a4b} \;+\; (3 \leftrightarrow 4) \,\big] f^{56c} d^{bc17} \\
& +\; 15 \text{ terms } C_{1,23,45,67} f^{23a} f^{45b} f^{67c} d^{1abc} \rangle
\end{aligned}
\tag{7.15}
$$

allow to read off the dictionary

$$
\begin{aligned}
C_{1,23,45,67} &\leftrightarrow f^{23a} f^{45b} f^{67c} d^{1abc} \\
C_{1,234,56,7} &\leftrightarrow f^{23a} f^{a4b} f^{56c} d^{bc17} \\
C_{1,2345,6,7} &\leftrightarrow f^{23a} f^{a4b} f^{b5c} d^{c167}.
\end{aligned}
\tag{7.16}
$$

In the next subsection, we shall put these observations into a more general context. Note that d^{123456} and $d^{12345a} f^{a67}$ tensors (or more generally $d^{a_1 \cdots a_6} (f^{bcd})^{n-6}$ and $d^{a_1 \ldots a_{2k}} (f^{bcd})^{n-2k}$ at $k \geq 3$) from the rank ≥ 6 traces do not contribute at $\mathcal{O}(\alpha'^2)$ because of the KK-like amplitude relations between $\mathcal{A}^{F^4}$.

7.2 Dual bases in color and kinematic space

We conclude from the calculations above that the BRST invariants $C_{1,\ldots}$ are natural objects to appear not only in the one-loop integrand but also in color-dressed tree-level amplitudes.

JHEP08(2014)099

According to (6.17), they are related to subamplitudes $\mathcal{A}^{F^4}$ at order α'^2 by a change of (S_3^{n-1} element) basis with coefficients ± 1 and automatically solve their KK-like relations. Moreover, the $C_{1,a_1\ldots a_p,b_1\ldots,b_q,c_1\ldots c_r}$ inherit all symmetry properties of the Berends-Giele current triplet $M^i_{a_1\ldots a_p} M^j_{b_1\ldots,b_q} M^k_{c_1\ldots c_r}$, see subsection 4.5. This makes their S_3^{n-1} basis under relations with constant coefficients manifest and leads to the observed harmony with the symmetries of color tensors $d^{a_1a_2a_3a_4}(f^{bcd})^{n-4}$.

In fact, arriving at the simple results (7.13), (7.14) and (7.15) for the α'^2 correction to the color-dressed amplitude crucially relies on the fact that the dimension of the basis for color factors and the kinematics matches. This fact has been exploited to choose "compatible" bases of color structures and corresponding kinematics, generalizing the tree-level correspondence (7.11) between color factors $(f^{bcd})^{n-2}$ and $\mathcal{A}^{\rm YM}$ in their $(n-2)!$ KK bases. In the SYM case, the reduction to $(n-2)!$ bases makes use of Jacobi identities on the color side and the KK relations for the subamplitudes.

We shall now explain why also the $d^{a_1a_2a_3a_4}(f^{bcd})^{n-4}$ color factors align into a basis of S_3^{n-1} elements. The reduction algorithm consists of two steps:

1. Move label number one to the symmetrized four-trace, i.e. $d^{ijk\cdots}(f^{bcd})^{n-4} \mapsto \sum d^{1pq\cdots}(f^{bcd})^{n-4}$, by repeated use of generalized Jacobi identities (7.8). At five-points, one applications is enough,

$$d^{234a} f^{1ab} = -d^{123a} f^{4ab} - d^{412a} f^{3ab} - d^{341a} f^{2ab}. \tag{7.17}$$

   For six-points there are two possibilities for the color factors which do not contain the label number one inside the symmetrized trace d^{ijkl}. The number of color space propagators δ^{ab} between leg number one and d^{ijkl} is either one (as in $f^{34a} f^{1ab} d^{b246}$) or two (as in $f^{12a} f^{a3b} d^{b456}$). For the one-propagator-link one uses the generalized Jacobi identity (7.17), whereas in the two-propagator-case the identity

$$f^{12a} f^{a3b} d^{b456} = -f^{12a} d^{a45b} f^{b63} - f^{12a} d^{a64b} f^{b53} - f^{12a} d^{a56b} f^{b43} \tag{7.18}$$

   reduces it to terms of one-propagator form where (7.17) can be applied. The analysis for higher-points is analogous, with more successive applications of (7.8) needed. The possibility to reduce $d^{ijkl}(f^{bcd})^{n-4} \mapsto \sum d^{1pqr}(f^{bcd})^{n-4}$ is the color dual of the finding in subsection 4.5 that any $C_{i,\ldots}$ with $i \neq 1$ can be expressed as a sum of $C_{1,\ldots}$ in the BRST cohomology.

2. Mod out the $d^{1pqr}(f^{bcd})^{n-4}$ by Jacobi identities (7.7) among the f^{n-4} factors: consider a generic color structure $d^{1\cdots}$ passing the first selection rule,

$$\begin{aligned} d^{1x_py_qz_r}\, & f^{a_1a_2x_2}\, f^{x_2a_3x_3}\, \ldots\, f^{x_{p-2}a_{p-1}x_{p-1}}\, f^{x_{p-1}a_px_p} \\ & \times f^{b_1b_2y_2}\, f^{y_2b_3y_3}\, \ldots\, f^{y_{q-2}b_{q-1}y_{q-1}}\, f^{y_{q-1}b_qy_q} \\ & \times f^{c_1c_2z_2}\, f^{z_2c_3z_3}\, \ldots\, f^{z_{r-2}c_{r-1}z_{r-1}}\, f^{z_{r-1}c_rz_r}. \end{aligned} \tag{7.19}$$

   Each of the remaining three slots of $d^{1x_py_qz_r}$ can adjoin a tree subdiagram with p, q and r external legs, respectively, such that $p+q+r = n-1$. Within the

JHEP08(2014)099

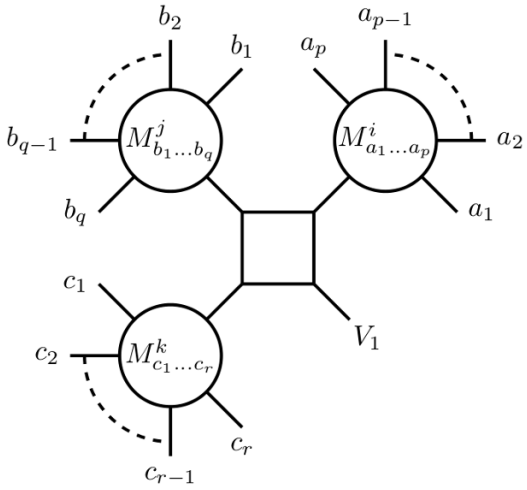

Figure 8. The diagram associated with the leading term of $C_{1,a_1\ldots a_p,b_1\ldots b_q,c_1\ldots c_r} = V_1\, M^i_{a_1\ldots a_p}\, M^j_{b_1\ldots b_q}\, M^k_{c_1\ldots c_r}$+BRST invariant completion. The Kleiss-Kuijf relations for the tree subdiagrams represented by $M^i_{a_1\ldots a_p}$, $M^j_{b_1\ldots b_q}$ and $M^k_{c_1\ldots c_r}$ yield all identities between the permutations σ, ρ, π in $C_{1,\sigma(a_1\ldots a_p),\rho(b_1\ldots b_q),\pi(c_1\ldots c_r)}$.

color tensors of each subdiagram like $f^{a_1a_2x_2}f^{x_2a_3x_3}\ldots f^{x_{p-2}a_{p-1}x_{p-1}}f^{x_{p-1}a_px_p}$, we can eliminate the Jacobi redundancy in analogy to the tree-level KK basis (7.11). This amounts to the convention that a_1 is kept fixed at the "outmost" structure constant $f^{a_1\sigma(a_2)x_2}$ whereas the remaining free indices $\{a_2, a_3, \ldots, a_p\}$ can appear in any of the $(p-1)!$ possible permutations. Then, the collection of $f^{a_1\sigma(a_2)x_2}f^{x_2\sigma(a_3)x_3}\ldots f^{x_{p-2}\sigma(a_{p-1})x_{p-1}}f^{x_{p-1}\sigma(a_p)x_p}$ with $\sigma \in S_{p-1}$ exhaust all Jacobi-independent half-ladder diagrams with fixed endpoints a_1 and x_p. The kinematical dual is the reduction of $C_{1,\sigma(a_1\ldots a_p),\ldots}$ to the smaller set of $C_{1,a_1\sigma(a_2\ldots a_p),\ldots}$.

After this two-step reduction, the basis dimension for the color factors $d^{ijkl}(f^{bcd})^{n-4}$ is manifestly equal to the unsigned Stirling number S_3^{n-1}, the same number which governs the number of independent BRST invariants $C_{1,\ldots}$.

A more intuitive understanding of the interplay between kinematic- and color basis can be found by inspecting the unique term $V_1M^i_{a_1\ldots a_p}M^j_{b_1\ldots b_q}M^k_{c_1\ldots c_r}$ in $C_{1,a_1\ldots a_p,b_1\ldots b_q,c_1\ldots c_r}$ with the standalone unintegrated vertex operator V_1, see the explicit expression in subsection 4.4. The ellipsis in $C_{1,a_1\ldots a_p,b_1\ldots b_q,c_1\ldots c_r} = V_1M^i_{a_1\ldots a_p}M^j_{b_1\ldots b_q}M^k_{c_1\ldots c_r} + \ldots$ obeys the same symmetry properties, so the first term is a valid representative for the symmetry analysis. Recall that the Berends-Giele currents $M^i_{a_1\ldots a_p}$ are color-ordered $(p+1)$-point amplitudes with leg number $p+1$ off-shell (corresponding to the color index x_p which is contracted into the box $d^{1x_py_qz_r}$), see figure 8. Within each of these three off-shell subamplitudes $M^i_{a_1\ldots a_p}$, we pick a Kleiss-Kuijf basis where, again, a_1 is kept fixed as the first subscript of $M^i_{a_1\ldots}$ and $a_{\geq 2}$ can appear in any permutation.

JHEP08(2014)099

Each of the KK basis elements $M^i_{a_1\sigma(a_2\ldots a_p)}$ is accompanied by a f^{p-1} color factor which is adapted to the permutation $\sigma \in S_{p-1}$ according to the tree-level rule (7.12):

$$M^i_{a_1\sigma(a_2\ldots a_p)} \quad \leftrightarrow \quad f^{a_1\sigma(a_2)x_2}\, f^{x_2\sigma(a_3)x_3}\, \ldots\, f^{x_{p-1}\sigma(a_p)x_p}\,.$$

The three chains of f corresponding to the $M^i_{\ldots}, M^j_{\ldots}$ and $M^k_{\ldots}$ are then contracted with the x_p, y_q, z_r indices of $d^{1x_py_qz_r}$, i.e. glued to the three corners of the box where leg one is not attached to. This amounts to the following rule how the dual S_3^{n-1} element bases for color- and kinematic factors enter $\mathcal{M}_n^{F^4}$: permutations of $C_{1,a_1\ldots a_p,b_1\ldots b_q,c_1\ldots c_r}$ for fixed sets $\{a_1,a_2,\ldots,a_p\}$, $\{b_1,b_2,\ldots,b_q\}$ and $\{c_1,c_2,\ldots,c_r\}$ always appear in the combination

$$\begin{aligned}\sum_{\sigma\in S_{p-1}}\sum_{\rho\in S_{q-1}}\sum_{\pi\in S_{r-1}} & C_{1,a_1\sigma(a_2\ldots a_p),b_1\rho(b_2\ldots b_q),c_1\pi(c_2\ldots c_r)} \times f^{a_1\sigma(a_2)x_2}\, f^{x_2\sigma(a_3)x_3}\, \ldots\, f^{x_{p-1}\sigma(a_p)x_p} \\ & \times f^{b_1\rho(b_2)y_2}\, f^{y_2\rho(b_3)y_3}\, \ldots\, f^{y_{q-1}\rho(b_q)y_q} \times f^{c_1\pi(c_2)z_2}\, f^{z_2\pi(c_3)z_3}\, \ldots\, f^{z_{r-1}\pi(c_r)z_r}\, d^{1x_py_qz_r}\,,\end{aligned} \tag{7.20}$$

in agreement with our results (7.13), (7.14) and (7.15) for $\mathcal{M}^{F^4}_{n\leq 7}$. This can be recognized as sum over the S_3^{n-1} partitions of legs $23\ldots n$ into three cycles, see subsection 5.4 for the associated set $\mathcal{S}_3^{n-1}$. Using the latter notation defined in (5.32), we can compactly write the n-point color-dressed amplitude as

$$\begin{aligned}\mathcal{M}_n^{F^4} = 6i^n \sum_{\sigma\times\pi\times\rho\in\mathcal{S}_3^{n-1}} & \langle C_{1,\sigma(23\ldots p),\pi(p+1\ldots q),\rho(q+1\ldots n)}\rangle\, d^{1x_py_qz_n}\, \sigma\left(f^{23x_3}\, f^{x_34x_4}\, \ldots\, f^{x_{p-1}px_p}\right) \\ & \pi\left(f^{p+1,p+2,y_{p+2}}\, f^{y_{p+2},p+3,y_{p+3}}\, \ldots\, f^{y_{q-1},q,y_q}\right)\, \rho\left(f^{q+1,q+2,z_{q+2}}\, f^{z_{q+2},q+3,z_{q+3}}\, \ldots\, f^{z_{n-1},n,z_n}\right).\end{aligned} \tag{7.21}$$

As in (5.34), the numbers p and q are defined through the cardinality of the permutations to be $p = |\sigma| + 1$ and $q - p = |\pi|$.

7.3 Duality between one-loop integrands and $\mathcal{M}_n^{F^4}$

This subsection is devoted to the close relationship between $\mathcal{M}_n^{F^4}$ and the one-loop kinematic factor K_n. Our final expressions (5.28), (5.30), (5.31) and (5.34) for K_5, K_6, K_7 and K_n can be obtained from the corresponding $\mathcal{M}_n^{F^4}$ using a well-defined one-to-one map between $d^{1pqr}(f^{bcd})^{n-4}$ colors factors and the $(X_{rs})^{n-4}$ polynomials in the worldsheet integrand. The color basis choice of having leg one attached to $d^{1\cdots}$ corresponds to integrating by parts on the worldsheet such that only X_{rs} with $r, s \neq 1$ enter the minimal form of K_n.

Let us start with lower order examples for the $d^{1pqr}(f^{bcd})^{n-4} \leftrightarrow (X_{rs})^{n-4}$ dictionary. First of all, $K_4 = \mathcal{A}^{F^4}(1,2,3,4)$ is related to $\mathcal{M}_4^{F^4}$ via

$$6d^{1234} \quad \longleftrightarrow \quad 1. \tag{7.22}$$

Comparing the representation (5.28) for K_5 with (7.13) yields the five-points dictionary,

$$6if^{23a}d^{a145} \quad \longleftrightarrow \quad X_{23}. \tag{7.23}$$

The two six-point topologies $C_{1,234,5,6}, C_{1,23,45,6}$ in K_6 and $\mathcal{M}_6^{F^4}$ (given by (5.30) and (7.14), respectively) are accompanied by

$$\begin{aligned}-6\, f^{23a}f^{a4b}d^{b156} \quad &\longleftrightarrow \quad X_{23}\,(X_{24}+X_{34}) \\ -6\, f^{23a}f^{45b}d^{ab16} \quad &\longleftrightarrow \quad X_{23}\,X_{45}\,,\end{aligned} \tag{7.24}$$

JHEP08(2014)099

and the $C_{1,2345,6,7}, C_{1,234,56,7}$ and $C_{1,23,45,67}$ at seven-points are dressed by

$$\begin{aligned} -6i\, f^{23a} f^{a4b} f^{b5c} d^{c167} &\quad\longleftrightarrow\quad X_{23}\,(X_{24}+X_{34})\,(X_{25}+X_{35}+X_{45}) \\ -6i\, f^{23a} f^{a4b} f^{56c} d^{bc17} &\quad\longleftrightarrow\quad X_{23}\,(X_{24}+X_{34})\,X_{56} \\ -6i\, f^{23a} f^{45b} f^{67c} d^{abc1} &\quad\longleftrightarrow\quad X_{23}\,X_{45}\,X_{67}, \end{aligned} \tag{7.25}$$

see (5.31) for K_7 and (7.15) for $\mathcal{M}_7^{F^4}$, respectively.

Both sides of the mappings (7.22) to (7.25) have the same symmetries in the labels $23\ldots n$ — the left hand side because of Jacobi identities, the right hand side due to algebraic identities such as $X_{23}(X_{24}+X_{34})+\text{cyc}(234)=0 \;\leftrightarrow\; f^{[23|a} f^{a|4]b}=0$ or

$$\begin{aligned} 0 = {} & X_{23}\,(X_{24}+X_{34})\,(X_{25}+X_{35}+X_{45}) \;-\; (4\leftrightarrow 5) \\ & +\; X_{45}\,(X_{42}+X_{52})\,(X_{43}+X_{53}+X_{23}) \;-\; (2\leftrightarrow 3) \end{aligned}$$

corresponding to $f^{23a} f^{a[4|b} f^{b|5]c} + f^{45a} f^{a[2|b} f^{b|3]c} = 0$ (which in turn reflects the "third" BRST symmetry $T_{23[45]}+T_{45[23]}=0$ under the map (2.15)).

More generally, the three independent cubic subdiagrams contracted with the x_p, y_q, z_r indices of $d^{1x_p y_q z_r}$ each correspond to a separate nested product of worldsheet functions like $\prod_{k=3}^{p}\sum_{m=2}^{k-1} X_{mk}$:

$$f^{23x_3}\, f^{x_3 4 x_4}\,\ldots\, f^{x_{p-1}px_p} \quad\longleftrightarrow\quad X_{23}\,(X_{24}+X_{34})\,\ldots\,(X_{2p}+X_{3p}+\ldots+X_{p-1,p})\,.$$

Combining the three subdiagrams with the central quartic vertex, we arrive at the following dictionary between $d^{1pqr}(f^{bcd})^{n-4}$ color tensors and $(X_{rs})^{n-4}$ worldsheet integrands:

$$\begin{aligned} & 6i^n\, d^{1x_p y_q z_n}\, f^{23x_3}\, f^{x_3 4 x_4}\,\ldots\, f^{x_{p-2},p-1,x_{p-1}}\, f^{x_{p-1}px_p} \\ & \quad\times\; f^{p+1,p+2,y_{p+2}}\, f^{y_{p+2},p+3,y_{p+3}}\,\ldots\, f^{y_{q-2},q-1,y_{q-1}}\, f^{y_{q-1},q,y_q} \\ & \quad\times\; f^{q+1,q+2,z_{q+2}}\, f^{z_{q+2},q+3,z_{q+3}}\,\ldots\, f^{z_{n-2},n-1,z_{n-1}}\, f^{z_{n-1},n,z_n} \\ & \longleftrightarrow\quad \Bigg(\prod_{k=3}^{p}\sum_{m=2}^{k-1} X_{mk}\Bigg)\Bigg(\prod_{k=p+2}^{q}\sum_{m=p+1}^{k-1} X_{mk}\Bigg)\Bigg(\prod_{k=q+2}^{n}\sum_{m=q+1}^{k-1} X_{mk}\Bigg)\,. \end{aligned} \tag{7.26}$$

Given the most general definition (7.26) of the double-arrow notation, the final forms (5.34) and (7.21) for K_n and $\mathcal{M}_n^{F^4}$, respectively, are related by

$$\mathcal{M}_n^{F^4} \quad\longleftrightarrow\quad K_n. \tag{7.27}$$

This map allows to construct the one-loop kinematic factor by knowledge of the corresponding color-dressed tree amplitude at order α'^2.

7.4 Proving total symmetry of K_n

In this subsection, we use the $\mathcal{M}_n^{F^4} \leftrightarrow K_n$ duality (7.26) to carry out the outstanding proof that K_n as given by (5.34) is completely symmetric in all labels $(12\ldots n)$.

Representing K_n and $\mathcal{M}_n^{F^4}$ in their minimal S_3^{n-1} basis hides the total permutation symmetry in $12\ldots n$. Leg number one is singled out in (5.34) and (7.21) on the level of both

JHEP08(2014)099

BRST invariants $C_{1,\dots}$ and color factors $d^{1pqr}(f^{bcd})^{n-4}$ and worldsheet functions $(X_{rs})^{n-4}$ since $r,s \neq 1$. Since the remaining legs $23\dots n$ enter on equal footing, it is sufficient to prove $1 \leftrightarrow 2$ symmetry of K_n and $\mathcal{M}_n^{F^4}$. The explicit check would require several changes of basis — firstly in kinematic space from $C_{1,\dots}$ to $C_{2,\dots}$ using the identities in subsection 4.5, secondly in color space $d^{1pqr} \mapsto d^{2pqr}$ and thirdly in the worldsheet integrand $(X_{rs})^{n-4}$ from $r,s \neq 1$ to $r,s \neq 2$. We will instead apply an indirect argument.

The mapping (7.26) between color factors and $(X_{rs})^{n-4}$ integrands respects not only the standard Jacobi identities (7.7) but also those relations which are required for the aforementioned change of basis: the generalized Jacobi relations (7.8) are dual to integration by parts. The simplest non-trivial example can be found at five-points,

$$f^{12a}d^{a345} + f^{13a}d^{a245} + f^{14a}d^{a235} + f^{15a}d^{a234} = 0 \quad \longleftrightarrow \quad X_{12} + X_{13} + X_{14} + X_{15} = 0$$

where the validity of the X_{1j} relation rests on integration against the Koba Nielsen factor, see subsection 5.3. At higher multiplicity, the form $\prod_{k=3}^{p}\sum_{m=2}^{k-1} X_{mk}$ of the worldsheet functions is sufficiently integration-by-parts-friendly such that they still obey four term identities of type (7.8), e.g.

$$f^{12a}f^{a3b}d^{b456} = f^{12a}(d^{a45b}f^{b36} + \text{cyc}(456)) \quad \longleftrightarrow \quad X_{12}(X_{13}+X_{23}) = X_{12}(X_{34}+X_{35}+X_{36})$$

as well as

$$f^{23a}f^{a1b}d^{b456} + \text{sym}(1456) = 0 \quad \longleftrightarrow \quad X_{23}(X_{21}+X_{24}+X_{25}+X_{26}+2 \leftrightarrow 3) = 0$$

at six-points. Generalizations to higher multiplicity are straightforward.

Since the mapping (7.26) preserves the generalized Jacobi relations (7.8), the hidden total symmetry of $\mathcal{M}_n^{F^4}$ implies that of K_n. Our computation of $\mathcal{M}_n^{F^4}$ started with the manifestly $1 \leftrightarrow 2$ symmetric expression (7.1) summing over all cyclically inequivalent permutations, so we can be sure that the representation (7.21) is totally symmetric. Our derivation of the final result (5.34) for K_n, on the other hand, started with the $V^1 \leftrightarrow U^2$ asymmetric prescription (3.1) and involved incomplete arguments about the absence of additional b-ghost contributions. It is quite assuring to see that (5.34) must be totally symmetric as well — if the b-ghost contributed to K_n via OPE contractions, then this would probably modify its symmetry properties due to the asymmetric response of V^1 and $U^{j\geq 2}$, suggesting their absence.

7.5 Correspondence between color and kinematics in $\mathcal{M}_n^{F^4}$

It was argued in [27] that the symmetric role of kinematic numerators and color factors in SYM amplitudes suggests to impose dual Jacobi identities in the kinematic sector. They have been successfully applied to simplify the calculation of multiloop amplitudes in both SYM and gravity [69, 70]. The BRST building blocks technique can be used to obtain local BCJ numerators at tree-level for any number of external legs [28] through the low energy limit of string amplitudes. Therefore, it seems worthwhile to search for possible BCJ generalizations at the next order in the momentum expansion of the superstring.

JHEP08(2014)099

multiplicity	4	5	6	7	8	n
diagrams per color-stripped $\mathcal{A}^{\rm YM}$	2	5	14	42	132	$C(n-2)$
diagrams per color-dressed $\mathcal{M}_n^{\rm YM}$	3	15	105	945	10395	$(2n-5)!!$
diagrams per color-stripped $\mathcal{A}^{F^4}$	1	5	21	84	330	$\frac{n-3}{2}\,C(n-2)$
diagrams per color-dressed $\mathcal{M}_n^{F^4}$	1	10	105	1260	17325	$\frac{n-3}{3}\,(2n-5)!!$

Table 2. Number of diagrams which compose the different types of amplitudes according to their kinematic pole structure. Here, $C(k)$ denotes the $k^{\rm th}$ Catalan number $C(k) = \frac{(2k)!}{k!(k+1)!}$.

So in this final subsection we show that the final form (7.21) for the color-dressed $\mathcal{O}(\alpha'^2)$ amplitude $\mathcal{M}_n^{F^4}$ is symmetric under exchange of color and kinematics. This observation has no direct relevance for one-loop amplitudes but it is an interesting generalization of the color-kinematic-symmetric representation [27]

$$\mathcal{M}_n^{\rm YM} = \sum_i^{(2n-5)!!} \frac{c_i\, n_i}{\prod_{\alpha_i}^{n-3} s_{\alpha_i}} \tag{7.28}$$

for color-dressed SYM tree amplitudes. The sum over i encompasses all cubic diagrams with $n-3$ propagators $\prod_{\alpha_i} s_{\alpha_i}^{-1}$, and c_i, n_i denote the associated color- and kinematic structures. One rewarding property of (7.28) is the fact that gravity tree amplitudes can be immediately obtained by replacing color factors $c_i \mapsto \tilde{n}_i$ by another copy $\tilde{n}_i$ of the kinematic numerators n_i, provided that the latter satisfy Jacobi identities dual to the color factors c_i.

This encouraged us to build the $\mathcal{M}_n^{F^4}$ analogue (7.32) of (7.28), we regard it as the first step towards a double copy construction that could ultimately yield a gravity analogue of $\mathcal{A}^{F^4}$ amplitudes. Instead of the cubic diagrams in (7.28), the diagrams in $\mathcal{M}_n^{F^4}$ are built from one totally symmetry quartic vertex and $n-4$ cubic vertices.

The expansion of $\mathcal{M}_n^{F^4}$ in terms of BRST invariants $C_{1,\ldots}$ takes a very compact form, but since each $C_{1,\ldots}$ encompasses several kinematic poles (i.e. diagrams of the form figure 4), it is not immediately obvious from (7.21) how the kinematic numerators associated to these poles combine with color factors. In section 4, we have constructed these numerators in pure spinor superspace, they are quartic expressions $\langle T_{d_1\ldots d_s} T^i_{a_1\ldots a_p} T^j_{b_1\ldots b_q} T^k_{c_1\ldots c_r}\rangle$ in tree subdiagrams $T_{\ldots}$ and $T^{i,j,k}_{\ldots}$ attached to a totally symmetric quartic vertex. As an artifact of inserting leg one via unintegrated vertex operator V^1, each numerator obeys $1 \in \{d_1, d_2 \ldots d_s\}$.

The number of diagrams per color-dressed $\mathcal{M}_n^{F^4}$ is displayed in the last line of table 2.[18] In order to resolve all of them, we start from the $\langle M_{d_1\ldots d_s} M^i_{a_1\ldots a_p} M^j_{b_1\ldots b_q} M^k_{c_1\ldots c_r}\rangle$ constituents of $C_{1,\ldots}$ and expand the Berends-Giele currents in terms of BRST building

[18] In order to arrive at the diagrams per topology, note that there are $(2p-3)!!$ subdiagrams within all the $T_{i_1 i_2\ldots i_p}$ permutations (at fixed set $\{i_1 i_2 \ldots i_p\}$), corresponding to the $(2n-5)!!$ cubic diagrams in an n-point color-dressed SYM tree amplitude. For instance, there are three different diagrams

$$\frac{T_{123}}{s_{12}s_{123}}, \quad \frac{T_{231}}{s_{23}s_{123}}, \quad \frac{T_{312}}{s_{13}s_{123}}$$

corresponding to the s-, t- and u channel in $\mathcal{M}_4^{\rm YM}$ and 15 different T_{pqrs}/s^3 subdiagrams.

JHEP08(2014)099

blocks $\langle T_{d_1\ldots d_s} T^i_{a_1\ldots a_p} T^j_{b_1\ldots b_q} T^k_{c_1\ldots c_r}\rangle$. Each individual kinematic diagram is associated with a separate color factor $d^{ijkl}(f^{bcd})^{n-4}$ which precisely matches its propagator structure. Of course, the color algebra makes use of the generalized Jacobi identities (7.7) and (7.8), e.g. the five-point result (7.13) yields

$$\begin{aligned}\frac{1}{6i}\mathcal{M}_5^{F^4} = &\frac{1}{s_{23}}\langle V_1 T^i_{23} T^j_4 T^k_5\rangle f^{23a} d^{a145} + \frac{1}{s_{24}}\langle V_1 T^i_{24} T^j_3 T^k_5\rangle f^{24a} d^{a135}\\ &+ \frac{1}{s_{25}}\langle V_1 T^i_{25} T^j_3 T^k_4\rangle f^{25a} d^{a134} + \frac{1}{s_{34}}\langle V_1 T^i_{34} T^j_2 T^k_5\rangle f^{34a} d^{a125}\\ &+ \frac{1}{s_{35}}\langle V_1 T^i_{35} T^j_2 T^k_4\rangle f^{35a} d^{a124} + \frac{1}{s_{45}}\langle V_1 T^i_{45} T^j_2 T^k_3\rangle f^{45a} d^{a123}\\ &+ \frac{1}{s_{12}}\langle T_{12} T^i_3 T^j_4 T^k_5\rangle f^{12a} d^{a345} + \frac{1}{s_{13}}\langle T_{13} T^i_2 T^j_4 T^k_5\rangle f^{13a} d^{a245}\\ &+ \frac{1}{s_{14}}\langle T_{14} T^i_2 T^j_3 T^k_5\rangle f^{14a} d^{a235} + \frac{1}{s_{15}}\langle T_{15} T^i_2 T^j_3 T^k_4\rangle f^{15a} d^{a234}.\end{aligned} \tag{7.29}$$

Similarly at six- and seven-points, (7.14) and (7.15) become

$$\begin{aligned}-\frac{1}{6}\,\mathcal{M}_6^{F^4} &= 45 \text{ terms } \Big[\, f^{12a}\, f^{34b}\, d^{ab56}\, \frac{1}{s_{12}\, s_{34}} \langle T_{12}\, T^i_{34}\, T^j_5\, T^k_6\rangle\Big]\\ &\quad +\ 60 \text{ terms } \Big[\, f^{12a}\, f^{a3b}\, d^{b456}\, \frac{1}{s_{12}\, s_{123}} \langle T_{123}\, T^i_4\, T^j_5\, T^k_6\rangle\Big] \end{aligned} \tag{7.30}$$

$$\begin{aligned}\frac{1}{6i}\,\mathcal{M}_7^{F^4} &= 105 \text{ terms } \Big[\, f^{12a}\, f^{34b}\, f^{56c}\, d^{abc7}\, \frac{1}{s_{12}\, s_{34}\, s_{56}} \langle T_{12}\, T^i_{34}\, T^j_{56}\, T^k_7\rangle\Big]\\ &\quad +\ 630 \text{ terms } \Big[\, f^{12a}\, f^{a3b}\, f^{45c}\, d^{bc67}\, \frac{1}{s_{12}\, s_{123}\, s_{45}} \langle T_{123}\, T^i_{45}\, T^j_6\, T^k_7\rangle\Big]\\ &\quad +\ 420 \text{ terms } \Big[\, f^{12a}\, f^{a3b}\, f^{b4c}\, d^{c567}\, \frac{1}{s_{12}\, s_{123}\, s_{1234}} \langle T_{1234}\, T^i_5\, T^j_6\, T^k_7\rangle\Big]\\ &\quad +\ 105 \text{ terms } \Big[\, f^{12a}\, f^{34b}\, f^{abc}\, d^{c567}\, \frac{1}{s_{12}\, s_{34}\, s_{1234}}\, 2\langle T_{12[34]}\, T^i_5\, T^j_6\, T^k_7\rangle\Big]\,. \end{aligned} \tag{7.31}$$

For each topology, we sum over all permutations that are inequivalent under the symmetries of $\langle T_{d_1\ldots d_s} T^i_{a_1\ldots a_p} T^j_{b_1\ldots b_q} T^k_{c_1\ldots c_r}\rangle$, up to the aforementioned rule that $1 \in \{d_1, d_2 \ldots d_s\}$ holds in each term. For instance, one of the suppressed terms in (7.30) reads $f^{23a} f^{45b} d^{ab61} \langle V_1 T^i_{23} T^j_{45} T^k_6\rangle/(s_{23} s_{45})$.

For higher multiplicity, this generalizes to

$$\mathcal{M}_n^{F^4} = \sum_I^{\frac{1}{3}(n-3)(2n-5)!!} \frac{C_I\, N_I}{\prod_{\alpha_I}^{n-4} s_{\alpha_I}}, \tag{7.32}$$

where the sum over I encompasses all box diagrams with four tree subdiagrams at the corners, $\prod_{\alpha_I} s_{\alpha_I}^{-1}$ denotes the associated $n-4$ propagators, and the numerator contains a color- $C_I \leftrightarrow d^{ijkl}(f^{bcd})^{n-4}$ and a kinematic factor $N_I \leftrightarrow \langle T_{\ldots} T^i_{\ldots} T^j_{\ldots} T^k_{\ldots}\rangle$.

Unfortunately, our superspace representation of these numerators N_I does not yet lead to kinematic Jacobi identities dual to the color relation $d^{a(ijk} f^{l)ab} = 0$, e.g.

$$\langle T_{12} T^i_3 T^j_4 T^k_5\rangle + \langle T_{13} T^i_2 T^j_4 T^k_5\rangle + \langle T_{14} T^i_2 T^j_3 T^k_5\rangle + \langle T_{15} T^i_2 T^j_3 T^k_4\rangle \neq 0. \tag{7.33}$$

JHEP08(2014)099

One could suspect that this is an artifact of the asymmetric role of label one in $\langle T_{12}T_3^iT_4^jT_5^k\rangle$ and $\langle T_1T_{23}^iT_4^jT_5^k\rangle$. It would be desirable to find an improved representation of the N_I such that a strict duality holds

$$C_I + C_J + C_K + C_L = 0 \quad \leftrightarrow \quad N_I + N_J + N_K + N_L = 0. \tag{7.34}$$

This is for instance achieved by the five-point box-numerators γ_{ij} in [81]. Finding such a duality-satisfying representation for n-point kinematics and studying the significance of the gravity amplitude obtained by replacing $C_I \mapsto \tilde{N}_I$ in (7.32) is left for future work.

8 Conclusions

In this article, we have derived BRST invariant worldsheet integrands K_n for one-loop open superstring amplitudes involving any number n of massless gauge multiplets. Our main result (5.34) is expressed in terms of kinematic building blocks $C_{1,\ldots}$ which are implicitly given in terms of $\mathcal{O}(\alpha'^2)$ tree subamplitudes via (6.17). Since we have used BRST invariance in determining the associated worldsheet functions, our setup is by construction blind to the hexagon anomaly [21, 22]. A superspace treatment of anomalous amplitude ingredients along the lines of [74] is left for future work.

Both the superspace kinematics $C_{1,\ldots}$ and the associated worldsheet functions fall into a basis of dimension S_3^{n-1}, an unsigned Stirling number of first kind. The same kind of symmetries also govern the color-dressed tree amplitude $\mathcal{M}_n^{F^4}$ at order α'^2, so we point out a duality between its minimal form (7.21) in a color basis and the one-loop integrand K_n given by (5.34). The link is a one-to-one dictionary (7.26) between color factors $d^{ijkl}(f^{bcd})^{n-4}$ (encompassing one symmetrized four-trace and structure constants otherwise) and worldsheet functions $X_{ij}^{n-4} \equiv (s_{ij}\eta_{ij})^{n-4}$ (built from $\eta_{ij} = \partial_i\langle x(z_i,\overline{z}_i)x(z_j,\overline{z}_j)\rangle$) present in K_n.

A detailed analysis of the S_3^{n-1} worldsheet integrals is left for future work. The only comment we want to make at this point is that the integrand structure closely parallels the tree-level result from [6, 7]: each $z_i \to z_j$ singularity in both the tree-level and the one-loop integrand is always accompanied by a corresponding Mandelstam numerator s_{ij}, i.e. we have $s_{ij}\eta_{ij} = s_{ij}/z_{ij} + \mathcal{O}(z_{ij})$. This guarantees that the integration does not introduce any poles in kinematic invariants, i.e. that the full propagator structure due to open string exchange is captured by the $C_{1,\ldots}$. On the other hand, loop amplitudes additionally involve non-analytic momentum dependencies, so the main challenge in further studying the worldsheet integrals is to identify the polylogarithms that arise in both leading and subleading orders in α'.

Acknowledgments

We wish to thank Johannes Brödel for intensive discussions and continuous inspiring email exchange. We are very grateful to Michael Green and Stephan Stieberger for numerous insightful conversations and valuable advice. We are indebted to Johannes Brödel, Henrik Johansson, Stephan Stieberger and Stefan Theisen for reading our draft and making helpful suggestions and to Bernhard Wurm for his creative ideas about notation. C.M.

JHEP08(2014)099

acknowledges support by the European Research Council Advanced Grant No. 247252 of Michael Green and would like to thank the Albert Einstein Institute in Potsdam for hospitality and financial support during completion of this work. O.S. is very grateful to Michael Green and his European Research Council Advanced Grant No. 247252 for hospitality and financial support during preparation and completion of this work. Moreover, he would like to thank DAMTP for creating a stimulating atmosphere during the visits in Cambridge.

A On the uniqueness of the b-ghost zero mode contribution

The computation of higher-point amplitudes at one-loop might involve different d_α zero-mode distributions among the picture changing operators, the b-ghost and the external vertices. In addition, the b-ghost might have OPE singularities with the other operators, resulting in yet other types of contributions.

However, the following argument supports that the *zero-mode* b-ghost contribution at one-loop is unique and given by $d^4\delta'(N)$. In order to see this note that the zero-mode contribution of the picture changing operators is fixed and given by $(d)^{10}(\lambda)^{10}\delta^{10}(N)\delta(J)(\theta)^{11}\delta^{11}(\lambda)$, which is responsible among other things for absorbing all 11 bosonic zero-modes of w_α [5]. Now assume that the b-ghost zero-mode contribution contains $(d)^n\delta^m(N)$ and note that performing the zero-mode integral $\int[\mathcal{D}N]\,d^{16}d(d)^{10+n}(\lambda)^{10}\delta^{11}(\lambda)\delta^m(N)\,\big[\text{vertices}\big]$ has the net effect of replacing $(d)^{6-n}(N)^m$ zero-modes from the external vertices by a function quadratic in λ,

$$(d)^{6-n}(N)^m \longrightarrow (\lambda)^2 \tag{A.1}$$

since $[\mathcal{D}N]$ has ghost number -8. From the expression (3.3) for the b-ghost it follows that the possible values are $n = 0, 1, 2, 4$ and $m = 1, 2, 3$. However $(d)^4(N)^m$ and $(d)^5(N)^m$ have no (00002) component for any value of m [34] and the zero-mode integral vanishes for $n = 1, 2$. Therefore group theory alone does not exclude the possibility of the b-ghost contributing either 0 or 4 zero modes of d_α with varying number of derivatives of delta functions. So let us analyze these possibilities in separate.

The possible zero-mode contribution from the b-ghost containing no d_α zero modes are given by

$$N\Pi^2\delta(N), \qquad N^2\Pi^2\delta'(N), \tag{A.2}$$

but they both vanish due to oversaturated N^{mn} zero modes using that $N\delta(N) = 0$. For the same reason, any contribution $\sim J$ and $\sim J^2$ from the b-ghost (e.g. $(d)^4JN\delta'''(N)$ or $(d)^4(J)^2\delta'''(N)$ at four d_α zero modes) is suppressed by the $\delta(J) = 0$ from the picture changing operator Z_J. That leaves the three contributions

$$(d)^4\delta'(N), \quad (d)^4N\delta''(N), \quad (d)^4(N)^2\delta'''(N) \tag{A.3}$$

of uniform type under integration by parts. Therefore the zero-mode contribution from the b-ghost is unique and given by $(d)^4\delta'(N)$. In this paper we studied the cohomology properties of precisely this class of terms in order to anticipate its appearance in the final expression for the superspace kinematic factors.

JHEP08(2014)099

When the b-ghost is allowed to contribute non-zero modes the number of possibilities increases, but only those which also contain either 0 or 4 zero modes of d_α can have a nonzero impact on the amplitude. As argued in [9], terms involving only one OPE contraction of the b-ghost vanish because they are proportional to a derivative with respect to the position z_0 of the b-ghost insertion. Since z_0 appears nowhere else in the correlation function, those terms are total derivatives which integrate to zero due to the doubling trick. Having excluded single OPEs with the b ghost, it follows that the five-point amplitude gets no contribution at all from b ghost OPEs [9], but from six-points onwards these terms are not excluded. For example, the b-ghost term $(d)^4 JN\delta'''(N)$ can in principle have simultaneous OPEs involving J and N with different external vertices leading to factors which are not manifestly total derivatives. This term requires two d's and three N's from the integrated vertices which can be provided in case of six and more external states.

B Symmetrized traces for six- and seven-point amplitudes

The six- and seven-point symmetrized traces can be computed using the `color` package of `FORM`. After rewriting the generated terms in the Kleiss-Kuijf basis of f^{n-2} and in the "Stirling" basis of $d^{ijkl}(f^{bcd})^{n-4}$ one gets for six-points

$$
\begin{aligned}
&\mathrm{Tr}(T^1T^2T^3T^4T^5T^6) \;+\; \mathrm{Tr}(T^6T^5T^4T^3T^2T^1) = 2d^{123456}\\
&+\frac{1}{5}f^{12a}f^{a3b}f^{b4c}f^{c56} - \frac{1}{20}f^{12a}f^{a3b}f^{b5c}f^{c46} - \frac{1}{20}f^{12a}f^{a4b}f^{b3c}f^{c56} - \frac{1}{20}f^{12a}f^{a4b}f^{b5c}f^{c36}\\
&-\frac{1}{20}f^{12a}f^{a5b}f^{b3c}f^{c46} + \frac{1}{30}f^{12a}f^{a5b}f^{b4c}f^{c36} - \frac{1}{20}f^{13a}f^{a2b}f^{b4c}f^{c56} + \frac{1}{30}f^{13a}f^{a2b}f^{b5c}f^{c46}\\
&-\frac{1}{20}f^{13a}f^{a4b}f^{b2c}f^{c56} - \frac{1}{20}f^{13a}f^{a4b}f^{b5c}f^{c26} - \frac{1}{20}f^{13a}f^{a5b}f^{b2c}f^{c46} + \frac{1}{30}f^{13a}f^{a5b}f^{b4c}f^{c26}\\
&-\frac{1}{20}f^{14a}f^{a2b}f^{b3c}f^{c56} + \frac{1}{30}f^{14a}f^{a2b}f^{b5c}f^{c36} + \frac{1}{30}f^{14a}f^{a3b}f^{b2c}f^{c56} + \frac{1}{30}f^{14a}f^{a3b}f^{b5c}f^{c26}\\
&-\frac{1}{20}f^{14a}f^{a5b}f^{b2c}f^{c36} + \frac{1}{30}f^{14a}f^{a5b}f^{b3c}f^{c26} - \frac{1}{20}f^{15a}f^{a2b}f^{b3c}f^{c46} + \frac{1}{30}f^{15a}f^{a2b}f^{b4c}f^{c36}\\
&+\frac{1}{30}f^{15a}f^{a3b}f^{b2c}f^{c46} + \frac{1}{30}f^{15a}f^{a3b}f^{b4c}f^{c26} + \frac{1}{30}f^{15a}f^{a4b}f^{b2c}f^{c36} - \frac{1}{20}f^{15a}f^{a4b}f^{b3c}f^{c26}\\
&-\frac{1}{2}f^{23a}d^{a14b}f^{b56} - \frac{1}{2}f^{23a}d^{a15b}f^{b46} - \frac{1}{2}f^{23a}d^{a16b}f^{b45} - \frac{1}{2}f^{24a}d^{a13b}f^{b56} - \frac{1}{2}f^{24a}d^{a15b}f^{b36}\\
&-\frac{1}{2}f^{24a}d^{a16b}f^{b35} - \frac{1}{2}f^{25a}d^{a13b}f^{b46} - \frac{1}{2}f^{25a}d^{a14b}f^{b36} - \frac{1}{2}f^{25a}d^{a16b}f^{b34} - \frac{1}{2}f^{26a}d^{a13b}f^{b45}\\
&-\frac{1}{2}f^{26a}d^{a14b}f^{b35} - \frac{1}{2}f^{26a}d^{a15b}f^{b34} - \frac{1}{2}f^{34a}d^{a12b}f^{b56} - \frac{1}{2}f^{35a}d^{a12b}f^{b46} - \frac{1}{2}f^{36a}d^{a12b}f^{b45}\\
&-\frac{1}{3}f^{23a}f^{a5b}d^{b146} - \frac{1}{3}f^{23a}f^{a6b}d^{b145} - \frac{1}{3}f^{24a}f^{a3b}d^{b156} - \frac{1}{3}f^{24a}f^{a5b}d^{b136}\\
&-\frac{1}{3}f^{24a}f^{a6b}d^{b135} - \frac{1}{3}f^{26a}f^{a5b}d^{b134} + \frac{2}{3}f^{34a}f^{a2b}d^{b156} - \frac{1}{3}f^{34a}f^{a5b}d^{b126}\\
&-\frac{1}{3}f^{34a}f^{a6b}d^{b125} + \frac{1}{3}f^{35a}f^{a2b}d^{b146} + \frac{1}{3}f^{36a}f^{a2b}d^{b145} - \frac{1}{3}f^{36a}f^{a5b}d^{b124}\\
&+\frac{1}{3}f^{45a}f^{a2b}d^{b136} + \frac{1}{3}f^{45a}f^{a3b}d^{b126} + \frac{1}{3}f^{46a}f^{a2b}d^{b135} + \frac{1}{3}f^{46a}f^{a3b}d^{b125}
\end{aligned}
$$

JHEP08(2014)099

$$-\frac{1}{3}f^{46a}f^{a5b}d^{b123}+\frac{2}{3}f^{56a}f^{a2b}d^{b134}+\frac{2}{3}f^{56a}f^{a3b}d^{b124}+\frac{2}{3}f^{56a}f^{a4b}d^{b123}.$$

The seven-point expression is too big to be illuminating and was therefore omitted.[19]

C The higher-multiplicity BRST invariants

In this appendix we list the explicit form of the BRST invariants which appear in the eight-point amplitude.

$$\begin{aligned}
C_{1,23456,7,8} = &\Big(M_1\,M^i_{23456}+M_{612}\,M^i_{345}+M_{56123}\,M^i_4\ +\big[M_{12345}\,M^i_6+M_{1234}\,M^i_{56}\\
&+M_{123}\,M^i_{456}+M_{12}\,M^i_{3456}+M_{6123}\,M^i_{45}+M_{61234}\,M^i_5+(2,3\leftrightarrow 6,5)\,\big]\Big)M^j_7M^k_8\\
C_{1,2345,67,8} = &\Big(M_1\,M^i_{2345}\,M^j_{67}+M_{215}\,M^i_{34}\,M^j_{67}\ +\big[\,M_{71}\,M^i_{2345}\,M^j_6-(6\leftrightarrow 7)\,\big]\\
&+\big[\,M_{12}\,M^i_{345}+M_{123}\,M^i_{45}+M_{1234}\,M^i_5+\ M_{4512}\,M^i_3-(2,3\leftrightarrow 5,4)\,\big]\,M^j_{67}\\
&+\big\{\,\big[\,M_{712}\,M^i_{345}\,M^j_6+M_{7123}\,M^i_{45}\,M^j_6+\ M_{71234}\,M^i_5\,M^j_6-M_{7125}\,M^i_{34}\,M^j_6\\
&+M_{75123}\,M^i_4\,M^j_6+M_{57123}\,M^i_4\,M^j_6-(2,3\leftrightarrow 5,4)\,\big]-(6\leftrightarrow 7)\,\big\}\Big)\,M^k_8\\
C_{1,234,567,8} = &\Big(M_1\,M^i_{234}\,M^j_{567}+\big[\,M_{12}\,M^i_{34}\,M^j_{567}+M_{14}\,M^i_{32}\,M^j_{567}+M_{123}\,M^i_4\,M^j_{567}\\
&+M_{143}\,M^i_2\,M^j_{567}+M_{214}\,M^i_3\,M^j_{567}+(2,3,4\leftrightarrow 5,6,7)\,\big]+\big[\,M_{712}\,M^i_{34}\,M^j_{56}\\
&+M_{7123}\,M^i_4\,M^j_{56}+M_{6712}\,M^i_{34}\,M^j_5+M_{6512}\,M^i_{34}\,M^j_7+M_{5124}\,M^i_3\,M^j_{67}\\
&+M_{2157}\,M^i_{34}\,M^j_6+M_{67123}\,M^i_4\,M^j_5-M_{65124}\,M^i_3\,M^j_7-M_{32157}\,M^i_4\,M^j_6\\
&+M_{24157}\,M^i_3\,M^j_6+(2\leftrightarrow 4)+(5\leftrightarrow 7)\,\big]\Big)\,M^k_8\\
C_{1,234,56,78} = &\Big(M_1\,M^i_{234}+M_{214}\,M^i_3+\big[\,M_{12}\,M^i_{34}+M_{123}\,M_4+(2\leftrightarrow 4)\,\big]\Big)\,M^j_{56}\,M^k_{78}\\
&+\Big(\big[\,M_{15}\,M^i_{234}\,M^j_6-M_{16}\,M^i_{234}\,M^j_5\,\big]\,M^k_{78}+\big[\,M_{216}\,M^i_{34}\,M^j_5+M_{6123}\,M^i_4\,M^j_5\\
&+M_{5124}\,M^i_3\,M^j_6+(2\leftrightarrow 4)-(5\leftrightarrow 6)\,\big]\,M^k_{78}+(5,6\leftrightarrow 7,8)\Big)\\
&+\big[\,M_{617}\,M^i_{234}\,M^j_5\,M^k_8-(5\leftrightarrow 6)-(7\leftrightarrow 8)\,\big]+\big[\,(M_{7126}+M_{7162})\,M^i_{34}\,M^j_5\,M^k_8\\
&+(M_{75123}+M_{57123})\,M^i_4\,M^j_6\,M^k_8-(M_{75124}+M_{57124})\,M^i_3\,M^j_6\,M^k_8\\
&+(2\leftrightarrow 4)-(5\leftrightarrow 6)-(7\leftrightarrow 8)\,\big]
\end{aligned}$$

References

[1] A.B. Goncharov, M. Spradlin, C. Vergu and A. Volovich, *Classical polylogarithms for amplitudes and Wilson loops*, *Phys. Rev. Lett.* **105** (2010) 151605 [arXiv:1006.5703] [INSPIRE].

[19] It is commented out in the TEX source file.

JHEP08(2014)099

[2] N. Arkani-Hamed, J.L. Bourjaily, F. Cachazo, S. Caron-Huot and J. Trnka, *The all-loop integrand for scattering amplitudes in planar $N = 4$ SYM*, *JHEP* **01** (2011) 041 [arXiv:1008.2958] [INSPIRE].

[3] S. Caron-Huot and S. He, *Jumpstarting the all-loop S-matrix of planar $N = 4$ super Yang-Mills*, *JHEP* **07** (2012) 174 [arXiv:1112.1060] [INSPIRE].

[4] N. Berkovits, *Super Poincaré covariant quantization of the superstring*, *JHEP* **04** (2000) 018 [hep-th/0001035] [INSPIRE].

[5] N. Berkovits, *Multiloop amplitudes and vanishing theorems using the pure spinor formalism for the superstring*, *JHEP* **09** (2004) 047 [hep-th/0406055] [INSPIRE].

[6] C.R. Mafra, O. Schlotterer and S. Stieberger, *Complete N-point superstring disk amplitude I. Pure spinor computation*, *Nucl. Phys.* **B 873** (2013) 419 [arXiv:1106.2645] [INSPIRE].

[7] C.R. Mafra, O. Schlotterer and S. Stieberger, *Complete N-point superstring disk amplitude II. Amplitude and hypergeometric function structure*, *Nucl. Phys.* **B 873** (2013) 461 [arXiv:1106.2646] [INSPIRE].

[8] C.R. Mafra, O. Schlotterer, S. Stieberger and D. Tsimpis, *A recursive method for SYM n-point tree amplitudes*, *Phys. Rev.* **D 83** (2011) 126012 [arXiv:1012.3981] [INSPIRE].

[9] C.R. Mafra and C. Stahn, *The one-loop open superstring massless five-point amplitude with the non-minimal pure spinor formalism*, *JHEP* **03** (2009) 126 [arXiv:0902.1539] [INSPIRE].

[10] E.P. Verlinde and H.L. Verlinde, *Chiral bosonization, determinants and the string partition function*, *Nucl. Phys.* **B 288** (1987) 357 [INSPIRE].

[11] M.B. Green and J.H. Schwarz, *Covariant description of superstrings*, *Phys. Lett.* **B 136** (1984) 367 [INSPIRE].

[12] M.B. Green and J.H. Schwarz, *Properties of the covariant formulation of superstring theories*, *Nucl. Phys.* **B 243** (1984) 285 [INSPIRE].

[13] M.B. Green, J.H. Schwarz and E. Witten, *Superstring theory. Vol. 2: loop amplitudes, anomalies and phenomenology*, Cambridge Monographs On Mathematical Physics, Cambdrige Univ. Pr., Cambridge U.K. (1987).

[14] C.R. Mafra, *PSS: a FORM program to evaluate pure spinor superspace expressions*, arXiv:1007.4999 [INSPIRE].

[15] J.A.M. Vermaseren, *New features of FORM*, math-ph/0010025 [INSPIRE].

[16] M. Tentyukov and J.A.M. Vermaseren, *The multithreaded version of FORM*, *Comput. Phys. Commun.* **181** (2010) 1419 [hep-ph/0702279] [INSPIRE].

[17] C.R. Mafra, *Towards field theory amplitudes from the cohomology of pure spinor superspace*, *JHEP* **11** (2010) 096 [arXiv:1007.3639] [INSPIRE].

[18] C.R. Mafra, O. Schlotterer, S. Stieberger and D. Tsimpis, *Six open string disk amplitude in pure spinor superspace*, *Nucl. Phys.* **B 846** (2011) 359 [arXiv:1011.0994] [INSPIRE].

[19] E. Witten, *Twistor-like transform in ten-dimensions*, *Nucl. Phys.* **B 266** (1986) 245 [INSPIRE].

[20] E. D'Hoker and D.H. Phong, *The geometry of string perturbation theory*, *Rev. Mod. Phys.* **60** (1988) 917 [INSPIRE].

JHEP08(2014)099

[21] M.B. Green and J.H. Schwarz, *Anomaly cancellation in supersymmetric $D = 10$ gauge theory and superstring theory*, *Phys. Lett.* **B 149** (1984) 117 [INSPIRE].

[22] M.B. Green and J.H. Schwarz, *The hexagon gauge anomaly in type I superstring theory*, *Nucl. Phys.* **B 255** (1985) 93 [INSPIRE].

[23] S. Stieberger, *Open & closed vs. pure open string disk amplitudes*, arXiv:0907.2211 [INSPIRE].

[24] N.E.J. Bjerrum-Bohr, P.H. Damgaard and P. Vanhove, *Minimal basis for gauge theory amplitudes*, *Phys. Rev. Lett.* **103** (2009) 161602 [arXiv:0907.1425] [INSPIRE].

[25] L. Brink, J.H. Schwarz and J. Scherk, *Supersymmetric Yang-Mills theories*, *Nucl. Phys.* **B 121** (1977) 77 [INSPIRE].

[26] M.B. Green, J.H. Schwarz and L. Brink, *$N = 4$ Yang-Mills and $N = 8$ supergravity as limits of string theories*, *Nucl. Phys.* **B 198** (1982) 474 [INSPIRE].

[27] Z. Bern, J.J.M. Carrasco and H. Johansson, *New relations for gauge-theory amplitudes*, *Phys. Rev.* **D 78** (2008) 085011 [arXiv:0805.3993] [INSPIRE].

[28] C.R. Mafra, O. Schlotterer and S. Stieberger, *Explicit BCJ numerators from pure spinors*, *JHEP* **07** (2011) 092 [arXiv:1104.5224] [INSPIRE].

[29] V. Del Duca, L.J. Dixon and F. Maltoni, *New color decompositions for gauge amplitudes at tree and loop level*, *Nucl. Phys.* **B 571** (2000) 51 [hep-ph/9910563] [INSPIRE].

[30] I. Oda and M. Tonin, *Y-formalism in pure spinor quantization of superstrings*, *Nucl. Phys.* **B 727** (2005) 176 [hep-th/0505277] [INSPIRE].

[31] N.E.J. Bjerrum-Bohr, P.H. Damgaard, H. Johansson and T. Sondergaard, *Monodromy-like relations for finite loop amplitudes*, *JHEP* **05** (2011) 039 [arXiv:1103.6190] [INSPIRE].

[32] T. van Ritbergen, A.N. Schellekens and J.A.M. Vermaseren, *Group theory factors for Feynman diagrams*, *Int. J. Mod. Phys.* **A 14** (1999) 41 [hep-ph/9802376] [INSPIRE].

[33] A. Bilal, *Higher derivative corrections to the non-Abelian Born-Infeld action*, *Nucl. Phys.* **B 618** (2001) 21 [hep-th/0106062] [INSPIRE].

[34] M.A.A. van Leeuwen, A.M. Cohen and B. Lisser, *LiE, a package for Lie group computations*, Computer Algebra Nederland, Amsterdam The Netherlands (1992) [ISBN:90-74116-02-7].

[35] D. Lüst, O. Schlotterer, S. Stieberger and T.R. Taylor, *The LHC string hunter's companion (II): five-particle amplitudes and universal properties*, *Nucl. Phys.* **B 828** (2010) 139 [arXiv:0908.0409] [INSPIRE].

[36] P.S. Howe, *Pure spinors lines in superspace and ten-dimensional supersymmetric theories*, *Phys. Lett.* **B 258** (1991) 141 [*Addendum ibid.* **B 259** (1991) 511] [INSPIRE].

[37] N. Berkovits, *Explaining pure spinor superspace*, hep-th/0612021 [INSPIRE].

[38] N. Berkovits, *Cohomology in the pure spinor formalism for the superstring*, *JHEP* **09** (2000) 046 [hep-th/0006003] [INSPIRE].

[39] N. Berkovits and B.C. Vallilo, *Consistency of super-Poincaré covariant superstring tree amplitudes*, *JHEP* **07** (2000) 015 [hep-th/0004171] [INSPIRE].

[40] N. Berkovits, *Super-Poincaré covariant two-loop superstring amplitudes*, *JHEP* **01** (2006) 005 [hep-th/0503197] [INSPIRE].

JHEP08(2014)099

[41] N. Berkovits and C.R. Mafra, *Equivalence of two-loop superstring amplitudes in the pure spinor and RNS formalisms*, *Phys. Rev. Lett.* **96** (2006) 011602 [hep-th/0509234] [INSPIRE].

[42] R. Kleiss and H. Kuijf, *Multi-gluon cross-sections and five jet production at hadron colliders*, *Nucl. Phys.* **B 312** (1989) 616 [INSPIRE].

[43] F.A. Berends and W.T. Giele, *Recursive calculations for processes with n-gluons*, *Nucl. Phys.* **B 306** (1988) 759 [INSPIRE].

[44] G. Policastro and D. Tsimpis, R^4*, purified*, *Class. Quant. Grav.* **23** (2006) 4753 [hep-th/0603165] [INSPIRE].

[45] J.P. Harnad and S. Shnider, *Constraints and field equations for ten-dimensional super Yang-Mills theory*, *Commun. Math. Phys.* **106** (1986) 183 [INSPIRE].

[46] P.A. Grassi and L. Tamassia, *Vertex operators for closed superstrings*, *JHEP* **07** (2004) 071 [hep-th/0405072] [INSPIRE].

[47] N. Berkovits, *ICTP lectures on covariant quantization of the superstring*, hep-th/0209059 [INSPIRE].

[48] H. Gomez, *One-loop superstring amplitude from integrals on pure spinors space*, *JHEP* **12** (2009) 034 [arXiv:0910.3405] [INSPIRE].

[49] H. Gomez and C.R. Mafra, *The overall coefficient of the two-loop superstring amplitude using pure spinors*, *JHEP* **05** (2010) 017 [arXiv:1003.0678] [INSPIRE].

[50] Y. Aisaka and N. Berkovits, *Pure spinor vertex operators in Siegel gauge and loop amplitude regularization*, *JHEP* **07** (2009) 062 [arXiv:0903.3443] [INSPIRE].

[51] R.M. Schabinger, *One-loop $N = 4$ super Yang-Mills scattering amplitudes in d dimensions, relation to open strings and polygonal Wilson loops*, *J. Phys.* **A 44** (2011) 454007 [arXiv:1104.3873] [INSPIRE].

[52] Z. Bern, L.J. Dixon, D.C. Dunbar and D.A. Kosower, *One loop n point gauge theory amplitudes, unitarity and collinear limits*, *Nucl. Phys.* **B 425** (1994) 217 [hep-ph/9403226] [INSPIRE].

[53] S. Stieberger and T.R. Taylor, *Amplitude for N-gluon superstring scattering*, *Phys. Rev. Lett.* **97** (2006) 211601 [hep-th/0607184] [INSPIRE].

[54] S. Stieberger and T.R. Taylor, *Multi-gluon scattering in open superstring theory*, *Phys. Rev.* **D 74** (2006) 126007 [hep-th/0609175] [INSPIRE].

[55] Z. Bern, G. Chalmers, L.J. Dixon and D.A. Kosower, *One loop N-gluon amplitudes with maximal helicity violation via collinear limits*, *Phys. Rev. Lett.* **72** (1994) 2134 [hep-ph/9312333] [INSPIRE].

[56] S.J. Parke and T.R. Taylor, *An amplitude for N-gluon scattering*, *Phys. Rev. Lett.* **56** (1986) 2459 [INSPIRE].

[57] C.R. Mafra, *Pure spinor superspace identities for massless four-point kinematic factors*, *JHEP* **04** (2008) 093 [arXiv:0801.0580] [INSPIRE].

[58] J. Bjornsson and M.B. Green, 5 *loops in* 24/5 *dimensions*, *JHEP* **08** (2010) 132 [arXiv:1004.2692] [INSPIRE].

[59] J. Bjornsson, *Multi-loop amplitudes in maximally supersymmetric pure spinor field theory*, *JHEP* **01** (2011) 002 [arXiv:1009.5906] [INSPIRE].

JHEP08(2014)099

[60] N. Berkovits, M.B. Green, J.G. Russo and P. Vanhove, *Non-renormalization conditions for four-gluon scattering in supersymmetric string and field theory*, *JHEP* **11** (2009) 063 [arXiv:0908.1923] [INSPIRE].

[61] N. Berkovits, *Pure spinor formalism as an $N = 2$ topological string*, *JHEP* **10** (2005) 089 [hep-th/0509120] [INSPIRE].

[62] N. Arkani-Hamed, F. Cachazo, C. Cheung and J. Kaplan, *The S-matrix in twistor space*, *JHEP* **03** (2010) 110 [arXiv:0903.2110] [INSPIRE].

[63] C.R. Mafra, *Four-point one-loop amplitude computation in the pure spinor formalism*, *JHEP* **01** (2006) 075 [hep-th/0512052] [INSPIRE].

[64] L.A. Barreiro and R. Medina, *5-field terms in the open superstring effective action*, *JHEP* **03** (2005) 055 [hep-th/0503182] [INSPIRE].

[65] A. Tsuchiya, *More on one loop massless amplitudes of superstring theories*, *Phys. Rev.* **D 39** (1989) 1626 [INSPIRE].

[66] J.J. Atick and A. Sen, *Covariant one loop fermion emission amplitudes in closed string theories*, *Nucl. Phys.* **B 293** (1987) 317 [INSPIRE].

[67] S. Stieberger and T.R. Taylor, *Non-Abelian Born-Infeld action and type I heterotic duality 2: nonrenormalization theorems*, *Nucl. Phys.* **B 648** (2003) 3 [hep-th/0209064] [INSPIRE].

[68] N. Arkani-Hamed, F. Cachazo and J. Kaplan, *What is the simplest quantum field theory?*, *JHEP* **09** (2010) 016 [arXiv:0808.1446] [INSPIRE].

[69] Z. Bern, J.J.M. Carrasco and H. Johansson, *Perturbative quantum gravity as a double copy of gauge theory*, *Phys. Rev. Lett.* **105** (2010) 061602 [arXiv:1004.0476] [INSPIRE].

[70] Z. Bern, J.J.M. Carrasco, L.J. Dixon, H. Johansson and R. Roiban, *Simplifying multiloop integrands and ultraviolet divergences of gauge theory and gravity amplitudes*, *Phys. Rev.* **D 85** (2012) 105014 [arXiv:1201.5366] [INSPIRE].

[71] J.M. Drummond, J.M. Henn and J. Plefka, *Yangian symmetry of scattering amplitudes in $N = 4$ super Yang-Mills theory*, *JHEP* **05** (2009) 046 [arXiv:0902.2987] [INSPIRE].

[72] J.M. Drummond, J. Henn, G.P. Korchemsky and E. Sokatchev, *Dual superconformal symmetry of scattering amplitudes in $N = 4$ super-Yang-Mills theory*, *Nucl. Phys.* **B 828** (2010) 317 [arXiv:0807.1095] [INSPIRE].

[73] R.H. Boels and R.S. Isermann, *Yang-Mills amplitude relations at loop level from non-adjacent BCFW shifts*, *JHEP* **03** (2012) 051 [arXiv:1110.4462] [INSPIRE].

[74] N. Berkovits and C.R. Mafra, *Some superstring amplitude computations with the non-minimal pure spinor formalism*, *JHEP* **11** (2006) 079 [hep-th/0607187] [INSPIRE].

[75] L.J. Dixon, *Scattering amplitudes: the most perfect microscopic structures in the universe*, *J. Phys.* **A 44** (2011) 454001 [arXiv:1105.0771] [INSPIRE].

[76] M.B. Green and J.H. Schwarz, *Supersymmetrical dual string theory. 3. Loops and renormalization*, *Nucl. Phys.* **B 198** (1982) 441 [INSPIRE].

[77] E. D'Hoker, M. Gutperle and D.H. Phong, *Two-loop superstrings and S-duality*, *Nucl. Phys.* **B 722** (2005) 81 [hep-th/0503180] [INSPIRE].

[78] E. D'Hoker and D.H. Phong, *Two-loop superstrings VI: non-renormalization theorems and the 4-point function*, *Nucl. Phys.* **B 715** (2005) 3 [hep-th/0501197] [INSPIRE].

JHEP08(2014)099

[79] Z. Bern, L.J. Dixon, D.C. Dunbar and D.A. Kosower, *One loop selfdual and $N = 4$ super Yang-Mills*, *Phys. Lett.* **B 394** (1997) 105 [hep-th/9611127] [INSPIRE].

[80] Z. Bern, J.J.M. Carrasco, L.J. Dixon, H. Johansson and R. Roiban, *The complete four-loop four-point amplitude in $N = 4$ super-Yang-Mills theory*, *Phys. Rev.* **D 82** (2010) 125040 [arXiv:1008.3327] [INSPIRE].

[81] J.J. Carrasco and H. Johansson, *Five-point amplitudes in $N = 4$ super-Yang-Mills theory and $N = 8$ supergravity*, *Phys. Rev.* **D 85** (2012) 025006 [arXiv:1106.4711] [INSPIRE].

[82] B.E.W. Nilsson, *Pure spinors as auxiliary fields in the ten-dimensional supersymmetric Yang-Mills theory*, *Class. Quant. Grav.* **3** (1986) L41 [INSPIRE].

[83] Z. Bern and D.A. Kosower, *Color decomposition of one loop amplitudes in gauge theories*, *Nucl. Phys.* **B 362** (1991) 389 [INSPIRE].

[84] Z. Bern and D.A. Kosower, *The computation of loop amplitudes in gauge theories*, *Nucl. Phys.* **B 379** (1992) 451 [INSPIRE].

[85] M.B. Green, private communication.

JHEP08(2014)099