

Towards the n-point one-loop superstring amplitude I: Pure spinors and superfield kinematics

Carlos R. Mafra[†] and Oliver Schlotterer^{‡,*}

[†]*Mathematical Sciences and STAG Research Centre, University of Southampton,
Highfield, Southampton, SO17 1BJ, UK*

[‡]*Max-Planck-Institut für Gravitationsphysik
Albert-Einstein-Institut, 14476 Potsdam, Germany*

**Perimeter Institute for Theoretical Physics, Waterloo, ON N2L 2Y5, Canada*

This is the first installment of a series of three papers in which we describe a method to determine higher-point correlation functions in one-loop open-superstring amplitudes from first principles. In this first part, we exploit the synergy between the cohomological features of pure-spinor superspace and the pure-spinor zero-mode integration rules of the one-loop amplitude prescription. This leads to the study of a rich variety of multiparticle superfields which are local, have covariant BRST variations, and are compatible with the particularities of the pure-spinor amplitude prescription. Several objects related to these superfields, such as their non-local counterparts and the so-called BRST pseudo-invariants, are thoroughly reviewed and put into new light. Their properties will turn out to be mysteriously connected to products of one-loop worldsheet functions in packages dubbed “generalized elliptic integrands”, whose prominence will be seen in the later parts of this series of papers.

December 2018

[†] email: c.r.mafra@soton.ac.uk

[‡] email: olivers@aei.mpg.de

Contents

1	Introduction	2
2	Basic formalism	5
	2.1. The pure-spinor amplitude prescription	6
	2.2. Chiral splitting of the Koba–Nielsen factor	13
3	Multiparticle SYM superfields	16
	3.1. Combinatorics on words	16
	3.2. Single-particle	19
	3.3. Two-particle	19
	3.4. Multiparticle	20
4	Pure-spinor superspace: local superfields	25
	4.1. Scalars	25
	4.2. Vectors	26
	4.3. Tensors	28
	4.4. Refined building blocks	31
5	Pure-spinor superspace: non-local superfields	38
	5.1. The Berends–Giele map	38
	5.2. BRST pseudo-invariants	42
	5.3. Anomaly counterparts of BRST invariants	45
	5.4. BRST cohomology identities	47
6	Conclusions	49
A	A new algorithm for the combinatorics of shuffle invariants	50
	A.1. Multi-word generalization of the rho map	50
	A.2. Unifying all C -like building blocks	52
	A.3. Change-of-basis identities	53
B	Empty BRST cohomology for manifestly local expressions	55
	B.1. BRST-closed expressions without momentum conservation	56
	B.2. BRST-closed expressions using momentum conservation	59
C	Eight-point anomalous building blocks	62
D	The BRST variations of local building blocks	64
	D.1. Scalar $T_{A,B,C}$	64
	D.2. Tensorial $T_{A,B,C,\dots}^{m\dots}$	65
	D.3. Refined $J_{A B,C,\dots}^{m\dots}$	66
	D.4. Anomaly building blocks $Y_{A,B,C,\dots}^{m\dots}$	67

1. Introduction

This is the first part of a series of papers [1] towards the derivation of n -point one-loop correlators of open- and closed-superstring states using the pure-spinor formalism [2,3]. When we refer to section and equation numbers from the papers II and III, these numbers will be prefixed by the roman numerals II and III accordingly.

A variety of recent developments revealed hidden simplicity and unexpectedly rich structures in scattering amplitudes of string theories. Many of these findings can be attributed to the boost in computational reach due to the manifestly supersymmetric pure-spinor formalism [2–7]. For instance, pure-spinor methods enabled the first three-loop computation in the low-energy limit of the four-point closed-string amplitude [8] and gave rise to a strikingly compact form of n -point tree-level amplitudes [9,10].

The pure-spinor computation at tree level paved the way for string-theory realizations and extensions of recent unifying relations among field-theory amplitudes. For instance, using a local representation of the massless n -point disk correlation function, manifestly local numerators satisfying the color-kinematics duality [11] in gauge-theory amplitudes were systematically constructed [12]¹. Locality refers to the absence of kinematic poles in the superspace kinematic factors of the correlator: All the propagators in the color-kinematics dual gauge-theory amplitudes stem from the field-theory limit of the moduli-space integrals over disk worldsheets. *The first main goal of this series of papers is to generalize these results to loop level and to construct local correlators on a genus-one surface.*

Moreover, the n -point disk amplitudes [9] were later on found in [15] to share the structure of the Kawai–Lewellen–Tye (KLT) formula [16] for supergravity trees. In relating the open superstring to supergravity, one copy of the color-ordered gauge-theory trees in the KLT formula are mapped to so-called Parke–Taylor integrands $(z_{12}z_{23}\dots z_{n1})^{-1}$ with $z_{ij} = z_i - z_j$. This mapping rests on the fact that disk integrals of Parke–Taylor type share the Bern–Carrasco–Johansson (BCJ) relations [11] of gauge-theory tree amplitudes [15]. We will refer to these phenomena as a double-copy structure of disk amplitudes and a duality between kinematics and worldsheet functions. *The second main goal of this series of papers is to find a one-loop incarnation of the duality between kinematics and worldsheet functions that results in a double-copy structure of open-superstring amplitudes [17].*

¹ Also see [13,14] for a string-theory derivation of the resulting Bern–Carrasco–Johansson relations among color-ordered gauge-theory amplitudes at tree level.

While tree-level correlators are completely determined by their singularity structure encoded in the OPEs of the vertex operators, the quest for genus-one correlators is guided by additional constraints: The homology cycles of the genus-one surface translate into a notion of double-periodicity in its complex coordinate. As we will see in part II, double-periodicity does not hold term-by-term in the genus-one correlators. Instead, the monodromies of individual terms cancel in similar patterns as the BRST variations of kinematic factors in pure-spinor superspace². This will not only be a crucial guiding principle in constructing local representations of genus-one correlators in part III but also furnish a key incarnation of the duality between kinematics and worldsheet functions.

Apart from the parallels in their BRST- and monodromy variations, the kinematic building blocks and worldsheet functions in this work resonate in their symmetry properties under exchange of external legs. In a local form of the correlators, kinematic building blocks exhibit Lie-symmetries in several groups of labels, which translate into kinematic Jacobi relations in the tree-level subdiagrams of the field-theory limit [12,18,19]. The worldsheet functions of part II in turn are designed to vanish under shuffle products in several groups of labels which is reminiscent of the Kleiss–Kuijf relations of gauge-theory tree amplitudes [20]. Combinations of Lie- and shuffle symmetric objects are tailor-made to realize the permutation invariance of the correlators. At the same time, this symmetry structure is well known in the mathematics literature from a theorem by Ree [21] in the context of *Lie polynomials* [22]. Therefore we say that the local one-loop correlators of part III have a *Lie-polynomial* structure³.

We will also explore manifestly BRST-invariant but non-local representations of the correlators in part III, where the kinematic building blocks are dressed with the tree-level propagators. The resulting supersymmetric Berends–Giele currents [18,23] and their BRST-invariant combinations [24] realize shuffle symmetries on the kinematic side. Similarly, monodromy invariance of the correlators can be manifested by organizing the worldsheet functions into so-called *generalized elliptic integrands* (GEIs), see [17] and part II. In contrast to conventional elliptic functions, GEIs may involve loop momenta of the chiral-splitting formalism [25,26,27] that transform as well when punctures are taken along the homology cycle and cancel the monodromies of Jacobi theta functions.

² BRST-invariance of a kinematic factor in pure-spinor superspace implies its components to be both gauge invariant and supersymmetric [2].

³ Note that this same Lie-polynomial structure is already present in the calculation of the tree-level correlator in [9], but it remained unnoticed until now.

In part III, we will present expressions for $(n \leq 7)$ -point correlators in terms of BRST-invariant superfields and GEIs such that both kinds of invariances are manifest. Given that BRST-invariant superfields and GEIs are shown in part II to obey the same kinds of relations, the role of kinematics and worldsheet functions can be freely interchanged. This generalizes the $(n-3)!$ -term representations of tree-level correlators [9], where gauge-theory trees and Parke–Taylor factors enter on symmetric footing [15]. In analogy to these tree-level results, the one-loop amplitudes computed from such correlators are said exhibit a double-copy structure [17]. In the same way as the double-copy representations of gravitational loop integrands [28,29] hinges on the color-kinematics duality in gauge theories, the duality between kinematics and worldsheet function reveals a double-copy structure in one-loop open-superstring amplitudes.

A brief executive summary for the combined parts [1] of this series is as follows. In the first sections we will have self-contained discussions to set up preliminary notions regarding the pure-spinor formalism (section 2), local superfields (sections 3 and 4), non-local superfields (section 5) and one-loop worldsheet functions (section II.2). Several important relations and interplays among these first sections are pointed out and thoroughly illustrated too. For instance, section II.4 discusses several parallels and dualities between the non-local kinematic building blocks of section 5 and worldsheet functions that are built from the constituents in section II.2. Then, in section III.2, after a brief discussion pointing out the Lie-polynomial structure of the local n -point *tree-level* correlators, we will argue that also the local n -point *one-loop* correlators of the open superstring have a Lie-polynomial form. Namely,

$$\mathcal{K}_n(\ell) = \sum_{r=0}^{n-4} \frac{1}{r!} \left(V_{A_1} T_{A_2, \dots, A_{r+4}}^{m_1 \dots m_r} \mathcal{Z}_{A_1, \dots, A_{r+4}}^{m_1 \dots m_r} + [12 \dots n | A_1, \dots, A_{r+4}] \right) + \text{corrections}, \quad (1.1)$$

where V_A and $T_{A_1, \dots, A_{r+3}}^{m_1 \dots m_r}$ are local kinematic building blocks satisfying Lie symmetries while $\mathcal{Z}_{A_1, \dots, A_{r+4}}^{m_1 \dots m_r}$ are functions on the genus-one worldsheet satisfying shuffle symmetries (the unconventional notation for the permutations is explained in detail in the appendix III.A).

The need for “+ corrections” at $n \geq 7$ points will be elaborated in detail invoking e.g., locality, BRST invariance, single-valuedness and several other related technical aspects introduced in the first sections. In section III.3, a multitude of representations for the correlators with $n = 4, 5, 6, 7$ is presented. The $n = 8$ correlator is proposed and, while it satisfies many non-trivial constraints, it fails to be BRST invariant by terms proportional

to the holomorphic Eisenstein series G_4 . Unfortunately, this points to a certain weakness of our method since any Eisenstein series is a monodromy-invariant function with no dependence on the worldsheet punctures. (We leave it as an open challenge for future work to determine the kinematic coefficients of G_k in $(n \geq 8)$ -point correlators.) Further representations for the correlators are presented in section III.4, which is concerned with the explicit integration of the loop momentum in both open- and closed-string one-loop amplitudes. Several rather technical discussions are left to the appendices.

Finally, one should not be overwhelmed by the total number of pages of this series; the wide areas of both mathematics and physics that it touches lead to several relationships and beautiful connections. The final results for the correlators are in fact quite compact.

2. Basic formalism

In this section we will review certain aspects of the one-loop amplitude prescription in the minimal⁴ pure-spinor formalism [3]. In the later sections, this prescription will be used as a basis to formulate a general approach to assemble integrands for n -point open-superstring amplitudes at one loop from standard constraints such as single-valuedness and BRST invariance, among others.

2.0.1. Conventions

Throughout this work, we will use the shorthands

$$z_{ij} \equiv z_i - z_j, \tag{2.1}$$

for the worldsheet positions and

$$k_\emptyset^m \equiv 0, \quad k_{12\dots p}^m \equiv k_1^m + k_2^m + \dots + k_p^m, \quad s_{12\dots p} \equiv \frac{1}{2}k_{12\dots p}^2 = \sum_{i<j}^p s_{ij} \tag{2.2}$$

for multiparticle momenta and Mandelstam invariants, where $s_{ij} = (k_i \cdot k_j)$. Vector and spinor indices of the ten-dimensional Lorentz group are denoted by $m, n, \dots = 0, 1, \dots, 9$ and $\alpha, \beta, \dots = 1, 2, \dots, 16$, respectively.

⁴ See [7] for the one-loop amplitude prescription in the non-minimal pure-spinor formalism.

Our convention for (anti)symmetrizing r vector indices does *not* include a factor of $\frac{1}{r!}$ and it always generates unit coefficients for each inequivalent term, even in the presence of symmetric tensors; for instance $A^{(m}B^{n)} \equiv A^m B^n + A^n B^m$ as well as

$$\delta^{(mn}k^{p)} \equiv \delta^{mn}k^p + \delta^{mp}k^n + \delta^{np}k^m . \quad (2.3)$$

2.1. The pure-spinor amplitude prescription

The prescription to compute n -point one-loop amplitudes for open superstrings is [3]

$$\mathcal{A}_n = \sum_{\text{top}} C_{\text{top}} \int_{D_{\text{top}}} d\tau \langle\langle (\mu, b) \mathcal{Z} V_1(z_1) \prod_{j=2}^n \int dz_j U_j(z_j) \rangle\rangle \quad (2.4)$$

where the Beltrami differential μ and the modulus τ encode the topological information of the genus-one surface. The sum runs over all open-string worldsheet topologies at one loop: the planar and non-planar cylinder as well as the Möbius strip. For each topology the integration domain D_{top} for the modulus t and the color factors C_{top} have to be adjusted and the region of integration over the z_j variables must reflect the ordering of the vertex operators insertions on its boundaries [30], see section 2.1.1 below. Using translation invariance of the path integral, the position z_1 can be fixed to $z_1 = 0$, but it is customary to carry it unfixed in the formulas.

Moreover, the prescription (2.4) uses picture-changing operators⁵ collectively denoted by \mathcal{Z} and detailed in [3] as well as a composite b -ghost of schematic form

$$\begin{aligned} b = & (\Pi d + N\partial\theta + J\partial\theta) d\delta(N) + (w\partial\lambda + J\partial N + N\partial J + N\partial N)\delta(N) \quad (2.5) \\ & + (N\Pi + J\Pi + \partial\Pi + d^2)(\Pi\delta(N) + d^2\delta'(N)) \\ & + (Nd + Jd)(\partial\theta\delta(N) + d\Pi\delta'(N) + d^3\delta''(N)) \\ & + (N^2 + JN + J^2)(d\partial\theta\delta'(N) + \Pi^2\delta'(N) + \Pi d^2\delta''(N) + d^4\delta'''(N)) , \end{aligned}$$

where the worldsheet fields on the right-hand side will be introduced below. The complicated expression (2.5) poses difficulties in a direct evaluation of (2.4) at multiplicities above four, especially when OPE contractions involving the b -ghost $b(z_0)$ are considered. For the five-point correlator, these contributions were shown to be total worldsheet derivatives with respect to z_0 and therefore could be dropped in the integrated amplitude [31].

⁵ The ingredients of $\mathcal{Z} = Z_J \prod_{P=2}^{10} Z_{B_P} \prod_{I=1}^{11} Y_{C_I}$ are explained in [3].

However, starting at six-points, the OPE contributions involving the b -ghost may introduce non-trivial functions of the punctures into the correlator although the dependence on z_0 ultimately drops out⁶. The expressions for one-loop correlators to be proposed in the later sections provide evidence that such OPE contractions can be reduced to zero-mode contributions.

When all the external states are massless, the unintegrated vertex operator $V_i(z_i)$ of conformal weight zero is given by⁷

$$V_i(z_i) = \lambda^\alpha A_\alpha^i(x, \theta) , \quad (2.6)$$

while the integrated vertices $U_i(z_i)$ are

$$U_i(z_i) = \partial\theta^\alpha A_\alpha^i(x, \theta) + \Pi^m A_m^i(x, \theta) + d_\alpha W_i^\alpha(x, \theta) + \frac{1}{2} N^{mn} F_{mn}^i(x, \theta) . \quad (2.7)$$

The vectorial and spinorial polarizations of the i^{th} gluon and gluino as well as their light-like momenta k_i enter through the ten-dimensional super-Yang–Mills superfields $[A_\alpha^i, A_m^i, W_i^\alpha, F_i^{mn}]$ to be reviewed in section 3.

The bosonic pure spinor λ^α in (2.6) has conformal weight zero and obeys

$$(\lambda\gamma^m\lambda) = 0 , \quad (2.8)$$

where $\gamma_{\alpha\beta}^m = \gamma_{\beta\alpha}^m$ denote the 16×16 Pauli matrices of $SO(1,9)$ subject to the Clifford algebra $\gamma_{\alpha\beta}^{(m}\gamma^{n)\beta\gamma} = \delta^{mn}\delta\gamma_\alpha$. Note that the symmetry properties of antisymmetrized Pauli matrices are $\gamma_{\alpha\beta}^{mnp} = -\gamma_{\beta\alpha}^{mnp}$ and $\gamma_{\alpha\beta}^{mnpqr} = \gamma_{\beta\alpha}^{mnpqr}$ for odd rank as well as $\gamma^{mn}{}_\alpha{}^\beta = -\gamma^{mn}{}_\alpha{}^\beta$ and $\gamma^{mnpq}{}_\alpha{}^\beta = \gamma^{mnpq}{}_\alpha{}^\beta$ for even rank.

Integrated vertices (2.7) involve the Lorentz-current N^{mn} of the pure-spinor ghost, and $\Pi^m = \partial x^m + \frac{1}{2}(\theta\gamma^m\partial\theta)$ as well as d_α are supersymmetric combinations of the matter variables in the pure-spinor worldsheet action [2].

⁶ An explicit example of a related cancellation can be found in appendix B of [32] in the parity-odd sector of one-loop amplitudes in the RNS formalism.

⁷ For ease of notation, the dependence on z_i via $\lambda^\alpha(z), x^m(z), \theta^\alpha(z)$ as well as $\partial\theta^\alpha(z), \Pi^m(z), d_\alpha(z), N^{mn}(z)$ is left implicit on the right-hand side of (2.6), (2.7) and later equations.

2.1.1. Open-string integration domains

The vertex-operator locations or *punctures* need to be integrated over the torus or over the boundary components of the cylinder and the Möbius strip. For open strings, the integration over the boundaries has to match the cyclic orderings of the accompanying color factors: Each external state carries color degrees of freedom encoded in Lie-algebra⁸ generators t^a , and the color dependence of the amplitude enters via traces.

Given a torus of modular parameter τ in the parameterization of figure fig. 1, one can obtain the open-string worldsheets via suitable involutions [35,36], resulting in a purely imaginary modular parameter τ in case of the cylinder. The cylinder boundaries $\mathcal{B}_{1,2}$ will be taken to be the B -cycle through the origin and the point $z = \frac{1}{2}$, respectively,

$$\mathcal{B}_1 = \{z = i\nu, 0 \leq \nu \leq |\tau|\}, \quad \mathcal{B}_2 = \{z = \frac{1}{2} + i\nu, 0 \leq \nu \leq |\tau|\}, \quad (2.9)$$

i.e. they are separated by half the A -cycle. After using translation invariance to fix one puncture as $z_1 = 0$, the integration domain associated with the single trace $\text{tr}(t^1 t^2 \dots t^n)$ is characterized by $0 < \text{Im}(z_2) < \text{Im}(z_3) < \dots < \text{Im}(z_n) < \text{Im}(\tau)$, and similar choices can be made for the Möbius strip [35,36]. Double traces $\text{tr}(t^1 t^2 \dots t^j) \text{tr}(t^{j+1} \dots t^n)$ are exclusive to the cylinder topology, and one may define analogous domains where the two cyclic orderings are implemented on $\mathcal{B}_{1,2}$ in (2.9).

Additionally, the modular parameters need to be integrated, e.g. over $\tau \in i\mathbb{R}_+$ in case of the cylinder. As indicated in (2.4), the integrations over the different topologies along with their color traces are denoted by the subscript $_{\text{top}}$.

2.1.2. Functional integration and OPEs

The worldsheet fields $[\partial\theta^\alpha, \Pi^m, d_\alpha, N^{mn}]$ in the integrated vertices (2.7) have conformal weight +1 and can be integrated after separating off the zero modes. Using $d_\alpha(z)$ as an example, in a genus- g surface one gets

$$d_\alpha(z) = \sum_{I=1}^g d_\alpha^I \omega_I(z) + \hat{d}_\alpha(z), \quad (2.10)$$

where $\omega_I(z)$ are g holomorphic one-forms normalized as $\oint_{A_I} dz \omega_J(z) = \delta_{IJ}$ when integrated around the A -cycles. The non-zero modes $\hat{d}_\alpha(z)$ in turn are characterized by $\oint_{A_I} dz \hat{d}_\alpha(z) =$

⁸ For the type-I superstring, the gauge group has to be chosen as $SO(32)$ in order to guarantee cancellation of infinities [33] and gauge anomalies [34].

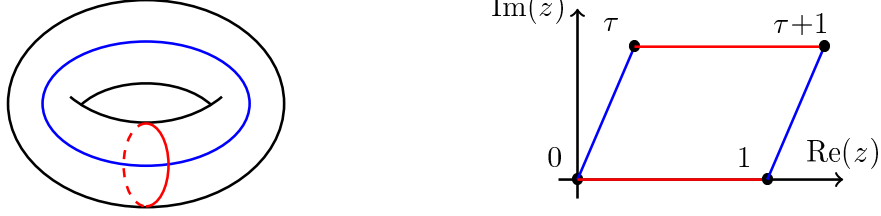


Fig. 1 Parameterization of the torus through the lattice $\mathbb{C}/(\mathbb{Z}+\tau\mathbb{Z})$ with an identification of points z with their translates $z+1$ and $z+\tau$ along the A - and B -cycle.

0. In addition, when the holomorphic one-forms are integrated around the B -cycles one gets the period matrix $\Omega_{IJ} = \oint_{B_I} dz \omega_J(z)$. Note that $\omega_1(z) = 1$ at genus one, and the parameterization of the torus in fig. 1 involves the period matrix $\oint_B dz = \tau$ with $\tau \in \mathbb{C}$ and $\text{Im } \tau > 0$.

The non-zero modes are functionally integrated through OPE contractions, in particular at genus one we have

$$\begin{aligned}
\hat{d}_\alpha(z)K(0) &\rightarrow D_\alpha K(0) g^{(1)}(z, \tau), & \hat{\Pi}^m(z)K(0) &\rightarrow -k^m K(0) g^{(1)}(z, \tau), \\
\hat{d}_\alpha(z)\hat{\Pi}^m(0) &\rightarrow (\gamma^m \partial\theta(0))_\alpha g^{(1)}(z, \tau) & \hat{d}_\alpha(z)\hat{d}_\beta(0) &\rightarrow -\gamma_{\alpha\beta}^m \Pi_m(0) g^{(1)}(z, \tau), \\
\hat{d}_\alpha(z)\theta^\beta(0) &\rightarrow \delta_\alpha^\beta g^{(1)}(z, \tau) & \hat{N}^{mn}(z)\lambda^\alpha(0) &\rightarrow -\frac{1}{2}(\lambda(0)\gamma^{mn})^\alpha g^{(1)}(z, \tau),
\end{aligned} \tag{2.11}$$

where the worldsheet singularities are captured by [3] ($\partial \equiv \frac{\partial}{\partial z}$)

$$g^{(1)}(z, \tau) \equiv \partial \log \theta_1(z, \tau), \tag{2.12}$$

and the standard odd Jacobi theta function with $q \equiv \exp(2\pi i\tau)$ is given by

$$\theta_1(z, \tau) \equiv 2q^{1/8} \sin(\pi z) \prod_{n=1}^{\infty} (1 - q^n)(1 - q^n e^{2\pi iz})(1 - q^n e^{-2\pi iz}). \tag{2.13}$$

In the above OPEs, $K(z)$ represents a generic superfield which depends on $\theta^\alpha(z)$ but not on any derivative $\partial^n \theta^\alpha(z)$ for $n \geq 1$, and whose x dependence is entirely contained in the plane-wave factor⁹ $e^{k \cdot x}$. The functional integration of the variables $x^m(z, \bar{z})$ gives rise to the so-called Koba–Nielsen factor and will be reviewed in the next subsection. The above OPEs can be read off from their tree-level counterparts [37,2] via the substitution $\frac{1}{z} \rightarrow g^{(1)}(z, \tau)$.

⁹ To avoid factors of i in the formulae, we depart from the standard conventions of real momenta and redefined $ik^m \rightarrow k^m$.

It turns out that the four-point amplitude computed with the pure-spinor formalism does not involve any OPE contraction; its outcome is based purely on zero-mode integrations [3]. Therefore, the n -point amplitude admits at most $n-4$ OPE contractions, and it will be explained in the later sections II.7 and III.4 that this gives rise to modular forms of weight $n-4$ after integrating over the zero modes of Π^m in (2.4).

2.1.3. Zero-mode integrations

The non-zero modes of the worldsheet fields are integrated out using OPE contractions (2.11), and a systematic procedure to capture the resulting tensor structures will be reviewed in section 3. Similarly, as in the tree-level correlator, we will assume that the OPE residues which feature double poles have been absorbed into single-pole residues using integration-by-parts identities [9]. The net effect of such manipulations can be accounted for via multiparticle superfields in the BCJ gauge [23] to be reviewed in subsection 3.4.

Once all non-zero modes are integrated out in that manner, the correlator (2.4) will depend only on the zero modes of the worldsheet variables $[d_\alpha(z), N^{mn}(z), \Pi^m(z)]$ and $[\lambda^\alpha(z), \theta^\alpha(z)]$. In this section we outline a practical procedure to integrate them using the pure-spinor zero-mode measures defined in [3]. Unlike the other worldsheet fields, the zero mode of $\Pi^m(z)$ is denoted with a different symbol ℓ^m , called the *loop momentum*. Its integration will be discussed in the context of *chiral splitting*, see subsection 2.2.

As explained in [3], in performing the path integral of the prescription (2.4) the role of the picture-changing operators \mathcal{Z} is to ensure that all bosonic and fermionic zero modes are absorbed correctly. In doing this it will be convenient to start integrating out d_α and N^{mn} while leaving the zero modes of Π^m , λ^α and θ^α to be dealt with at later stages. The reason for this is that the integrations over d_α and N^{mn} can be performed, under mild assumptions, using group-theory arguments alone. This will lead to effective rules which will then be used as input on section 4 to define local kinematic building blocks based on the multiparticle SYM superfields of section 3.

To see how this comes about, the first thing to note is that in integrating out the zero modes d_α and N^{mn} using the pure-spinor measures of [3] one introduces two pure spinors λ^α into the rest of the path integral. In addition, since the picture-changing operators contribute a fixed number of ten fermionic d_α zero modes, an additional six zero modes must come from the b -ghost and the external vertices (2.7). We will consider the contributions from two classes of terms in the b -ghost (2.5), given by $b^{(2)} \equiv \Pi d^2 \delta(N)$ and $b^{(4)} \equiv d^4 \delta'(N)$.

When the b -ghost contributes $b^{(4)} \equiv d^4\delta'(N)$, the external vertices must provide the remaining two d_α zero modes. However, there must also be a N^{mn} zero mode due to the factor of $\delta'(N)$ [3]. Therefore, the non-vanishing configuration of zero modes from the external vertices must be proportional to $d_\alpha d_\beta N^{mn}$. Given the expression (2.7) for integrated vertex operators, this contribution is of the schematic form $U_2 U_3 U_4 \rightarrow d_\alpha d_\beta N_{mn} W_2^\alpha W_3^\beta F_4^{mn}$, see section 3.4 for multiparticle generalizations. Taking into account the introduction of two pure spinors from the pure-spinor measures, the integration of the factor $d_\alpha d_\beta N_{mn} W_2^\alpha W_3^\beta F_4^{mn}$ from the external vertices can be summarized as

$$\int d^4\delta'(N) d_\alpha d_\beta N^{mn} \rightarrow (\lambda\gamma^{[m})_\alpha (\lambda\gamma^{n]})_\beta, \quad (2.14)$$

where the integral sign represents the integration using the zero-mode measures of [3]. Up to an overall coefficient, this is the unique outcome because there is only one two-form irreducible representation in the tensor product of two pure spinors and two Weyl spinors¹⁰

$$[0, 0, 0, 0, 2] \otimes [0, 0, 0, 0, 1]^{\wedge 2} = [0, 1, 0, 0, 0] + \dots \quad (2.15)$$

Hence, the net contribution from this sector to the correlator is given by a unique Lorentz-scalar combination of superfields

$$\int b^{(4)} U_2 U_3 U_4 \rightarrow (\lambda\gamma_m W_2)(\lambda\gamma_n W_3) F_4^{mn} + \text{cyc}(2, 3, 4). \quad (2.16)$$

Similarly, when the b -ghost contribution comes from the term $b^{(2)} \equiv \Pi d^2\delta(N)$, the external vertices must provide four d_α zero modes, and this time there is no need for an additional N^{mn} . Therefore, $b^{(2)}$ requires the external vertices to contribute $d_\alpha d_\beta d_\gamma d_\delta W_2^\alpha W_3^\beta W_4^\gamma W_5^\delta$, see section 3.4 for multiparticle generalizations. One can show that, up to an overall coefficient, the effective rule for integrating these zero modes is given by

$$\int \Pi d^2\delta(N) d_\alpha d_\beta d_\gamma d_\delta \rightarrow \ell_m (\lambda\gamma^a)_{[\alpha} (\lambda\gamma^b)_{\beta} (\gamma^{abm})_{\gamma\delta]}. \quad (2.17)$$

The argument to see this is similar to (2.15): there is a single vector representation in the tensor product of $\lambda^2 W^4$,

$$[0, 0, 0, 0, 2] \otimes [0, 0, 0, 0, 1]^{\wedge 4} = [1, 0, 0, 0, 0] + \dots \quad (2.18)$$

¹⁰ We acknowledge the use of the LiE program in these decompositions [38].

Therefore the unique Lorentz-invariant overall contribution from this zero-mode sector can be summarized by the following superfield combination

$$\int b^{(2)} U_2 U_3 U_4 U_5 \rightarrow \ell_m (\lambda \gamma_a W_2) (\lambda \gamma_b W_3) (W_4 \gamma^{abm} W_5) + \text{perm}(2, 3, 4, 5). \quad (2.19)$$

In section 4 we will see how (2.19) motivates the introduction of tensorial local building blocks that capture the kinematics of one-loop correlators.

The above rules are readily generalized for additional instances of zero modes of Π^m ,

$$\int \Pi^{m_1} \Pi^{m_2} \dots \Pi^{m_r} d^4 \delta'(N) d_\alpha d_\beta N^{pq} \rightarrow \ell^{m_1} \ell^{m_2} \dots \ell^{m_r} (\lambda \gamma^{[p})_\alpha (\lambda \gamma^{q]})_\beta, \quad (2.20)$$

$$\int \Pi^{m_1} \dots \Pi^{m_r} \Pi d^2 \delta(N) d_\alpha d_\beta d_\gamma d_\delta \rightarrow \ell^{m_1} \dots \ell^{m_r} \ell_n (\lambda \gamma^a)_{[\alpha} (\lambda \gamma^b)_{\beta} (\gamma^{abn})_{\gamma \delta]}. \quad (2.21)$$

The analogues of (2.16) and (2.19) for the remaining terms of the b -ghost (2.5) besides $b^{(2)}$ and $b^{(4)}$ are currently out of reach. Instead, we will infer their contributions to one-loop correlators from first principles to be detailed in section III.2.4. Up to integration-by-parts equivalent terms, $b^{(2)}$ and $b^{(4)}$ provide the highest numbers of zero modes of d_α, N^{mn} and therefore start to contribute at the lowest multiplicities. Using these zero-mode considerations it follows that the loop integrand for n -point open-string amplitudes (2.4) is a *polynomial* of degree $n-4$ in the loop momentum ℓ .

2.1.4. Pure-spinor superspace

The angle brackets $\langle\langle \dots \rangle\rangle$ in the amplitude prescription (2.4) represent the complete path integral over all the worldsheet degrees of freedom. After integrating the zero modes of d_α, N^{mn} and all the other variables except for λ^α and θ^α , these $\langle\langle \dots \rangle\rangle$ are replaced by $\langle \dots \rangle$ which represent the remaining functional integration over zero modes of λ^α and θ^α . In integrating the variables in this order, the kinematic factors become expressions in *pure-spinor superspace* as defined in [39]. Pure-spinor superspace compactly encodes all states in the supermultiplet, and the components can be extracted using the prescription [2]

$$\langle (\lambda \gamma^m \theta) (\lambda \gamma^n \theta) (\lambda \gamma^p \theta) (\theta \gamma_{mnp} \theta) \rangle = 2880 \quad (2.22)$$

for integration over zero modes of λ^α and θ^α . In fact, the amplitudes exhibit their most convenient form when written in pure-spinor superspace, i.e. *without* performing the integration in (2.22), and will be represented as such in this series of papers.

A key feature of the measure (2.22) is its interplay with the BRST operator

$$Q \equiv \lambda^\alpha D_\alpha, \quad D_\alpha \equiv \frac{\partial}{\partial \theta^\alpha} + \frac{1}{2}(\gamma^m \theta)_\alpha \frac{\partial}{\partial x^m}. \quad (2.23)$$

As pointed out in [2], the measure (2.22) is only sensitive to the cohomology of Q : BRST-closed superfields $QS(x, \theta, \lambda) = 0$ are mapped to gauge invariant and supersymmetric components in $\langle S(x, \theta, \lambda) \rangle$, whereas BRST-exact terms decouple, i.e. $\langle E(x, \theta, \lambda) \rangle = 0$ if $E(x, \theta, \lambda) = Q\Sigma(x, \theta, \lambda)$. This cohomology structure can be exploited to obtain non-trivial relations among seemingly different superspace expressions including amplitudes at different loop orders [40].

2.2. Chiral splitting of the Koba–Nielsen factor

The zero-mode integrations of the matter variables $x^m(z, \bar{z})$ or equivalently $\Pi^m(z)$ is performed employing the techniques of the *chiral-splitting* formalism of [25,26,27]. The idea is to defer the zero-mode integration for Π^m within the path integral in (2.4) to the last step of the amplitude computation¹¹ and to interpret it as a string-theory antecedent of the loop momentum in Feynman integrals, to be denoted by

$$\ell^m \equiv \oint_A dz \Pi^m(z). \quad (2.24)$$

In this setting, the contributions from the plane-wave factors $e^{k \cdot x}$ can be reproduced from the *Koba–Nielsen factor*

$$\mathcal{I}_n(\ell) \equiv \exp \left(\sum_{i < j}^n s_{ij} \log \theta_1(z_{ij}, \tau) + \sum_{j=1}^n z_j(\ell \cdot k_j) + \frac{\tau}{4\pi i} \ell^2 \right), \quad (2.25)$$

and our notation $\mathcal{I}_n(\ell)$ for the Koba–Nielsen factor omits its dependence on the variables z_j, k_j, τ . Chiral splitting can be easily undone:

In a closed-string context, the loop integration comprised by the path integral $\langle\langle \dots \rangle\rangle_{\text{closed}}$ over left- and right movers reproduces the more conventional and modular invariant form of the Koba–Nielsen factor,

$$\begin{aligned} \hat{\mathcal{I}}_n &= \left\langle\left\langle \prod_{j=1}^n e^{k_j \cdot x(z_j, \bar{z}_j)} \right\rangle\right\rangle_{\text{closed}} = \int d^D \ell |\mathcal{I}_n(\ell)|^2 \\ &= \frac{(2\pi i)^D}{(2 \operatorname{Im} \tau)^{\frac{D}{2}}} \exp \left(\sum_{i < j}^n s_{ij} \left[\log |\theta_1(z_{ij}, \tau)|^2 - \frac{2\pi}{\operatorname{Im} \tau} (\operatorname{Im} z_{ij})^2 \right] \right). \end{aligned} \quad (2.26)$$

¹¹ More formally, chiral splitting is implemented by inserting the integrated delta function $1 = \int d^D \ell \delta^D(\ell^m - \oint_A dz \Pi^m(z))$ into the path integral.

On the one hand, the combination $\log |\theta_1(z, \tau)|^2 - \frac{2\pi}{\text{Im } \tau} (\text{Im } z)^2$ in the exponent exhibits double periodicity under translations $z \rightarrow z + 1$ and $z \rightarrow z + \tau$ around the homology cycles of the Riemann surface. On the other hand, its second term $\sim \frac{(\text{Im } z_{ij})^2}{\text{Im } \tau}$ obstructs holomorphic factorization of the moduli-space integrand for closed-string amplitudes.

In an open-string context, the path integral $\langle\langle \dots \rangle\rangle$ in (2.4) only comprises half the non-zero modes of x^m as compared to its closed-string counterpart $\langle\langle \dots \rangle\rangle_{\text{closed}}$ in (2.26). Accordingly, the plane-wave correlator of the open string yields half of the Koba–Nielsen exponent,

$$\begin{aligned} \hat{\mathcal{I}}_n^{\text{open}} &= \left\langle\left\langle \prod_{j=1}^n e^{k_j \cdot x(z_j, \bar{z}_j)} \right\rangle\right\rangle = \int d^D \ell |\mathcal{I}_n(\ell)| \\ &= \frac{(2\pi i)^D}{(\text{Im } \tau)^{\frac{D}{2}}} \exp \left(\sum_{i < j} s_{ij} \left[\log |\theta_1(z_{ij}, \tau)| - \frac{\pi (\text{Im } z_{ij})^2}{\text{Im } \tau} \right] \right), \end{aligned} \quad (2.27)$$

while the loop integration is the same as in (2.26) since the loop momentum is a joint zero mode of left- and right movers in $\langle\langle \dots \rangle\rangle_{\text{closed}}$. The purpose of writing the last line of (2.27) in terms of $\text{Im } z_j$ and $\text{Im } \tau$ is to have a universal expression for all the topologies of open-string amplitudes in (2.4): While the punctures and modulus of a planar cylinder diagram are accounted for by purely imaginary choices of z_j, τ , the non-planar cylinder and the Möbius strip also introduce real parts for some of z_j, τ , see e.g. [35,36]. Still, one has to keep in mind that there is no distinction between holomorphic and antiholomorphic variables in an open-string setup when taking total derivatives of the Koba–Nielsen factor (2.27). Accordingly, open and closed strings give rise to the same equivalence classes of correlators with respect to total-derivative relations as discussed in section II.2.3.

Zero-mode integrations at multiplicities higher than four require generalizations of (2.26) and (2.27) and will be discussed in sections II.7 and III.4.

2.2.1. Definition of open-string correlators

The main challenge to be addressed in this work is to determine the dependence of the open-string amplitude (2.4) on the polarizations and momenta. The universal Koba–Nielsen factor (2.25) due to plane waves will be stripped off from

$$\langle\langle (\mu, b) \mathcal{Z} V_1(z_1) \prod_{j=2}^n U_j(z_j) \rangle\rangle = \int d^D \ell |\mathcal{I}_n(\ell)| \langle \mathcal{K}_n(\ell) \rangle. \quad (2.28)$$

The residual task is to identify kinematic factors $\mathcal{K}_n(\ell)$ in pure-spinor superspace that depend on the loop momentum as well as the zero modes of $\lambda^\alpha, \theta^\alpha$ and capture the superfield kinematics arising from the path integral. Given their origin from integrating out all the non-zero modes as well as the zero modes of d_α and N^{mn} , we will henceforth refer to these kinematic factors $\mathcal{K}_n(\ell)$ as *correlators*. They carry the key information on the amplitudes

$$\mathcal{A}_n = \sum_{\text{top}} C_{\text{top}} \int_{D_{\text{top}}} d\tau dz_2 dz_3 \dots dz_n \int d^D \ell |\mathcal{I}_n(\ell)| \langle \mathcal{K}_n(\ell) \rangle, \quad (2.29)$$

and the computational methods and organizing principles for correlators to be developed in this work are tailored to reveal hidden double-copy structures.

2.2.2. Closed-string correlators and amplitudes

By virtue of chiral splitting, left- and right-moving worldsheet degrees of freedom completely decouple at the level of the loop integrand, and closed-string correlators are obtained from holomorphic squares of their open-string instances. More precisely, the n -point closed-string amplitude reads

$$\mathcal{M}_n = \int_{\mathcal{F}} d^2 \tau d^2 z_2 d^2 z_3 \dots d^2 z_n \int d^D \ell |\mathcal{I}_n(\ell)|^2 \langle \mathcal{K}_n(\ell) \rangle \langle \tilde{\mathcal{K}}_n(-\ell) \rangle, \quad (2.30)$$

where \mathcal{F} denotes the fundamental domain of the modular group $SL_2(\mathbb{Z})$ and the punctures z_j are integrated over a torus of modular parameter τ . The reflection $\ell \rightarrow -\ell$ in the right-moving correlator is due to our normalization conventions for external momenta. Finally, the tilde in $\tilde{\mathcal{K}}_n$ instructs to replace the super-Yang–Mills superfields $[A_\alpha^i, A_m^i, W_i^\alpha, F_i^{mn}]$ by another copy, where the Weyl spinors have the same (opposite) chirality for type-IIB (type-IIA) superstrings.

In situations where both $\mathcal{K}_n(\ell)$ and $\tilde{\mathcal{K}}_n(-\ell)$ depend on ℓ , we will see in section III.4 that quadratic and higher terms in the loop momentum introduces vector contractions between left- and right-moving superfields proportional to $\pi/(\text{Im } \tau)$, see e.g. [41,42] and [43,44] in cases of maximal and reduced supersymmetry, respectively. This exemplifies how the double-copy structure of the closed-string integrand in (2.30) disappears after performing the loop integration [45]: While $\langle \mathcal{K}_n(\ell) \rangle \langle \tilde{\mathcal{K}}_n(-\ell) \rangle$ is evidently a holomorphic square of open-string correlators, its loop integral over $\int d^D \ell |\mathcal{I}_n(\ell)|^2$ no longer factorizes. That is why chiral splitting is a convenient framework to study the double-copy properties of gravity amplitudes from a string-theory perspective.

By the appearance of open-string correlators \mathcal{K}_n in closed-string amplitudes (2.30), they need to be well-defined functions on the torus, at least after integration over ℓ . In particular – after stripping off a global factor of $(\text{Im } \tau)^{-5}$ – the loop integral over $|\mathcal{I}_n|^2 \mathcal{K}_n$ and $|\mathcal{I}_n|^2 \mathcal{K}_n \tilde{\mathcal{K}}_n$ must have modular weight $(n-4, 0)$ and $(n-4, n-4)$, respectively.

3. Multiparticle SYM superfields

After introducing a convenient notation we review the recursive construction of multiparticle super-Yang–Mills superfields of [18]. A special emphasis will be given to their *local* representatives, as they will play an essential role in the construction of one-loop correlators in later sections.

3.1. Combinatorics on words

Let us first introduce a notation based on words and review a few associated results that will be used in the rest of this work. Good introductions to the combinatorics on words and related subjects can be found in [46,22].

In dealing with objects that contain multiple particle labels, one is faced with many permutations and associated operations acting on the labels of the participating particles. A convenient framework to handle such things is to use the notion of words and linear maps acting on them. As such, permutations¹² of particle labels referring to the external legs are encoded in *words* composed from *letters* in the *alphabet* of natural numbers; $\{1, 2, 3, \dots\}$.

Words will be written in upper-case (e.g. $P = 134256$) and its letters in lower-case (e.g. $i = 3$). The *length* of the word $P = p_1 p_2 \dots p_n$ is the number n of its letters and is denoted by $|P|$. The reversal of the word $P = p_1 p_2 \dots p_n$ is the word $\tilde{P} = p_n \dots p_2 p_1$.

The *concatenation* product of the words $P = p_1 \dots p_n$ and $Q = q_1 \dots q_m$ is the word $PQ = p_1 \dots p_n q_1 \dots q_m$. The *empty word* is denoted by \emptyset and it is the unit with respect to the concatenation, i.e. $P\emptyset = \emptyset P = P$. Unless otherwise noted, labeled objects are defined to be zero when their label is the empty word (such as the momentum $k_\emptyset^m \equiv 0$).

The *deconcatenation* of a word P into two words is denoted $P = XY$ and is given by all pairs of words X, Y such that $P = XY$ under concatenation (with obvious generalization for $P = XYZ$ etc). For example, the deconcatenation of $P = XY$ when $P = 312$ is given by the the words (X, Y) : $(\emptyset, 312)$, $(3, 12)$, $(31, 2)$ and $(312, \emptyset)$. The deconcatenation map often occurs as a summation condition, e.g.

$$T_P = \sum_{P=XY} F_X F_Y \quad \Longrightarrow \quad T_{312} = F_\emptyset F_{312} + F_3 F_{12} + F_{31} F_2 + F_{312} F_\emptyset \quad (3.1)$$

for arbitrary labeled objects T and F . The *shuffle product* of words of length n and m is defined recursively by

$$\emptyset \sqcup A = A \sqcup \emptyset = A, \quad A \sqcup B \equiv a_1(a_2 \dots a_n \sqcup B) + b_1(b_2 \dots b_m \sqcup A), \quad (3.2)$$

¹² Words with repeated letters do not appear in the context of scattering amplitudes.

and it generates all $\frac{(n+m)!}{n!m!}$ possible ways to interleave the letters of A and B without changing their orderings within A and B . For example,

$$\begin{aligned} 1 \sqcup 2 &= 12 + 21, & 12 \sqcup 3 &= 123 + 132 + 312, \\ 12 \sqcup 34 &= 1234 + 1324 + 1342 + 3142 + 3124 + 3412. \end{aligned} \quad (3.3)$$

A word P is said to be a *shuffle* of X and Y if it appears in their shuffle product, i.e. if $P \in X \sqcup Y$. From the examples (3.3) it follows that 3142 is a shuffle of 12 and 34.

The *deshuffle* of P is denoted $P = X \sqcup Y$ and is the sum of all pairs of words X, Y such that P is a shuffle of X and Y . An efficient algorithm that generates X, Y in the deshuffle of P follows from the linear map $\delta(P) = X \otimes Y$ defined by

$$\delta(a_1 a_2 \dots a_n) \equiv \delta(a_1) \delta(a_2) \dots \delta(a_n), \quad \delta(a_i) \equiv \emptyset \otimes a_i + a_i \otimes \emptyset, \quad \delta(\emptyset) \equiv \emptyset \otimes \emptyset. \quad (3.4)$$

For example,

$$\begin{aligned} \delta(1) &= \emptyset \otimes 1 + 1 \otimes \emptyset, \\ \delta(12) &= \delta(1) \delta(2) = (\emptyset \otimes 1 + 1 \otimes \emptyset)(\emptyset \otimes 2 + 2 \otimes \emptyset) = \emptyset \otimes 12 + 1 \otimes 2 + 2 \otimes 1 + 12 \otimes \emptyset, \\ \delta(123) &= \delta(12) \delta(3) = (\emptyset \otimes 12 + 1 \otimes 2 + 2 \otimes 1 + 12 \otimes \emptyset)(\emptyset \otimes 3 + 3 \otimes \emptyset) \\ &= \emptyset \otimes 123 + 1 \otimes 23 + 2 \otimes 13 + 12 \otimes 3 + 3 \otimes 12 + 13 \otimes 2 + 23 \otimes 1 + 123 \otimes \emptyset. \end{aligned} \quad (3.5)$$

An alternative characterization is $\delta(P) = \sum_{X,Y} \langle P, X \sqcup Y \rangle X \otimes Y$ where $\langle \cdot, \cdot \rangle$ denotes the scalar product on words defined by

$$\langle A, B \rangle \equiv \delta_{A,B}, \quad \delta_{A,B} = \begin{cases} 1, & \text{if } A = B \\ 0, & \text{otherwise} \end{cases}. \quad (3.6)$$

We will see in section 3.4.4 that the deshuffle coproduct describes the BRST variation of local multiparticle superfields just like the deconcatenation describes the BRST variation of their non-local counterparts.

As words are restricted to be permutations of the letters $\{1, 2, 3, \dots\}$, an explicit sum over permutations is often represented by a sum over words, e.g.

$$\sum_P T_P \equiv \sum_{|P|=1}^{\infty} \sum_{\alpha \in \{p_1, \dots, p_{|P|}\}} T_\alpha. \quad (3.7)$$

Furthermore, two common operations on words are given by the left-to-right bracketing map $\ell(A)$ and the rho map $\rho(A)$. They are defined recursively as [22]

$$\ell(123\dots n) \equiv \ell(123\dots n-1)n - n\ell(123\dots n-1), \quad \ell(i) \equiv i, \quad \ell(\emptyset) \equiv 0, \quad (3.8)$$

$$\rho(123\dots n) \equiv 1\rho(23\dots n) - n\rho(123\dots n-1), \quad \rho(i) \equiv i, \quad \rho(\emptyset) \equiv 0, \quad (3.9)$$

for example $\ell(123) = 123 - 132 - 231 + 321$ and $\rho(123) = 123 - 132 - 312 + 321$. In sections 3.4 and 5.1 these maps will be used, among other applications, in the discussion of superfields in the BCJ gauge and as a practical prescription to convert non-local Berends–Giele currents into their local counterparts. There is a vast literature dealing with these and similar maps in the context of free Lie algebras, see for instance [22].

In addition, unless otherwise noted every labeled object considered in this series of papers is linear on words, e.g. $T_{A+B} \equiv T_A + T_B$. This linearity will be frequently exploited to avoid unnecessary summation symbols, for instance

$$T_{A \sqcup B} \equiv \sum_{\sigma \in A \sqcup B} T_\sigma, \quad T_{\ell(A)} \equiv \sum_{\sigma \in \ell(A)} T_\sigma. \quad (3.10)$$

To further illustrate the above points, the Kleiss–Kuijf amplitude relations [20] among Yang–Mills tree amplitudes become $A_{P1Qn} = (-1)^{|P|} A_{1(\tilde{P} \sqcup Q)n}$, while the symmetry [47,23] obeyed by the Berends–Giele currents [48] is written as $\mathcal{K}_{A \sqcup B} = 0$.

In this work we use the convention that whenever words of external-state labels in a subscript are separated through a comma (rather than a vertical bar), the parental object is understood to be symmetric under exchange of these words. For example,

$$T_{A,B,C} = T_{A,C,B} = T_{B,A,C}. \quad (3.11)$$

In addition to denoting a sum over permutations with standard notations such as

$$T_A T_{B,C,D} + (A \leftrightarrow B, C, D) \equiv T_A T_{B,C,D} + T_B T_{A,C,D} + T_C T_{A,B,D} + T_D T_{A,B,C}, \quad (3.12)$$

more general permutations will be handled with the notation $+(A, B|A, B, C, D)$; it instructs to sum over all ordered combinations of the words A, B taken from the set $\{A, B, C, D\}$, for example

$$\begin{aligned} T_A T_B T_{C,D} + (A, B|A, B, C, D) &\equiv T_A T_B T_{C,D} + T_A T_C T_{B,D} + T_A T_D T_{B,C} \\ &+ T_B T_C T_{A,D} + T_B T_D T_{A,C} + T_C T_D T_{A,B}. \end{aligned} \quad (3.13)$$

Generalizations of the form $+(A_1, \dots, A_n|A_1, \dots, A_{n+m})$ for a total number $\binom{n+m}{n}$ of terms follow similarly.

3.2. Single-particle

A ten-dimensional covariant description of the SYM equations of motion makes use of four types of superfields seen in the vertex operators (2.6) and (2.7): the gluino and gluon potentials $A_\alpha(x, \theta)$, $A^m(x, \theta)$ and their field-strengths $W^\alpha(x, \theta)$, and $F^{mn}(x, \theta)$. They satisfy the following linearized equations of motion [49,50]

$$\begin{aligned} D_\alpha A_\beta + D_\beta A_\alpha &= \gamma_{\alpha\beta}^m A_m & D_\alpha A_m &= (\gamma_m W)_\alpha + \partial_m A_\alpha \\ D_\alpha F_{mn} &= \partial_{[m} (\gamma_{n]} W)_\alpha & D_\alpha W^\beta &= \frac{1}{4} (\gamma^{mn})_\alpha{}^\beta F_{mn} , \end{aligned} \quad (3.14)$$

see (2.23) for the supersymmetric derivative D_α in $D = 10$ superspace¹³.

We will use the collective notation

$$K_i \in \{A_\alpha^i(x, \theta), A_i^m(x, \theta), W_i^\alpha(x, \theta), F_i^{mn}(x, \theta)\} \quad (3.15)$$

for the four types of superfields describing the i^{th} external leg in an open-string amplitude (2.29). The superfields K_i will be referred to as *single-particle* superfields.

3.3. Two-particle

The vertex operators (2.6) and (2.7) for massless states in the pure-spinor formalism are expanded in terms of single-particle superfields. The computation of OPEs among the above vertex operators as required by the CFT amplitude prescription in the pure-spinor formalism leads to a natural definition of *multiparticle superfields* [18]. In contrast to the standard description of (3.14), these superfields encompass more than a single particle label. For example, after absorbing the double poles into total derivatives, the single pole in the OPE $U_1(z_1)U_2(z_2)$ can be written as [51]

$$U_{12} \equiv \partial\theta^\alpha A_\alpha^{12} + \Pi^m A_m^{12} + d_\alpha W_{12}^\alpha + \frac{1}{2} N^{mn} F_{mn}^{12} , \quad (3.16)$$

where the two-particle superfields are given by

$$\begin{aligned} A_\alpha^{12} &= -\frac{1}{2} [A_\alpha^1(k^1 \cdot A^2) + A_m^1 (\gamma^m W^2)_\alpha - (1 \leftrightarrow 2)] , \\ A_m^{12} &= \frac{1}{2} [A_p^1 F_{pm}^2 - A_m^1 (k^1 \cdot A^2) + (W^1 \gamma_m W^2) - (1 \leftrightarrow 2)] , \\ W_{12}^\alpha &= \frac{1}{4} (\gamma^{mn} W^2)^\alpha F_{mn}^1 + W_2^\alpha (k^2 \cdot A^1) - (1 \leftrightarrow 2) , \\ F_{mn}^{12} &= k_m^{12} A_n^{12} - k_n^{12} A_m^{12} - (k_1 \cdot k_2) (A_m^1 A_n^2 - A_n^1 A_m^2) . \end{aligned} \quad (3.17)$$

¹³ We will freely swap $k_m \leftrightarrow \partial_m$ without warning due to our convention $ik_m \rightarrow k_m$.

The last line involves the two-particle momentum $k_{12}^m = k_1^m + k_2^m$, see (2.2) for the definition of multiparticle momenta. By virtue of (3.14), one can check that the two-particle superfields (3.17) satisfy the following equations of motion:

$$\begin{aligned}
D_{(\alpha} A_{\beta)}^{12} &= \gamma_{\alpha\beta}^m A_m^{12} + (k_1 \cdot k_2)(A_\alpha^1 A_\beta^2 + A_\beta^1 A_\alpha^2), \\
D_\alpha A_m^{12} &= (\gamma_m W^{12})_\alpha + k_m^{12} A_\alpha^{12} + (k_1 \cdot k_2)(A_\alpha^1 A_m^2 - A_\alpha^2 A_m^1), \\
D_\alpha W_{12}^\beta &= \frac{1}{4}(\gamma^{mn})_\alpha{}^\beta F_{mn}^{12} + (k_1 \cdot k_2)(A_\alpha^1 W_2^\beta - A_\alpha^2 W_1^\beta), \\
D_\alpha F_{mn}^{12} &= k_m^{12}(\gamma_n W^{12})_\alpha - k_n^{12}(\gamma_m W^{12})_\alpha + (k_1 \cdot k_2)(A_\alpha^1 F_{mn}^2 - A_\alpha^2 F_{mn}^1) \\
&\quad + (k_1 \cdot k_2)(A_n^1(\gamma_m W^2)_\alpha - A_n^2(\gamma_m W^1)_\alpha - A_m^1(\gamma_n W^2)_\alpha + A_m^2(\gamma_n W^1)_\alpha),
\end{aligned} \tag{3.18}$$

which augment the linearized equations of motion in (3.14) by contact terms $\sim k_1 \cdot k_2$. Up to BRST exact terms [18], the two-particle version

$$V_{12} \equiv \lambda^\alpha A_\alpha^{12} \tag{3.19}$$

of the unintegrated vertex operator also appears in the OPE $V_1(z_1)U_2(z_2)$. Written in terms of the BRST charge $Q = \lambda^\alpha D_\alpha$, the equations of motion (3.18) become

$$\begin{aligned}
QV_{12} &= (k_1 \cdot k_2)V_1V_2, \\
QA_{12}^m &= (\lambda\gamma_m W_{12}) + k_{12}^m V_{12} + (k_1 \cdot k_2)(V_1 A_2^m - V_2 A_1^m), \\
QW_{12}^\beta &= \frac{1}{4}(\lambda\gamma_{mn})^\beta F_{12}^{mn} + (k_1 \cdot k_2)(V_1 W_2^\beta - V_2 W_1^\beta), \\
QF_{12}^{mn} &= k_{12}^m(\lambda\gamma^n W_{12}) - k_{12}^n(\lambda\gamma^m W_{12}) + (k_1 \cdot k_2)(V_1 F_2^{mn} - V_2 F_1^{mn}) \\
&\quad + (k_1 \cdot k_2)(A_1^n(\lambda\gamma^m W_2) - A_2^n(\lambda\gamma^m W_1) - A_1^m(\lambda\gamma^n W_2) + A_2^m(\lambda\gamma^n W_1)),
\end{aligned} \tag{3.20}$$

where the term $(\lambda\gamma^m\lambda)A_m^{12}$ is absent by the pure-spinor constraint (2.8).

3.4. Multiparticle

Higher-point amplitudes can be elegantly described by *multiparticle superfields* that contain information on multiple particles at once. These superfields not only played a fundamental role in the derivation of the n -point disk amplitude in [9] but will also simplify the description of one-loop correlators.

3.4.1. Multiparticle vertex operators

As shown in [18], there is a multiparticle generalization of the the above superfields,

$$K_P \in \{A_\alpha^P, A_m^P, W_P^\alpha, F_P^{mn}\}, \quad (3.21)$$

that is suggested by iterated OPE calculations among vertex operators. Up to total derivatives and BRST-exact terms, the results of iterated OPEs boil down to the generalization

$$V_P \equiv \lambda^\alpha A_\alpha^P, \quad U_P \equiv \partial\theta^\alpha A_\alpha^P + \Pi^m A_m^P + d_\alpha W_P^\alpha + \frac{1}{2} N^{mn} F_{mn}^P, \quad (3.22)$$

of (3.19) and (3.16). The appearance of d_α and N_{mn} in (3.22) immediately addresses the generalization of the zero-mode integration in section 2.1.3 to higher multiplicity, where (2.16) and (2.19) become

$$\int b^{(4)} U_A U_B U_C \rightarrow (\lambda\gamma_m W_A)(\lambda\gamma_n W_B) F_C^{mn} + \text{cyc}(A, B, C), \quad (3.23)$$

$$\int b^{(2)} U_A U_B U_C U_D \rightarrow \ell_m (\lambda\gamma_a W_A)(\lambda\gamma_b W_B)(W_C \gamma^{abm} W_D) + \text{perm}(A, B, C, D), \quad (3.24)$$

see section 4 for the systematic construction of tensorial superfield building blocks.

3.4.2. Lie symmetries of multiparticle superfields

The construction of multiparticle superfields (3.21) is detailed in section 3 of [18]: Recursive equations following the structure of (3.17) are augmented by certain algorithmic redefinitions, which conspire to total derivatives or BRST-exact terms and were later identified as standard non-linear gauge transformations in [23]. More importantly, the symmetries resulting from these redefinitions are characterized by the *generalized Jacobi identities* or *Lie symmetries* (see section 8.6.7 of [22]),

$$K_{A\ell(B)C} + K_{B\ell(A)C} = 0, \quad A, B \neq \emptyset, \quad \forall C, \quad (3.25)$$

where $\ell(A)$ is the left-to-right bracketing (3.8). These are the same symmetries obtained by the following string of structure constants,

$$K_{1234\dots p} \leftrightarrow f^{12a_2} f^{a_2 3a_3} f^{a_3 4a_4} \dots f^{a_{p-1} p a_p}, \quad (3.26)$$

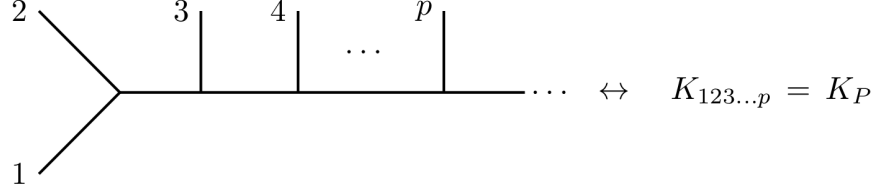


Fig. 2 The correspondence between local multiparticle superfields $K_{123\dots p} = K_P$ and tree-level subdiagrams.

and their simplest examples read

$$\begin{aligned}
K_{12C} + K_{21C} &= 0, \quad \forall C \\
K_{123C} + K_{231C} + K_{312C} &= 0, \quad \forall C \\
K_{1234C} + K_{2143C} + K_{3412C} + K_{4321C} &= 0, \quad \forall C.
\end{aligned} \tag{3.27}$$

By the correspondence (3.26) with contracted structure constants, the first two lines of (3.27) are the kinematic counterparts of the antisymmetry $f^{12a} = -f^{21a}$ and the Jacobi identities $f^{12a} f^{a3b} + \text{cyc}(1, 2, 3) = 0$. More generally, the correspondence (3.26) between the symmetries of color and kinematic factors lines up with the BCJ duality between color and kinematics [11]: Multiparticle superfields K_P implement the BCJ duality in the tree-level subdiagram of fig. 2 [18]. Accordingly, superfields that satisfy the symmetries (3.25) are said to be in the *BCJ gauge* [23].

Since the symmetries (3.25) are unchanged for any suffix word C , multiparticle superfields K_P preserve the symmetries of their lower-multiplicity counterparts. For instance, the symmetry $K_{12} + K_{21} = 0$ (when $C = \emptyset$) carries over to $K_{123} + K_{213} = 0$ (when $C = 3$) and so forth for arbitrary C .

3.4.3. Nested bracket notation for superfields in BCJ gauge

Since the superfields K_P in the BCJ gauge satisfy the same generalized Jacobi symmetries as nested brackets $\ell(P) \equiv [[\dots [p_1, p_2], p_3], \dots], p_n]$, it is convenient to use a notation where this is manifest. To this effect, a word P is understood as having a nested bracket structure $P \rightarrow \ell(P)$ and we define¹⁴

$$K_P \equiv K_{\ell(P)}, \tag{3.28}$$

¹⁴ Note, however, that in the definition $K_P \equiv K_{\ell(P)}$ one must not expand the Dynkin bracket as it would imply that $K_{\ell(P)} = |P|K_P$ since K_P satisfies the generalized Jacobi identities. So it is important to stress that (3.28) is a notational device.

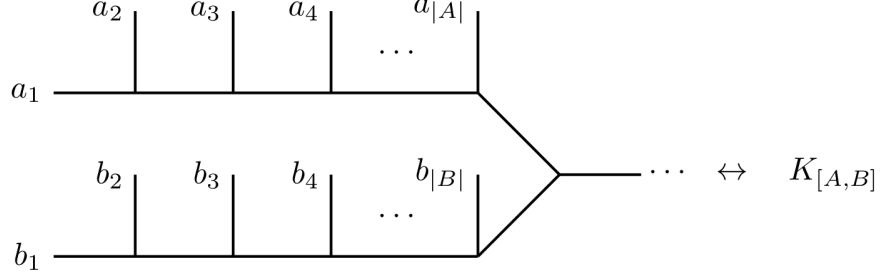


Fig. 3 The planar binary tree associated with the multiparticle superfield $K_{[A,B]}$.

for instance, $K_{12} = K_{[1,2]}$ and $K_{123} = K_{[[1,2],3]}$. The Jacobi symmetry allows the definition of local superfields with a even more general bracketing structure such as $K_{[A,[B,C]]}$. It then follows from Baker's identity [22],

$$[\ell(A), \ell(B)] = \ell(A\ell(B)), \quad (3.29)$$

that it is always possible to flatten brackets within local superfields,

$$K_{[A,B]} \equiv K_{[\ell(A), \ell(B)]} = K_{\ell(A\ell(B))} \equiv K_{A\ell(B)}. \quad (3.30)$$

For example,

$$\begin{aligned} K_{[1,2]} &= K_{1\ell(2)} = K_{12}, \\ K_{[12,3]} &= K_{12\ell(3)} = K_{123}, \\ K_{[12,34]} &= K_{12\ell(34)} = K_{1234} - K_{1243}, \\ K_{[1,[[2,3],4]]} &= K_{1\ell(234)} = K_{1234} - K_{1324} - K_{1423} + K_{1432}. \end{aligned} \quad (3.31)$$

Of course, one can check that the right-hand side of the last identity can also be written as $-K_{2341}$. The above relations are equivalent to the Jacobi identities used in the context of kinematic numerators subject to the BCJ duality. They can be visualized as flattening out of the planar binary tree associated with the two branches A and B (see fig. 3). In the context of the pure-spinor superstring, the identities (3.31) have been firstly derived in [10,9] for the special case $K_P = V_P$ as a consequence of the BRST algebra obeyed by V_P to be reviewed below.

3.4.4. BRST variation of BCJ-gauge superfields

In evaluating the BRST variations of multiparticle superfields one is faced with an interesting pattern. Explicit calculations using the equations of motion of the single-particle superfields in the generalization of the definitions (3.17) to a multiparticle setup reveals the following behavior, for example [18]

$$\begin{aligned}
QV_1 &= 0, \\
QV_{12} &= (k_1 \cdot k_2)V_1V_2, \\
QV_{123} &= (k_1 \cdot k_2)[V_1V_{23} + V_{13}V_2] + (k_{12} \cdot k_3)V_{12}V_3, \\
QV_{1234} &= (k_1 \cdot k_2)[V_1V_{234} + V_{13}V_{24} + V_{14}V_{23} + V_{134}V_2] \\
&\quad + (k_{12} \cdot k_3)[V_{12}V_{34} + V_{124}V_3] + (k_{123} \cdot k_4)V_{123}V_4, \\
QV_{12345} &= (k^1 \cdot k^2)[V_1V_{2345} + V_{13}V_{245} + V_{134}V_{25} + V_{1345}V_2 \\
&\quad + V_{135}V_{24} + V_{14}V_{235} + V_{145}V_{23} + V_{15}V_{234}] \\
&\quad + (k^{12} \cdot k^3)[V_{12}V_{345} + V_{124}V_{35} + V_{1245}V_3 + V_{125}V_{34}] \\
&\quad + (k^{123} \cdot k^4)[V_{123}V_{45} + V_{1235}V_4] \\
&\quad + (k^{1234} \cdot k^5)V_{1234}V_4.
\end{aligned} \tag{3.32}$$

It turns out that the deconcatenation and deshuffle maps defined in section 2 can be used to capture not only these identities for V_P but also for the other multiparticle superfields in a precise manner¹⁵. That is, one can show that multiparticle superfields K_P in the BCJ gauge satisfy the following BRST variations ($k_\emptyset \equiv 0$) [18]

$$\begin{aligned}
QV_P &= \sum_{\substack{P=XjY \\ Y=R\cup S}} (k_X \cdot k_j) V_{XR}V_{jS}, \\
QA_P^m &= (\lambda\gamma^m W_P) + k_P^m V_P + \sum_{\substack{P=XjY \\ Y=R\cup S}} (k_X \cdot k_j) [V_{XR}A_{jS}^m - V_{jR}A_{XS}^m], \\
QW_P^\beta &= \frac{1}{4}(\lambda\gamma^{mn})^\beta F_{mn}^P + \sum_{\substack{P=XjY \\ Y=R\cup S}} (k_X \cdot k_j) [V_{XR}W_{jS}^\beta - V_{jR}W_{XS}^\beta], \\
QF_P^{mn} &= 2k_P^{[m} (\lambda\gamma^{n]} W_P) + \sum_{\substack{P=XjY \\ Y=R\cup S}} (k_X \cdot k_j) [V_{XR}F_{jS}^{mn} + 2A_{XR}^{[n} (\lambda\gamma^{m]} W_{jS}) - (X \leftrightarrow j)],
\end{aligned} \tag{3.33}$$

¹⁵ In previous papers these BRST variations were formulated using *sets* and the powerset operation. Since sets are by definition unordered, this characterization was imprecise. This is rectified in (3.33) by using *words* together with the deshuffle map.

where $P = XjY$ denotes the deconcatenation of the word P into the word X , a single letter j , and a word Y . Moreover, $Y = R \sqcup S$ denotes the deshuffle of the word Y into the words R and S , see section 2 for more details and examples. To illustrate how these definitions are used, the first line of $QV1234$ in (3.32) is generated by one of the deconcatenation terms of $P=1234=XjY$, namely $X=1, j=2, Y=34$, and gives rise to four terms $V_{XR}V_{jS}$ according to the deshuffle $\delta(34) = \{(\emptyset, 34), (3, 4), (4, 3), (34, \emptyset)\}$.

Note that when applying the formula (3.33) to QV_1 , the deconcatenation in $1 = XjY$ implies that at least two words among X , j and Y are empty. By defining $k_\emptyset \equiv 0$ the momentum contraction $(k_X \cdot k_j)$ vanishes and we get the correct answer.

4. Pure-spinor superspace: local superfields

In this section we motivate and define a family of *local* kinematic building blocks for one-loop open-string amplitudes. On the one hand, they will be demonstrated to capture the contributions to the correlators (2.28) in pure-spinor superspace arising from the b -ghost sectors described in section 2.1.3. On the other hand, these building blocks are intertwined by their BRST variations: This defining property will be referred to as *BRST covariance* and shown to be the suitable starting point for local and BRST invariant correlators.

The non-local counterparts of the subsequent building blocks in the form of supersymmetric Berends–Giele currents have been considered in [24], and we now complete that discussion by explicitly presenting their *local* versions. In the appendix D we display the BRST variations of every local building block relevant for correlators up to multiplicity eight. A subset of these local building blocks has been used in the construction of the four, five and six-point one-loop amplitudes of ten-dimensional SYM in [19] and the six-point string amplitudes in [32].

4.1. Scalars

The zero-mode integrations in the one-loop amplitude prescription (2.4) select certain superfields from the vertex operators according to their associated worldsheet variables. For example, the b -ghost zero-mode contribution of the form $b^{(4)} = d^4 \delta'(N)$ was argued to require the zero modes $d_\alpha d_\beta N^{mn}$ from the external vertices, see section 2.1.3. From the expression (3.22) for the multiparticle integrated vertex operator, these zero modes are

accompanied by superfields $W_A^\alpha W_B^\beta F_C^{mn}$. Then, by the resulting effective rule (3.23), one is naturally led to the definition [24]

$$T_{A,B,C} \equiv \frac{1}{3}(\lambda\gamma_m W_A)(\lambda\gamma_n W_B)F_C^{mn} + \text{cyclic}(A, B, C). \quad (4.1)$$

Using the BRST variation of the multiparticle superfields (3.33), it follows that the BRST variation of (4.1) is given by ($k_\emptyset \equiv 0$)

$$QT_{A,B,C} = \sum_{\substack{A=XjY \\ Y=R\omega S}} (k_X \cdot k_j) [V_{XR}T_{jS,B,C} - V_{jR}T_{XS,B,C}] + (A \leftrightarrow B, C), \quad (4.2)$$

where the notation for the sums is explained below (3.33). For example, the BRST variations of all $T_{A,B,C}$ up to multiplicity five are given by

$$\begin{aligned} QT_{1,2,3} &= 0, \\ QT_{12,3,4} &= (k_1 \cdot k_2) [V_1 T_{2,3,4} - V_2 T_{1,3,4}], \\ QT_{123,4,5} &= (k_1 \cdot k_2) [V_1 T_{23,4,5} + V_{13} T_{2,4,5} - V_2 T_{13,4,5} - V_{23} T_{1,4,5}] \\ &\quad + (k_{12} \cdot k_3) [V_{12} T_{3,4,5} - V_3 T_{12,4,5}], \\ QT_{12,34,5} &= (k_1 \cdot k_2) [V_1 T_{2,34,5} - V_2 T_{1,34,5}] + (12 \leftrightarrow 34), \end{aligned} \quad (4.3)$$

while the multiplicity-six and -seven BRST variations will be listed in the appendix D. Since the right-hand side of the BRST variation (4.2) involves the same class of objects $T_{B,C,D}$ as seen on the left-hand side, the family of building blocks (4.1) is said to be *BRST covariant*. The appearance of V_A on the right-hand side is inherited from the multiparticle equations of motion (3.33) and an integral part of our notion of BRST covariance.

Note that $T_{A,B,C}$ is symmetric in A, B, C by its definition (4.1), in agreement with the convention (3.11) adopted throughout this work.

4.2. Vectors

Vectorial building blocks can be defined from the zero-mode integrations of correlators that contain a single loop momentum ℓ^m (the zero mode of Π^m). In this case, there are two different classes of terms in the correlator (2.4), see section 2.1.3:

- i) the b -ghost contributes $b^{(4)} = d^4 \delta'(N)$ zero modes and the external vertices $\ell^m d^2 N^{mn}$
- ii) the b -ghost contributes $\ell^m d^2 \delta(N)$ via $b^{(2)}$ and the external vertices d^4

According to the zero-mode integrations (2.16) and (2.19), the superfield expressions for the two cases above are given by

$$\text{i) } A_A^m T_{B,C,D} \quad \text{and} \quad \text{ii) } W_{A,B,C,D}^m, \quad (4.4)$$

where the effective rule (3.24) gives rise to

$$W_{A,B,C,D}^m \equiv \frac{1}{12} (\lambda \gamma_n W_A) (\lambda \gamma_p W_B) (W_C \gamma^{mnp} W_D) + (A, B | A, B, C, D). \quad (4.5)$$

The relative coefficient of these superfields is uniquely fixed as

$$T_{A,B,C,D}^m \equiv [A_A^m T_{B,C,D} + (A \leftrightarrow B, C, D)] + W_{A,B,C,D}^m \quad (4.6)$$

once we impose the covariant BRST transformation

$$QT_{A,B,C,D}^m = k_A^m V_A T_{B,C,D} + \sum_{\substack{A=XjY \\ Y \in R \sqcup S}} (k_X \cdot k_j) [V_{XR} T_{jS,B,C,D}^m - V_{jR} T_{XS,B,C,D}^m] + (A \leftrightarrow B, C, D), \quad (4.7)$$

for example, the BRST variations of all $T_{A,B,C,D}^m$ up to multiplicity six are given by

$$\begin{aligned} QT_{1,2,3,4}^m &= k_1^m V_1 T_{2,3,4} + (1 \leftrightarrow 2, 3, 4), \quad (4.8) \\ QT_{12,3,4,5}^m &= [k_{12}^m V_{12} T_{3,4,5} + (12 \leftrightarrow 3, 4, 5)] + (k_1 \cdot k_2) (V_1 T_{2,3,4,5}^m - V_2 T_{1,3,4,5}^m), \\ QT_{123,4,5,6}^m &= [k_{123}^m V_{123} T_{4,5,6} + (123 \leftrightarrow 4, 5, 6)] \\ &\quad + (k_1 \cdot k_2) [V_1 T_{23,4,5,6}^m + V_{13} T_{2,4,5,6}^m - (1 \leftrightarrow 2)] \\ &\quad + (k_{12} \cdot k_3) [V_{12} T_{3,4,5,6}^m - (12 \leftrightarrow 3)], \\ QT_{12,34,5,6}^m &= [k_{12}^m V_{12} T_{34,5,6} + (12 \leftrightarrow 34, 5, 6)] \\ &\quad + (k_1 \cdot k_2) [V_1 T_{2,34,5,6}^m - (1 \leftrightarrow 2)] + (k_3 \cdot k_4) [V_3 T_{12,4,5,6}^m - (3 \leftrightarrow 4)], \end{aligned}$$

while the examples at multiplicity seven are listed in the appendix D. In order to track the origin of BRST covariance, we first compute the BRST variations of the superfields in (4.4) to obtain

$$QA_A^m T_{B,C,D} = k_A^m V_A T_{B,C,D} + (\lambda \gamma^m W_A) T_{B,C,D} \quad (4.9)$$

$$\begin{aligned} &+ \sum_{\substack{A=XjY \\ Y \in R \sqcup S}} (k_X \cdot k_j) [V_{XR} A_{jS}^m - (X \leftrightarrow j)] T_{B,C,D} \\ &+ \sum_{\substack{B=XjY \\ Y \in R \sqcup S}} (k_X \cdot k_j) [V_{XR} A_A^m T_{jS,C,D} - (X \leftrightarrow j)] + (B \leftrightarrow C, D), \end{aligned}$$

$$QW_{A,B,C,D}^m = -(\lambda \gamma^m W_A) T_{B,C,D} + (A \leftrightarrow B, C, D) \quad (4.10)$$

$$+ \sum_{\substack{A=XjY \\ Y \in R \sqcup S}} (k_X \cdot k_j) [V_{XR} W_{jS,B,C,D}^m - (X \leftrightarrow j)] + (A \leftrightarrow B, C, D).$$

The linear combination in (4.6) is tailored to cancel the non-covariant term $(\lambda\gamma^m W_A)T_{B,C,D}$ in which the vector index is carried by a gamma matrix and one arrives at (4.7). The remaining terms in (4.7) are compatible with the notion of BRST covariance: deshuffle sums of $T_{A,B,C,D}^m$ itself or terms of the form $k_A^m V_A T_{B,C,D}$, where the vector index is carried by momenta. As firstly observed in [42,8], BRST covariant vector building blocks are crucial for BRST invariance of closed-string amplitudes that contain vector contractions between left- and right-movers.

The non-local counterparts of $T_{A,B,C}$ and $T_{A,B,C,D}^m$ can be found in section 2.4 of [24].

4.3. Tensors

Local building blocks of higher tensor ranks can be defined from the zero-mode integrations of correlators that contain higher powers of loop momenta. For instance, with two loop momenta¹⁶

$$\begin{aligned} T_{A,B,C,D,E}^{mn} &\equiv A_A^{(m} A_B^{n)} T_{C,D,E} + (A, B|A, B, C, D, E) \\ &\quad + A_A^{(m} W_{B,C,D,E}^{n)} + (A \leftrightarrow B, C, D, E) \\ &= A_A^m W_{B,C,D,E}^n + A_A^n T_{B,C,D,E}^m + (A \leftrightarrow B, C, D, E), \end{aligned} \quad (4.11)$$

or three loop momenta,

$$\begin{aligned} T_{A,B,C,D,E,F}^{mnp} &\equiv A_A^{(m} A_B^n A_C^p) T_{D,E,F} + (A, B, C|A, B, C, D, E, F) \\ &\quad + A_A^{(m} A_B^n W_{C,D,E,F}^p) + (A, B|A, B, C, D, E, F), \end{aligned} \quad (4.12)$$

and in general,

$$\begin{aligned} T_{B_1, B_2, \dots, B_{r+3}}^{m_1 \dots m_r} &\equiv T_{B_1, B_2, B_3} A_{B_4}^{(m_1} A_{B_5}^{m_2} \dots A_{B_{r+3}}^{m_r)} + (B_1, B_2, B_3|B_1, B_2, \dots, B_{r+3}) \\ &\quad + W_{B_1, B_2, B_3, B_4}^{(m_1} A_{B_5}^{m_2} \dots A_{B_{r+3}}^{m_r)} + (B_1, \dots, B_4|B_1, B_2, \dots, B_{r+3}). \end{aligned} \quad (4.13)$$

Similarly as before, the terms in the first line originate from the $\Pi^{m_1} \Pi^{m_2} \dots \Pi^{m_r} d_\alpha d_\beta N_{pq}$ zero-mode coefficient in the external vertices under the integration rules (2.20). The second line in turn originates from the b -ghost sector linear in Π^m , see (2.21).

¹⁶ Unless otherwise noted or when written inside [...], the convention for summing over the permutations $(B_1 \leftrightarrow B_2, \dots, B_p)$ applies only to the line in which the permutation appears.

Straightforward but tedious calculations using the BRST variations (3.33) of multi-particle superfields imply the rank-two variation,

$$\begin{aligned}
QT_{A,B,C,D,E}^{mn} &= \delta^{mn} Y_{A,B,C,D,E} \\
&+ k_A^{(m} V_A T_{B,C,D,E}^{n)} + (A \leftrightarrow B, C, D, E) \\
&+ \sum_{\substack{A=XjY \\ Y=R\omega S}} (k_X \cdot k_j) [V_{XR} T_{jS,B,C,D,E}^{mn} - (X \leftrightarrow j)] + (A \leftrightarrow B, C, D, E),
\end{aligned} \tag{4.14}$$

where the anomaly building block $Y_{A,B,C,D,E}$ and its generalizations will be introduced in the next subsection. Similarly, we find the following variation at rank three:

$$\begin{aligned}
QT_{B_1,B_2,\dots,B_6}^{mnp} &= \delta^{(mn} Y_{B_1,\dots,B_6}^{p)} \\
&+ k_{B_1}^{(m} V_{B_1} T_{B_2,\dots,B_6}^{np)} + (B_1 \leftrightarrow B_2, \dots, B_6) \\
&+ \sum_{\substack{B_1=XjY \\ Y=R\omega S}} (k_X \cdot k_j) [V_{XR} T_{jS,B_2,\dots,B_6}^{mnp} - (X \leftrightarrow j)] + (B_1 \leftrightarrow B_2, \dots, B_6).
\end{aligned} \tag{4.15}$$

In general, the BRST variation is given by

$$\begin{aligned}
QT_{B_1,B_2,\dots,B_{r+3}}^{m_1\dots m_r} &= \delta^{(m_1 m_2} Y_{B_1,B_2,\dots,B_{r+3}}^{m_3\dots m_r)} \\
&+ k_{B_1}^{(m_1} V_{B_1} T_{B_2,\dots,B_{r+3}}^{m_2\dots m_r)} + (B_1 \leftrightarrow B_2, \dots, B_{r+3}) \\
&+ \sum_{\substack{B_1=XjY \\ Y=R\omega S}} (k_X \cdot k_j) [V_{XR} T_{jS,B_2,\dots,B_{r+3}}^{m_1\dots m_r} - (X \leftrightarrow j)] + (B_1 \leftrightarrow B_2, \dots, B_{r+3}).
\end{aligned} \tag{4.16}$$

For example, the BRST variations of the two- and three-tensor $T_{A,B,\dots}^{m\dots}$ up to multiplicity six are given by

$$\begin{aligned}
QT_{1,2,3,4,5}^{mn} &= \delta^{mn} Y_{1,2,3,4,5} + [k_1^{(m} V_1 T_{2,3,4,5}^{n)} + (1 \leftrightarrow 2, 3, 4, 5)], \\
QT_{12,3,4,5,6}^{mn} &= \delta^{mn} Y_{12,3,4,5,6} + [k_{12}^{(m} V_{12} T_{3,4,5,6}^{n)} + (12 \leftrightarrow 3, 4, 5, 6)] \\
&+ (k_1 \cdot k_2) [V_1 T_{2,3,4,5,6}^{mn} - (1 \leftrightarrow 2)], \\
QT_{1,2,3,4,5,6}^{mnp} &= \delta^{(mn} Y_{1,2,3,4,5,6}^{p)} + [k_1^{(m} V_1 T_{2,3,4,5,6}^{np)} + (1 \leftrightarrow 2, 3, 4, 5, 6)],
\end{aligned} \tag{4.17}$$

while the multiplicity-seven examples will be listed in the appendix D. Here, our notion of BRST covariance is extended as follows: The admissible terms on the right-hand sides of the variations are either deshuffle sums of the terms $T_{B_1,B_2,\dots,B_{r+3}}^{m_1\dots m_r}$ on left-hand side along with V_A , or they comprise an anomalous superfield $Y_{B_1,\dots,B_{r+5}}^{m_1\dots m_r}$ to be introduced next.

The non-local counterparts of $T_{B_1,\dots,B_{r+3}}^{m_1\dots m_r}$ can be found in sections 3 and 4 of [24].

4.3.1. Anomalous building blocks

One-loop amplitudes of the open superstring at $n \geq 6$ points are known to exhibit a gauge anomaly before combining the worldsheet topologies [34], also see [52]. The supersymmetric kinematic factor of the six-point anomaly derived with the pure-spinor formalism in [53] was given in terms of the pure-spinor superspace expression $(\lambda\gamma^m W_2)(\lambda\gamma^n W_3)(\lambda\gamma^p W_4)(W_5\gamma_{mnp} W_6)$. By promoting the W_i to multiparticle superfields, one arrives at its higher-point extension

$$\begin{aligned} Y_{A,B,C,D,E} &\equiv \frac{1}{2}(\lambda\gamma^m W_A)(\lambda\gamma^n W_B)(\lambda\gamma^p W_C)(W_D\gamma_{mnp} W_E) \\ &= (\lambda\gamma_m W_A)W_{B,C,D,E}^m, \end{aligned} \quad (4.18)$$

as well as its tensorial generalization

$$\begin{aligned} Y_{B_1,B_2,\dots,B_{r+5}}^{m_1\dots m_r} &\equiv Y_{B_1,\dots,B_5} A_{B_6}^{(m_1} A_{B_7}^{m_2} \dots A_{B_{r+5}}^{m_r)} + (B_1, \dots, B_5 | B_1, \dots, B_{r+5}) \\ &= A_{B_1}^{m_1} Y_{B_2,B_3,\dots,B_{r+5}}^{m_2\dots m_r} + (B_1 \leftrightarrow B_2, B_3, \dots, B_{r+5}), \end{aligned} \quad (4.19)$$

and their symmetry in B_1, B_2, \dots, B_{r+5} follows from the pure-spinor constraint and group-theory arguments [24]. These definitions enter the BRST variations (4.14) to (4.16) of the higher-rank building blocks introduced above. By the arguments in appendix B.5 of [24], the bosonic components $\langle Y_{B_1,B_2,\dots,B_{r+5}}^{m_1\dots m_r} \rangle$ are parity odd, i.e. proportional to the ten-dimensional Levi–Civita symbol.

The BRST variations of the anomaly building blocks (4.19) themselves are covariant as well: They follow the structure of $QT_{A,B,C}$, $QT_{A,B,C,D}^m$ and $QT_{B_1,\dots,B_{r+3}}^{m_1\dots m_r}$ in (4.2), (4.7) and (4.16), respectively¹⁷,

$$\begin{aligned} QY_{A,B,C,D,E} &= \sum_{\substack{A=XjY \\ Y=R\omega S}} (k_X \cdot k_j) \left[V_{XR} Y_{jS,B,C,D,E} - (X \leftrightarrow j) \right] + (A \leftrightarrow B, C, D, E), \\ QY_{A,B,C,D,E,F}^m &= k_A^m V_A Y_{B,C,D,E,F} + (A \leftrightarrow B, C, D, E, F) \\ &\quad + \sum_{\substack{A=XjY \\ Y=R\omega S}} (k_X \cdot k_j) \left[V_{XR} Y_{jS,B,C,D,E,F}^m - (X \leftrightarrow j) \right] + (A \leftrightarrow B, C, D, E, F), \\ QY_{B_1,B_2,\dots,B_{r+5}}^{m_1\dots m_r} &= k_{B_1}^{(m_1} V_{B_1} Y_{B_2,\dots,B_{r+5}}^{m_2\dots m_r)} + (B_1 \leftrightarrow B_2, \dots, B_{r+5}) \\ &\quad + \sum_{\substack{B_1=XjY \\ Y=R\omega S}} (k_X \cdot k_j) \left[V_{XR} Y_{jS,B_2,\dots,B_{r+5}}^{m_1\dots m_r} - (X \leftrightarrow j) \right] + (B_1 \leftrightarrow B_2, \dots, B_{r+5}). \end{aligned} \quad (4.20)$$

¹⁷ The symmetry of $Y_{A,B,C,D,E}$ in A, B, \dots, E and the BRST variation of $Y_{A,B,C,D,E,F}^m$ rely on the group-theory fact that the tensor $t_{\alpha_1\dots\alpha_5} \equiv (\lambda\gamma_m)_{\alpha_1}(\lambda\gamma_n)_{\alpha_2}(\lambda\gamma_p)_{\alpha_3}\gamma_{\alpha_4\alpha_5}^{mnp}$ is totally antisymmetric in α_1 to α_5 and that the vector $t_{[\alpha_1\dots\alpha_5}(\lambda\gamma^m)_{\alpha_6]}$ vanishes [24].

For example, the BRST variations of the above anomaly building blocks up to multiplicity six are given by

$$\begin{aligned}
QY_{1,2,3,4,5} &= 0, \\
QY_{12,3,4,5,6} &= (k_1 \cdot k_2) [V_1 Y_{2,3,4,5,6} - V_2 Y_{1,3,4,5,6}], \\
QY_{1,2,3,4,5,6}^m &= k_1^m V_1 Y_{2,3,4,5,6} + (1 \leftrightarrow 2, 3, 4, 5, 6),
\end{aligned} \tag{4.21}$$

while the multiplicity-seven and -eight examples are listed in the appendix D. The non-local counterparts of $Y_{B_1, \dots, B_{r+5}}^{m_1 \dots m_r}$ can be found in sections 3 and 4 of [24].

4.4. Refined building blocks

In this section, we extend the system of $T_{B_1, \dots, B_{r+3}}^{m_1 \dots m_r}$ by additional building blocks that preserve the key property of BRST covariance. This extension is initiated by the observation that the five-point linear combination $k_1^m V_1 T_{2,3,4,5}^m + [V_{12} T_{3,4,5} + (2 \leftrightarrow 3, 4, 5)]$ is BRST closed [42]. Indeed, one can identify a local BRST generator,

$$\begin{aligned}
J_{1|2,3,4,5} &\equiv A_1^m T_{2,3,4,5}^m - \frac{1}{2} [(A_1 \cdot A_2) T_{3,4,5} + (2 \leftrightarrow 3, 4, 5)] \\
&= \frac{1}{2} A_1^m (T_{2,3,4,5}^m + W_{2,3,4,5}^m),
\end{aligned} \tag{4.22}$$

which reproduces the above terms along with an anomaly building block (4.18):

$$QJ_{1|2,3,4,5} = k_1^m V_1 T_{2,3,4,5}^m + [V_{12} T_{3,4,5} + (2 \leftrightarrow 3, 4, 5)] + Y_{1,2,3,4,5}. \tag{4.23}$$

Although the emergence of the expression (4.22) for $J_{1|2,3,4,5}$ from the amplitude prescription is unclear, its independent study is motivated by the connection with the earlier building blocks via BRST covariance.

We emphasize that label 1 enters (4.22) on special footing, i.e it does not participate in the symmetrization of the other labels 2, 3, 4, 5. That is why the notation for this *refined* label 1 separates it from the rest by a vertical bar¹⁸. The refined building block $J_{1|2,3,4,5}$ can

¹⁸ Note that $J_{A|B,C,D,E}$ can be interpreted as the refinement of $T_{B,C,D,E}^m$ and should be denoted by $T_{A|B,C,D,E}$ just like the other refined building blocks discussed below which share the parental notation, see e.g. (4.32). This inconsistency in the notation is a hysterical artifact.

be generalized to multiparticle labels. On top of promoting the superfields on the right-hand side of (4.22) to their multiparticle versions A_A^m and $T_{B,C,D,E}^m$, BRST covariance requires additional corrections $\sim H_{[A,B]}$ in

$$J_{A|B,C,D,E} \equiv A_A^m T_{B,C,D,E}^m - \left[(H_{[A,B]} + \frac{1}{2}(A_A \cdot A_B)) T_{C,D,E} + (B \leftrightarrow C, D, E) \right]. \quad (4.24)$$

A partial list of explicit expressions for local $H_{[A,B]}$ can be found in [23]. The appearance of the *redefining* superfields $H_{[A,B]}$ in (4.24) is needed in order to write the following BRST variation in terms of multiparticle superfields V_A in the BCJ gauge:

$$\begin{aligned} QJ_{A|B,C,D,E} &= k_A^m V_A T_{B,C,D,E}^m + [V_{[A,B]} T_{C,D,E} + (B \leftrightarrow C, D, E)] + Y_{A,B,C,D,E} \\ &+ \sum_{\substack{A=XjY \\ Y=R\cup S}} (k_X \cdot k_j) [V_{XR} J_{jS|B,C,D,E} - (X \leftrightarrow j)] \\ &+ \sum_{\substack{B=XjY \\ Y=R\cup S}} (k_X \cdot k_j) [V_{XR} J_{A|jS,C,D,E} - (X \leftrightarrow j)] + (B \leftrightarrow C, D, E). \end{aligned} \quad (4.25)$$

The brackets of the term $V_{[A,B]}$ in the first line can be flattened via (3.30). For example, the BRST variations of $J_{A|B,C,D,E}$ up to multiplicity six are given by

$$\begin{aligned} QJ_{1|2,3,4,5} &= k_1^m V_1 T_{2,3,4,5}^m + [V_{12} T_{3,4,5} + (2 \leftrightarrow 3, 4, 5)] + Y_{1,2,3,4,5}, \\ QJ_{12|3,4,5,6} &= k_{12}^m V_{12} T_{3,4,5,6}^m + [V_{123} T_{4,5,6} + (3 \leftrightarrow 4, 5, 6)] + Y_{12,3,4,5,6} \\ &+ (k_1 \cdot k_2)(V_1 J_{2|3,4,5,6} - V_2 J_{1|3,4,5,6}), \\ QJ_{1|23,4,5,6} &= k_1^m V_1 T_{23,4,5,6}^m - V_{231} T_{4,5,6} + [V_{14} T_{23,5,6} + (4 \leftrightarrow 5, 6)] \\ &+ Y_{1,23,4,5,6} + (k_2 \cdot k_3)(V_2 J_{1|3,4,5,6} - V_3 J_{1|2,4,5,6}), \end{aligned} \quad (4.26)$$

while the multiplicity-seven examples will be listed in the appendix D. In checking the BRST variations of the above refined building block it is convenient to note,

$$QD_{A,B} = \hat{V}_{(A,B)} + \left[\sum_{\substack{A=XjY \\ Y=R\cup S}} (k_X \cdot k_j) [V_{XR} D_{jS,B} - (X \leftrightarrow j)] + (A \leftrightarrow B) \right], \quad (4.27)$$

where

$$D_{A,B} \equiv \frac{1}{2}(A_A \cdot A_B), \quad \hat{V}_{(A,B)} \equiv \frac{1}{2} [V_A (k_A \cdot A_B) + A_A^m (\lambda \gamma_m W_B) + (A \leftrightarrow B)]. \quad (4.28)$$

4.4.1. Higher-rank tensors

One can also define higher-rank tensors following the same logic,

$$J_{A|B_1, \dots, B_{r+4}}^{m_1 \dots m_r} \equiv A_A^p T_{B_1, \dots, B_{r+4}}^{pm_1 \dots m_r} - \left[(H_{[A, B_1]} + \frac{1}{2} (A_A \cdot A_{B_1})) T_{B_2, \dots, B_{r+4}}^{m_1 \dots m_r} + (B_1 \leftrightarrow B_2, \dots, B_{r+4}) \right]. \quad (4.29)$$

In doing so, the word A separated by a vertical bar is said to be *refined*. Straightforward but long and tedious calculations show that

$$\begin{aligned} QJ_{A|B_1, \dots, B_{r+4}}^{m_1 \dots m_r} &= k_A^p V_A T_{B_1, \dots, B_{r+4}}^{pm_1 \dots m_r} + \delta^{(m_1 m_2} Y_{A|B_1, \dots, B_{r+4}}^{m_3 \dots m_r)} + Y_{A, B_1, \dots, B_{r+4}}^{m_1 \dots m_r} \\ &+ V_{[A, B_1]} T_{B_2, \dots, B_{r+4}}^{m_1 \dots m_r} + k_{B_1}^{(m_1} V_{B_1} J_{A|B_2, \dots, B_{r+4}}^{m_2 \dots m_r)} + (B_1 \leftrightarrow B_2, \dots, B_{r+4}) \\ &+ \sum_{\substack{A=XjY \\ Y=R\omega S}} (k_X \cdot k_j) [V_{XR} J_{jS|B_1, \dots, B_{r+4}}^{m_1 \dots m_r} - (X \leftrightarrow j)] \\ &+ \sum_{\substack{B_1=XjY \\ Y=R\omega S}} (k_X \cdot k_j) [V_{XR} J_{A|jS, B_2, \dots, B_{r+4}}^{m_1 \dots m_r} - (X \leftrightarrow j)] + (B_1 \leftrightarrow B_2, \dots, B_{r+4}), \end{aligned} \quad (4.30)$$

where the additional class of anomaly superfields $Y_{A|B_1, \dots, B_{r+4}}^{m_3 \dots m_r}$ in the first line will be defined below. For example, the BRST variations of the above superfields up to multiplicity seven are given by

$$\begin{aligned} QJ_{1|2,3,4,5,6}^m &= k_1^p T_{2,3,4,5,6}^{pm} + Y_{1,2,3,4,5,6}^m \\ &+ V_{12} T_{3,4,5,6}^m + k_2^m V_2 J_{1|3,4,5,6} + (2 \leftrightarrow 3, 4, 5, 6), \\ QJ_{12|3,4,5,6,7}^m &= k_{12}^p V_{12} T_{3,4,5,6,7}^{pm} + Y_{12,3,4,5,6,7}^m \\ &+ V_{123} T_{4,5,6,7}^m + k_3^m V_3 J_{12|4,5,6,7} + (3 \leftrightarrow 4, 5, 6, 7) \\ &+ (k_1 \cdot k_2) [V_1 J_{2|3,4,5,6,7}^m - (1 \leftrightarrow 2)], \\ QJ_{1|23,4,5,6,7}^m &= k_1^p V_1 T_{23,4,5,6,7}^{pm} + Y_{1,23,4,5,6,7}^m \\ &- V_{231} T_{4,5,6,7}^m + k_{23}^m V_{23} J_{1|4,5,6,7} \\ &+ V_{14} T_{23,5,6,7}^m + k_4^m V_4 J_{1|23,5,6,7} + (4 \leftrightarrow 5, 6, 7) \\ &+ (k_2 \cdot k_3) (V_2 J_{1|3,4,5,6,7}^m - V_3 J_{1|2,4,5,6,7}^m), \\ QJ_{1|2,3,4,5,6,7}^{mn} &= k_1^p V_1 T_{2,3,4,5,6,7}^{mnp} + \delta^{mn} Y_{1|2,3,4,5,6,7} + Y_{1,2,3,4,5,6,7}^{mn} \\ &+ V_{12} T_{3,4,5,6,7}^{mn} + k_2^{(m} V_2 J_{1|3,4,5,6,7}^{n)} + (2 \leftrightarrow 3, 4, 5, 6, 7). \end{aligned} \quad (4.31)$$

The inclusion of $J_{A|B_1, \dots, B_{r+4}}^{m_1 \dots m_r}$ and their generalizations into our system of ghost-number two building blocks is essential to rule out local cohomology objects: Up to and including

multiplicity eight, they allow to identify a BRST generator for each local BRST-invariant at ghost number three which is constructed from the alphabet of building blocks introduced in this section, see the appendix III.B for more details.

The non-local counterparts of $J_{A|B_1, \dots, B_{r+4}}^{m_1 \dots m_r}$ can be found in sections 5 and 6 of [24].

4.4.2. Refined anomaly building blocks

One can also repeat the analysis above and define the refinement of the anomaly building blocks. A covariant BRST variation written in terms of BCJ-gauge superfields fixes their general definition to be

$$Y_{A|B_1, \dots, B_{r+6}}^{m_1 \dots m_r} \equiv \frac{1}{2} A_A^p Y_{B_1, \dots, B_{r+6}}^{pm_1 \dots m_r} - [H_{[A, B_1]} Y_{B_2, \dots, B_{r+6}}^{m_1 \dots m_r} + (B_1 \leftrightarrow B_2, \dots, B_{r+6})]. \quad (4.32)$$

One can show that the BRST variation is of the same structure as $QJ_{A|B_1, \dots, B_{r+4}}^{m_1 \dots m_r}$ in (4.30),

$$\begin{aligned} QY_{A|B_1, \dots, B_{r+6}}^{m_1 \dots m_r} &= k_A^p V_A Y_{B_1, \dots, B_{r+6}}^{pm_1 \dots m_r} \\ &+ V_{[A, B_1]} Y_{B_2, \dots, B_{r+6}}^{m_1 \dots m_r} + k_{B_1}^{(m_1} V_{B_1} Y_{A|B_2, \dots, B_{r+6}}^{m_2 \dots m_r)} + (B_1 \leftrightarrow B_2, \dots, B_{r+6}) \\ &+ \sum_{\substack{B_1 = XjY \\ Y = R\omega S}} (k_X \cdot k_j) [V_{XR} Y_{A|jS, B_2, \dots, B_{r+6}}^{m_1 \dots m_r} - (X \leftrightarrow j)] + (B_1 \leftrightarrow B_2, \dots, B_{r+6}) \\ &+ \sum_{\substack{A = XjY \\ Y = R\omega S}} (k_X \cdot k_j) [V_{XR} Y_{jS|B_1, \dots, B_{r+6}}^{m_1 \dots m_r} - (X \leftrightarrow j)]. \end{aligned} \quad (4.33)$$

For example, the BRST variations of the above superfields up to multiplicity eight are

$$\begin{aligned} QY_{1|2,3,4,5,6,7} &= k_1^p V_1 Y_{2,3,4,5,6,7}^p + [V_{12} Y_{3,4,5,6,7} + (2 \leftrightarrow 3, 4, 5, 6, 7)], \\ QY_{12|3,4,5,6,7,8} &= k_{12}^p V_{12} Y_{3,4,5,6,7,8}^p + [V_{123} Y_{4,5,6,7,8} + (3 \leftrightarrow 4, 5, 6, 7, 8)] \\ &+ (k_1 \cdot k_2) (V_1 Y_{2|3,4,5,6,7,8} - (1 \leftrightarrow 2)), \\ QY_{1|23,4,5,6,7,8} &= k_1^p V_1 Y_{23,4,5,6,7,8}^p - V_{231} Y_{4,5,6,7,8} + [V_{14} Y_{23,5,6,7,8} + (4 \leftrightarrow 5, 6, 7, 8)] \\ &+ (k_2 \cdot k_3) (V_2 Y_{1|3,4,5,6,7,8} - (2 \leftrightarrow 3)), \\ QY_{1|2,3,4,5,6,7,8}^m &= k_1^p V_1 Y_{2,3,4,5,6,7,8}^{pm} + [V_{12} Y_{3,4, \dots, 8}^m + k_2^m V_2 Y_{1|3,4, \dots, 8} + (2 \leftrightarrow 3, \dots, 8)]. \end{aligned} \quad (4.34)$$

The non-local counterparts of $Y_{A|B_1, \dots, B_{r+6}}^{m_1 \dots m_r}$ can be found in section 6 of [24].

4.4.3. Higher-refinement building blocks

It is possible to generalize the degree of refinement of multiparticle building blocks in a straightforward manner by contracting refined superfields with additional instances of A_B^m .

4.4.3.1. Jacobi currents

The first non-trivial instance of an additional refined slot can be obtained by considering the term $A_m^1 J_{2|3,4,5,6,7}^m$. The BRST covariance principle suggests the definition to be

$$J_{1,2|3,4,5,6,7} \equiv A_m^1 J_{2|3,4,5,6,7}^m - \frac{1}{2} [(A_1 \cdot A_3) J_{2|4,5,6,7} + (3 \leftrightarrow 4, 5, 6, 7)], \quad (4.35)$$

since its Q -variation can be expressed in terms of simpler building blocks,

$$\begin{aligned} QJ_{1,2|3,4,5,6,7} &= k_1^m V_1 J_{2|3,4,5,6,7}^m + k_2^m V_2 J_{1|3,4,5,6,7}^m + Y_{1|2,3,4,5,6,7} + Y_{2|1,3,4,5,6,7} \\ &+ [V_{13} J_{2|4,5,6,7} + (3 \leftrightarrow 4, 5, 6, 7)] + [V_{23} J_{1|4,5,6,7} + (3 \leftrightarrow 4, 5, 6, 7)] \end{aligned} \quad (4.36)$$

and therefore generalizes (4.23). The notation $J_{1,2|3,4,5,6,7} = J_{2,1|3,4,5,6,7}$ reflects a symmetry in the refined slots 1, 2 which is not manifest from the definition (4.35).

To define a general recursion for arbitrary tensor ranks and arbitrary degree d of refinement, it will be convenient to introduce

$$\begin{aligned} W_{A_1, \dots, A_d | B_1, \dots, B_{d+r+3}}^{m_1 \dots m_{r-1} | m_r} &\equiv \frac{1}{2} A_{A_1}^p W_{A_2, \dots, A_d | B_1, \dots, B_{d+r+3}}^{p m_1 \dots m_{r-1} | m_r} \\ &- [H_{[A_1, B_1]} W_{A_2, \dots, A_d | B_2, \dots, B_{d+r+3}}^{m_1 \dots m_{r-1} | m_r} + (B_1 \leftrightarrow B_2, \dots, B_{d+r+3})]. \end{aligned} \quad (4.37)$$

Using this auxiliary superfield the recursion for refined currents $J_{A_1, \dots, A_d | B_1, \dots, B_{d+r+3}}^{m_1 \dots m_r}$ of arbitrary refinement becomes

$$\begin{aligned} J_{B_1, \dots, B_{r+3}}^{m_1 \dots m_r} &\equiv T_{B_1, \dots, B_{r+3}}^{m_1 \dots m_r}, \\ J_{A_1, \dots, A_d | B_1, \dots, B_{d+r+3}}^{m_1 \dots m_r} &\equiv \frac{1}{2} A_{A_1}^p \left[J_{A_2, \dots, A_d | B_1, \dots, B_{d+r+3}}^{p m_1 \dots m_r} + W_{A_2, \dots, A_d | B_1, \dots, B_{d+r+3}}^{m_1 \dots m_r | p} \right] \\ &- H_{[A_1, B_1]} J_{A_2, \dots, A_d | B_2, \dots, B_{d+r+3}}^{m_1 \dots m_r} + (B_1 \leftrightarrow B_2, \dots, B_{d+r+3}). \end{aligned} \quad (4.38)$$

In general, for scalars of refinement $d = 2$, one finds

$$\begin{aligned} QJ_{A,B|C,D,E,F,G} &= Y_{A|B,C,D,E,F,G} + V_A k_A^m J_{B|C,D,E,F,G}^m + (A \leftrightarrow B) \\ &+ V_{[A,C]} J_{B|D,E,F,G} + V_{[B,C]} J_{A|D,E,F,G} + (C \leftrightarrow D, E, F, G) \\ &+ \sum_{\substack{A=XjY \\ Y=R\cup S}} (k_X \cdot k_j) [V_{XR} J_{jS,B|C,D,E,F,G} - (X \leftrightarrow j)] + (A \leftrightarrow B) \\ &+ \sum_{\substack{C=XjY \\ Y=R\cup S}} (k_X \cdot k_j) [V_{XR} J_{A,B|jS,D,E,F,G} - (X \leftrightarrow j)] + (C \leftrightarrow D, \dots, G), \end{aligned} \quad (4.39)$$

and this will be the maximum degree of refinement present in the eight-point correlator. For completeness, even higher degrees of refinement and tensor ranks are possible,

$$\begin{aligned}
QJ_{A_1, \dots, A_d | B_1, \dots, B_{d+r+3}}^{m_1 \dots m_r} &= \delta^{(m_1 m_2 Y_{A_1, \dots, A_d | B_1, \dots, B_{d+r+3}}^{m_3 \dots m_r})} \\
&+ k_{B_1}^{(m_1)} V_{B_1} J_{A_1, \dots, A_d | B_2, \dots, B_{d+r+3}}^{m_2 \dots m_r} + (B_1 \leftrightarrow B_2, \dots, B_{d+r+3}) \\
&+ V_{[A_1, B_1]} J_{A_2, \dots, A_d | B_2, \dots, B_{d+r+3}}^{m_1 \dots m_r} + \left(\begin{array}{l} A_1 \leftrightarrow A_2, A_3, \dots, A_d \\ B_1 \leftrightarrow B_2, \dots, B_{d+r+3} \end{array} \right) \\
&+ Y_{A_2, \dots, A_d | A_1, B_1, \dots, B_{d+r+3}}^{m_1 \dots m_r} + k_{A_1}^p V_{A_1} J_{A_2, \dots, A_d | B_1, \dots, B_{d+r+3}}^{p m_1 \dots m_r} + (A_1 \leftrightarrow A_2, \dots, A_d) \\
&+ \sum_{\substack{A_1 = X_j Y \\ Y = R \cup S}} (k_X \cdot k_j) \left[V_{XR} J_{jS, A_2, \dots, A_d | B_1, \dots, B_{d+r+3}}^{m_1 \dots m_r} - (X \leftrightarrow j) \right] + (A_1 \leftrightarrow A_2, \dots, A_d) \\
&+ \sum_{\substack{B_1 = X_j Y \\ Y = R \cup S}} (k_X \cdot k_j) \left[V_{XR} J_{A_1, \dots, A_d | jS, B_2, \dots, B_{d+r+3}}^{m_1 \dots m_r} - (X \leftrightarrow j) \right] + (B_1 \leftrightarrow B_2, \dots, B_{d+r+3}),
\end{aligned} \tag{4.40}$$

where the objects $Y_{A_1, \dots, A_d | B_1, \dots, B_{d+r+3}}^{m_3 \dots m_r}$ in the first line will be defined next. The non-local counterparts of $W_{A_1, \dots, A_d | B_1, \dots, B_{d+r+3}}^{m_1 \dots m_{r-1} | m_r}$ and $J_{A_1, \dots, A_d | B_1, \dots, B_{d+r+3}}^{m_1 \dots m_r}$ can be found in section 6 of [24].

4.4.3.2. Anomalous building blocks

The higher-refinement generalization of the local superfields discussed above can also be applied to the anomalous building blocks. However, it turns out that already the simplest scalar building block with a double refinement can only appear starting at nine points,

$$Y_{1,2|3,4,5,6,7,8,9} \equiv \frac{1}{2} A_1^m Y_{2|3,4,5,6,7,8,9}^m. \tag{4.41}$$

More generally, defining anomaly building blocks with higher degree of refinement by

$$\begin{aligned}
Y_{A_1, \dots, A_d | B_1, \dots, B_{d+r+5}}^{m_1 \dots m_r} &\equiv \frac{1}{2} A_{A_1}^p Y_{A_2, \dots, A_d | B_1, \dots, B_{d+r+5}}^{p m_1 \dots m_r} \\
&- \left[H_{[A_1, B_1]} Y_{A_2, \dots, A_d | B_2, \dots, B_{d+r+5}}^{m_1 \dots m_r} + (B_1 \leftrightarrow B_2, \dots, B_{d+r+5}) \right],
\end{aligned} \tag{4.42}$$

their bosonic components are parity odd (cf. appendix B.5 of [24]) and their BRST variations inherit the structure of $QJ_{A_1, \dots, A_d | B_1, \dots, B_{d+r+3}}^{m_1 \dots m_r}$ in (4.40),

$$QY_{A_1, \dots, A_d | B_1, \dots, B_{d+r+5}}^{m_1 \dots m_r} = V_{[A_1, B_1]} Y_{A_2, \dots, A_d | B_2, \dots, B_{d+r+5}}^{m_1 \dots m_r} + \left(\begin{array}{l} A_1 \leftrightarrow A_2, A_3, \dots, A_d \\ B_1 \leftrightarrow B_2, \dots, B_{d+r+5} \end{array} \right)$$

$$\begin{aligned}
& + \left[\sum_{\substack{A_1=XjY \\ Y=R\omega S}} (k_X \cdot k_j) (V_{XR} Y_{jS, A_2, \dots, A_d | B_1, \dots, B_{d+r+5}}^{m_1 \dots m_r} - (X \leftrightarrow j)) \right. \\
& \quad \left. + k_{A_1}^p V_{A_1} Y_{A_2, \dots, A_d | B_1, \dots, B_{d+r+5}}^{pm_1 \dots m_r} + (A_1 \leftrightarrow A_2, \dots, A_d) \right] \quad (4.43) \\
& + \left[\sum_{\substack{B_1=XjY \\ Y=R\omega S}} (k_X \cdot k_j) (V_{XR} Y_{A_1, \dots, A_d | jS, B_2, \dots, B_{d+r+5}}^{m_1 \dots m_r} - (X \leftrightarrow j)) \right. \\
& \quad \left. + k_{B_1}^{(m_1)} V_{B_1} Y_{A_1, \dots, A_d | B_2, \dots, B_{d+r+5}}^{m_2 \dots m_r} + (B_1 \leftrightarrow B_2, \dots, B_{d+r+5}) \right].
\end{aligned}$$

For example,

$$\begin{aligned}
QY_{1,2|3,4,5,6,7,8,9} &= k_1^p V_1 Y_{2|3,4,5,6,7,8,9}^p + k_2^p V_2 Y_{1|3,4,5,6,7,8,9}^p \quad (4.44) \\
&+ V_{13} Y_{2|4,5,6,7,8,9} + V_{23} Y_{1|4,5,6,7,8,9} + (3 \leftrightarrow 4, 5, 6, 7, 8, 9).
\end{aligned}$$

The non-local counterparts of $Y_{A_1, \dots, A_d | B_1, \dots, B_{d+r+3}}^{m_1 \dots m_r}$ can be found in section 6 of [24].

4.4.4. Trace relations

As an immediate consequence of their definition (4.24), refined scalar building blocks are related to traces of unrefined tensors [24],

$$\delta_{mn} T_{A,B,C,D,E}^{mn} = 2 [J_{A|B,C,D,E} + (A \leftrightarrow B, C, D, E)]. \quad (4.45)$$

By the definitions (4.29) and (4.38), this generalizes to higher tensor rank

$$\delta_{np} T_{B_1, \dots, B_{r+5}}^{npm_1 \dots m_r} = 2 [J_{B_1 | B_2, \dots, B_{r+5}}^{m_1 \dots m_r} + (B_1 \leftrightarrow B_2, \dots, B_{r+5})], \quad (4.46)$$

and to higher degree of refinement, respectively [24]

$$\delta_{np} J_{A_1, \dots, A_d | B_1, \dots, B_{d+r+5}}^{npm_1 \dots m_r} = 2 [J_{A_1, \dots, A_d, B_1 | B_2, \dots, B_{d+r+5}}^{m_1 \dots m_r} + (B_1 \leftrightarrow B_2, \dots, B_{d+r+5})]. \quad (4.47)$$

The same structures arise for anomaly building blocks

$$\delta_{np} Y_{A_1, \dots, A_d | B_1, \dots, B_{d+r+7}}^{npm_1 \dots m_r} = 2 [Y_{A_1, \dots, A_d, B_1 | B_2, \dots, B_{d+r+7}}^{m_1 \dots m_r} + (B_1 \leftrightarrow B_2, \dots, B_{d+r+7})], \quad (4.48)$$

and both of (4.47) and (4.48) can be straightforwardly iterated to express double traces such as $\delta_{mn} \delta_{pq} J_{A_1, \dots, A_d | \dots}^{mnpq \dots}$ in terms of objects with degree of refinement $d+2$.

5. Pure-spinor superspace: non-local superfields

The goal of this work is to assemble one-loop correlators in pure-spinor superspace from kinematic building blocks and their associated worldsheet functions. When comparing these two classes of ingredients, one discovers surprising parallels in their structures and relations which will be referred to as a *duality between worldsheet functions and kinematics*, see part II. One incarnation of this duality is based on the BRST pseudo-invariants discussed in [24] and has been pioneered in [17]. The purpose of this section is to review the BRST pseudo-invariants from the perspective of the above local building blocks. They will be related to their non-local Berends–Giele representations of [24] via the so-called Berends–Giele map.

5.1. The Berends–Giele map

Every local building block discussed in section 4 can be mapped to its non-local counterpart studied in [24]. This mapping is induced by a relation among the local superfields K_P and their non-local *Berends–Giele superfield* \mathcal{K}_P given¹⁹ by the Berends–Giele (BG) map:

$$K_{iA} = \sum_B S(A|B)_i \mathcal{K}_{iB}, \quad \mathcal{K}_{iA} = \sum_B \Phi(A|B)_i K_{iB}, \quad (5.1)$$

where $S(A|B)_i$ is the KLT matrix [54] (also known as the *momentum kernel* [55]) and $\Phi(A|B)_i$ corresponds to its inverse [56],

$$\delta_{A,B} = \sum_C S(A|C)_i \Phi(C|B)_i, \quad (5.2)$$

where $\delta_{A,B}$ is equal to one if $A = B$ and zero otherwise, see (3.6). Both matrices S and Φ are symmetric and subject to the conditions $\Phi(A|B)_i = S(A|B)_i \equiv 0$ if A is not a permutation of B and they admit the following recursive forms [57,58]

$$\begin{aligned} S(P, j|Q, j, R)_i &= (k_{iQ} \cdot k_j) S(P|Q, R)_i, \quad S(\emptyset|\emptyset)_i = 1 \\ \phi(P|Q) &= \frac{1}{s_P} \sum_{\substack{XY=P \\ AB=Q}} (\phi(X|A)\phi(Y|B) - \phi(Y|A)\phi(X|B)), \quad \phi(i|j) = \delta_{ij}, \end{aligned} \quad (5.3)$$

where $\Phi(A|B)_i \equiv \phi(iA|iB)$. The first instances are given by,

$$\begin{aligned} S(2|2)_1 &= (k_1 \cdot k_2), \quad S(23|23)_1 = (k_{12} \cdot k_3)(k_1 \cdot k_2), \quad S(23|32)_1 = (k_1 \cdot k_3)(k_1 \cdot k_2), \\ \Phi(2|2)_1 &= \frac{1}{s_{12}}, \quad \Phi(23|23)_1 = \frac{1}{s_{12}s_{123}} + \frac{1}{s_{23}s_{123}}, \quad \Phi(23|32)_1 = -\frac{1}{s_{23}s_{123}}, \end{aligned}$$

¹⁹ For historic reasons the BG superfield associated with V_P is denoted M_P rather than \mathcal{V}_P . Similarly, the BG image of the local building block $T_{A,B,\dots}^{m_1,\dots}$ is denoted by $M_{A,B,\dots}^{m_1,\dots}$.

where $S(32|32)_1$ and $\Phi(32|32)_1$ follow from $S(23|23)_1$ and $\Phi(23|23)_1$, respectively, by relating $2 \leftrightarrow 3$. In (5.1) the notation \sum_B instructs to sum over all words B ; the condition that $S(A|B)_i$ and $\Phi(A|B)_i$ are zero if B is not a permutation of A leading to a finite sum.

The simplest applications of (5.1) to $K_P \rightarrow V_P$ are

$$M_{12} = \Phi(2|2)_1 V_{12} = \frac{V_{12}}{s_{12}}, \quad (5.4)$$

$$M_{123} = \Phi(23|23)_1 V_{123} + \Phi(23|32)_1 V_{132} = \frac{V_{123}}{s_{12}s_{123}} + \frac{V_{123} - V_{132}}{s_{23}s_{123}},$$

and similarly $V_{12} = S(2|2)_1 M_{12} = s_{12} M_{12}$ as well as

$$V_{123} = S(23|23)_1 M_{123} + S(23|32)_1 M_{132} = (s_{13} + s_{23})s_{12} M_{123} + s_{13}s_{12} M_{132}. \quad (5.5)$$

Consistency of the above relations can be checked by plugging the expressions (5.4) into (5.5) and in general follows from (5.2).

It is interesting to observe that the generalized Jacobi symmetries obeyed by the local superfields are translated to shuffle symmetries under the BG map. More explicitly, the BG superfields \mathcal{K}_P related to K_P by (5.1) obey [47,23]

$$\mathcal{K}_{A \sqcup B} = 0, \quad \forall A, B \neq \emptyset. \quad (5.6)$$

Note that in writing (5.1) one needs to fix the first letter of the word P in both the local K_P and non-local \mathcal{K}_P representatives to be the same. This can be done with

$$K_{BiA} = -K_{i\ell(B)A}, \quad \mathcal{K}_{BiA} = (-1)^{|B|} \mathcal{K}_{i(\tilde{B} \sqcup A)}. \quad (5.7)$$

The first relation follows from Baker's identity (3.29) while the second was proven in [59].

In general, applying the BG map to each individual slot in a local building block gives rise to its non-local Berends–Giele version. Therefore knowing one representation suffices to obtain the other, for example

$$M_{aA,bB,cC} \equiv \sum_{A',B',C'} \Phi(A|A')_a \Phi(B|B')_b \Phi(C|C')_c T_{aA',bB',cC'}, \quad (5.8)$$

$$T_{aA,bB,cC} \equiv \sum_{A',B',C'} S(A|A')_a S(B|B')_b S(C|C')_c M_{aA',bB',cC'}.$$

However, it is conceptually simpler (but equivalent to (5.8)) to define the BG counterpart of $T_{A,B,C}$ by directly using non-local multiparticle superfields in (5.1), for example [24]

$$M_{A,B,C} \equiv \frac{1}{3} (\lambda \gamma_m \mathcal{W}_A) (\lambda \gamma_n \mathcal{W}_B) \mathcal{F}_C^{mn} + \text{cyclic}(A, B, C). \quad (5.9)$$

The relation between $M_{A,B,C}$ and $T_{A,B,C}$ straightforwardly generalizes to the tensorial, refined and anomalous kinematic factors in section 4. We will use parental letters $M_{B_1, \dots, B_{r+3}}^{m_1 \dots m_r}$, $\mathcal{J}_{A_1, \dots, A_d | B_1, \dots, B_{d+r+3}}^{m_1 \dots m_r}$ and $\mathcal{Y}_{A_1, \dots, A_d | B_1, \dots, B_{d+r+5}}^{m_1 \dots m_r}$ for the Berends–Giele versions of $T_{B_1, \dots, B_{r+3}}^{m_1 \dots m_r}$, $\mathcal{J}_{A_1, \dots, A_d | B_1, \dots, B_{d+r+3}}^{m_1 \dots m_r}$ and $Y_{A_1, \dots, A_d | B_1, \dots, B_{d+r+5}}^{m_1 \dots m_r}$, respectively. The shuffle symmetry (5.6) applies to every slot of a BG current, e.g. $M_{R \sqcup S, B, C} = 0$ or $\mathcal{Y}_{A, B, R \sqcup S, D, \dots}^{mn \dots} = 0$ for $R, S \neq \emptyset$.

In addition to changing the symmetry properties within each word, the BG map also modifies the behavior under a BRST variation. The characteristic terms proportional to the momentum contraction ($k_X \cdot k_j$) in the BRST variation of the local superfields become a simpler deconcatenation sum,

$$QV_P = \sum_{\substack{P=XjY \\ Y=R \sqcup S}} (k_X \cdot k_j) V_{XR} V_{jS} \iff QM_P = \sum_{P=XY} M_X M_Y, \quad (5.10)$$

where M_P is related to V_P by (5.1). In general, one can show that the BRST variations of K_P in (3.33) are mapped to the following variations of their Berends–Giele counterparts (note $M_P \equiv \mathcal{V}_P$) [18]

$$\begin{aligned} QM_{12 \dots p} &= \sum_{j=1}^{p-1} M_{12 \dots j} M_{j+1 \dots p}, \quad (5.11) \\ Q\mathcal{A}_{12 \dots p}^m &= (\lambda \gamma^m \mathcal{W}_{12 \dots p}) + k_{12 \dots p}^m M_{12 \dots p} + \sum_{j=1}^{p-1} (M_{12 \dots j} \mathcal{A}_{j+1 \dots p}^m - M_{j+1 \dots p} \mathcal{A}_{12 \dots j}^m), \\ Q\mathcal{W}_{12 \dots p}^\alpha &= \frac{1}{4} (\lambda \gamma_{mn})^\alpha \mathcal{F}_{12 \dots p}^{mn} + \sum_{j=1}^{p-1} (M_{12 \dots j} \mathcal{W}_{j+1 \dots p}^\alpha - M_{j+1 \dots p} \mathcal{W}_{12 \dots j}^\alpha), \\ Q\mathcal{F}_{12 \dots p}^{mn} &= k_{12 \dots p}^{[m} (\lambda \gamma^{n]} \mathcal{W}_{12 \dots p}) + \sum_{j=1}^{p-1} (M_{12 \dots j} \mathcal{F}_{j+1 \dots p}^{mn} - M_{j+1 \dots p} \mathcal{F}_{12 \dots j}^{mn}) \\ &\quad + \sum_{j=1}^{p-1} [\mathcal{A}_{12 \dots j}^{[n} (\lambda \gamma^{m]} \mathcal{W}_{j+1 \dots p}) - \mathcal{A}_{j+1 \dots p}^{[n} (\lambda \gamma^{m]} \mathcal{W}_{12 \dots j})]. \end{aligned}$$

In the appendix of [18] an alternative relation between the local and the non-local superfields was given in terms of a diagrammatic map using planar binary trees. Yet one more relation between these objects will be given below in terms of the so-called *S-map*. In summary, there is a multitude of perspectives on how these superfields are defined and the relations among them.

5.1.1. The S-map between local and non-local superfields

In this subsection we will describe the so-called S-map which relates local and non-local superfield representations. This map originally appeared in the appendix of [18] as a way to encode the BCJ relations among tree amplitudes and to rewrite scalar BRST cohomology objects in terms of super Yang–Mills trees.

After defining a *weighted* concatenation product \otimes^s of Berends–Giele superfields by

$$\mathcal{K}_{Ai} \otimes^s \mathcal{K}_{jB} \equiv s_{ij} \mathcal{K}_{AijB}, \quad (5.12)$$

the definition of the S-map can be written as

$$\mathcal{K}_{S[A,B]} \equiv (-1)^{|B|+1} \mathcal{K}_{\rho(A)} \otimes^s \mathcal{K}_{\tilde{\rho}(B)}, \quad (5.13)$$

where $\rho(B)$ is defined in (3.9) and $\tilde{\rho}(B)$ denotes the reversal of $\rho(B)$. For example, given $\rho(123) = 123 - 132 - 312 + 321$ and $\tilde{\rho}(45) = 54 - 45$, the S-map $\mathcal{K}_{S[123,45]}$ yields

$$\begin{aligned} \mathcal{K}_{S[123,45]} &= (-1)^3 \mathcal{K}_{(123-132-312+321)} \otimes^s \mathcal{K}_{(54-45)} = \mathcal{K}_{(123-132-312+321)} \otimes^s \mathcal{K}_{(45-54)} \\ &= s_{34} \mathcal{K}_{12345} - s_{35} \mathcal{K}_{12354} - s_{24} \mathcal{K}_{13245} + s_{25} \mathcal{K}_{13254} - s_{24} \mathcal{K}_{31245} \\ &\quad + s_{25} \mathcal{K}_{31254} + s_{14} \mathcal{K}_{32145} - s_{15} \mathcal{K}_{32154}, \end{aligned} \quad (5.14)$$

and simpler cases include

$$\mathcal{K}_{S[1,2]} = s_{12} \mathcal{K}_{12}, \quad \mathcal{K}_{S[12,3]} = s_{23} \mathcal{K}_{123} - s_{13} \mathcal{K}_{213}. \quad (5.15)$$

A curious property of the S-map is that its iteration over all letters in a given word yields a translation between the Berends–Giele currents and its local counterparts in a way that preserves the bracketing structure. More precisely,

$$\begin{aligned} \mathcal{K}_{S[1,2]} &= K_{[1,2]}, \quad \mathcal{K}_{S[S[1,2],3]} = K_{[[1,2],3]}, \quad \mathcal{K}_{S[S[S[1,2],3],4]} = K_{[[[1,2],3],4]} \\ \mathcal{K}_{S[1,S[2,3],4]} &= K_{[1,[[2,3],4]]}, \end{aligned} \quad (5.16)$$

see (3.31) for a discussion on the bracketing notation for local superfields in the BCJ gauge. Note that the S-map plays a key role in deriving BCJ relations [11] of SYM tree amplitudes from the BRST cohomology [60], and that the definition (5.13) is equivalent²⁰ to

$$\mathcal{K}_{S[A,B]} \equiv \sum_{i=1}^{|A|} \sum_{j=1}^{|B|} (-1)^{i-j+|A|-1} s_{a_i b_j} \mathcal{K}_{(a_1 a_2 \dots a_{i-1} \sqcup a_{|A|} a_{|A|-1} \dots a_{i+1}) a_i b_j (b_{j-1} \dots b_2 b_1 \sqcup b_{j+1} \dots b_{|B|})} \quad (5.17)$$

for $A = a_1 a_2 \dots a_{|A|}$ and $B = b_1 b_2 \dots b_{|B|}$.

²⁰ Their equivalence follows from the identity $\rho(A) = \sum_{X \sqcup Y = A} (X \sqcup \tilde{Y}) j(-1)^{|Y|}$ [61].

5.2. BRST pseudo-invariants

Following [24], let us now consider the non-local versions of the local building blocks discussed above. As mentioned in the previous section, they are denoted by $M_{A,\dots}^m$ or by the calligraphic letter of its local counterpart. The BRST variations for the simplest cases can be written as

$$\begin{aligned}
QM_{A,B,C} &= \sum_{XY=A} [M_X M_{Y,B,C} - (X \leftrightarrow Y)] + (A \leftrightarrow B, C), \\
QM_{A,B,C,D}^m &= k_A^m M_A M_{B,C,D} + \sum_{XY=A} [M_X M_{Y,B,C,D}^m - (X \leftrightarrow Y)] + (A \leftrightarrow B, C, D), \\
QM_{A,B,C,D,E}^{mn} &= \delta^{mn} \mathcal{Y}_{A,B,C,D,E} \\
&\quad + k_A^{(m} M_A M_{B,C,D,E}^{n)} + (A \leftrightarrow B, C, D, E) \\
&\quad + \sum_{XY=A} [M_X M_{Y,B,C,D,E}^{mn} - (X \leftrightarrow Y)] + (A \leftrightarrow B, C, D, E),
\end{aligned} \tag{5.18}$$

and mirror the structure of their local counterparts (4.2), (4.7) and (4.14), respectively. For refined building blocks, the appearance of $V_{[A,B]}$ on the right-hand side of $QJ_{A|B,\dots}$ and its generalizations translates into $M_{S[A,B]}$ under the Berends–Giele map, for instance

$$\begin{aligned}
QJ_{A|B,C,D,E} &= \mathcal{Y}_{A,B,C,D,E} + k_A^m M_A M_{B,C,D,E}^m \\
&\quad + M_{S[A,B]} M_{C,D,E} + (B \leftrightarrow C, D, E) \\
&\quad + \sum_{XY=B} (M_X J_{A|Y,C,D,E} - (X \leftrightarrow Y)) + (B \leftrightarrow C, D, E) \\
&\quad + \sum_{XY=A} (M_X J_{Y|B,C,D,E} - (X \leftrightarrow Y)).
\end{aligned} \tag{5.19}$$

See [24] for more details and examples.

5.2.1. BRST invariants

The Berends–Giele currents were shown in [24] to be the natural building blocks in constructing recursion relations for BRST (pseudo-)invariants. For instance, it was shown using (5.18) that the following definitions are BRST invariant:

$$\begin{aligned}
C_{1|2,3,4} &\equiv M_1 M_{2,3,4}, \\
C_{1|23,4,5} &\equiv M_1 M_{23,4,5} + M_{12} M_{3,4,5} - M_{13} M_{2,4,5}, \\
C_{1|234,5,6} &\equiv M_1 M_{234,5,6} + M_{12} M_{34,5,6} + M_{123} M_{4,5,6} - M_{124} M_{3,5,6} \\
&\quad - M_{14} M_{23,5,6} - M_{142} M_{3,5,6} + M_{143} M_{2,5,6}, \\
C_{1|23,45,6} &\equiv M_1 M_{23,45,6} + M_{12} M_{45,3,6} - M_{13} M_{45,2,6} + M_{14} M_{23,5,6} - M_{15} M_{23,4,6} \\
&\quad - M_{412} M_{3,5,6} + M_{314} M_{2,5,6} + M_{215} M_{3,4,6} - M_{315} M_{2,4,6}.
\end{aligned} \tag{5.20}$$

Similarly, their vectorial upgrades are also BRST closed:

$$\begin{aligned}
C_{1|2,3,4,5}^m &\equiv M_1 M_{2,3,4,5}^m + [k_2^m M_{12} M_{3,4,5} + (2 \leftrightarrow 3, 4, 5)], \\
C_{1|23,4,5,6}^m &\equiv M_1 M_{23,4,5,6}^m + M_{12} M_{3,4,5,6}^m - M_{13} M_{2,4,5,6}^m + k_3^m M_{123} M_{4,5,6} - k_2^m M_{132} M_{4,5,6} \\
&\quad + [k_4^m M_{14} M_{23,5,6} - k_4^m M_{214} M_{3,5,6} + k_4^m M_{314} M_{2,5,6} + (4 \leftrightarrow 5, 6)].
\end{aligned} \tag{5.21}$$

These superfields are in the BRST cohomology and were dubbed *BRST invariants* in [18]. In general, a recursion relation was written down in [24] for these scalar and vector cohomology elements at arbitrary multiplicities. An alternative algorithm to generate the above combinations of Berends–Giele currents (and those of the subsequent tensorial generalizations $C_{1|A_1, \dots, A_{r+3}}^{m_1 \dots m_r}$) is described in appendix A.2.

5.2.2. BRST pseudo-invariants

The BRST variations of higher-rank tensors no longer vanish but they are proportional to superfields with an anomalous factor of $\mathcal{Y}_{A_1, \dots, A_d | B_1, \dots, B_{d+r+5}}^{m_1 \dots m_r}$ in each term. Superspace expressions with a purely anomalous Q variation are referred to as *BRST pseudo-invariant*. Similarly, a general recursion was written down in [24] and the first non-trivial instance is given by

$$\begin{aligned}
C_{1|2,3,4,5,6}^{mn} &= M_1 M_{2,3,4,5,6}^{mn} + [k_2^m M_{12} M_{3,4,5,6}^n + k_2^n M_{12} M_{3,4,5,6}^m + (2 \leftrightarrow 3, 4, 5, 6)] \\
&\quad - [(k_2^m k_3^n + k_2^n k_3^m) M_{213} M_{4,5,6} + (2, 3 | 2, 3, 4, 5, 6)]
\end{aligned} \tag{5.22}$$

and it satisfies

$$QC_{1|2,3,4,5,6}^{mn} = -\delta^{mn} M_1 \mathcal{Y}_{2,3,4,5,6}. \tag{5.23}$$

Unrefined pseudo-invariants $C_{1|A_1, \dots, A_{r+3}}^{m_1 \dots m_r}$ of arbitrary tensor ranks can be characterized by a leading term $M_1 M_{A_1, \dots, A_{r+3}}^{m_1 \dots m_r} + \dots$ as in (5.20) to (5.22). Then, the recursions of [24] adjoin a tail of completions via $\sim M_{1B}$ with $B \neq \emptyset$ such that $QC_{1|A_1, \dots, A_{r+3}}^{m_1 \dots m_r}$ is purely anomalous. Similarly, one can start from a refined leading term $M_1 \mathcal{J}_{A|B_1, \dots, B_{r+4}}^{m_1 \dots m_r}$, and the recursions of [24] generate completions such as

$$P_{1|2|3,4,5,6} = M_1 \mathcal{J}_{2|3,4,5,6} + M_{12} k_2^m M_{3,4,5,6}^m + [s_{23} M_{123} M_{4,5,6} + (3 \leftrightarrow 4, 5, 6)], \tag{5.24}$$

where refined and unrefined terms are mixed and the BRST variation becomes purely anomalous

$$QP_{1|2|3,4,5,6} = -M_1 \mathcal{Y}_{2,3,4,5,6}. \tag{5.25}$$

The same logic applies to higher degrees d of refinement where pseudo-invariants $P_{1|A_1, \dots, A_d | B_1, \dots, B_{d+r+3}}^{m_1 \dots m_r}$ are defined from recursively generated completions of the leading term $M_1 \mathcal{J}_{A_1, \dots, A_d | B_1, \dots, B_{d+r+3}}^{m_1 \dots m_r}$.

5.2.2.1. Symmetries of pseudo-invariants

Following our convention for subscripts with words A_1, A_2, \dots separated by commas, the most general pseudo-invariant $P_{1|A_1, \dots, A_d|B_1, \dots, B_{d+r+3}}^{m_1 \dots m_r}$ is separately symmetric in the refined slots A_i and the unrefined slots B_j but not under exchange of A_i with B_j . Also, the shuffle-symmetries (5.6) of their leading terms propagate to the pseudo-invariants, e.g.

$$C_{1|A_1, \dots, R \sqcup S, \dots, A_{r+3}}^{m_1 m_2 \dots m_r} = 0, \quad \forall R, S \neq \emptyset, \quad (5.26)$$

and the same is true for both types of slots A_i and B_j of $P_{1|A_1, \dots, A_d|B_1, \dots, B_{d+r+3}}^{m_1 \dots m_r}$.

5.2.3. Q variations of pseudo-BRST invariants

The Q variations (5.23) and (5.25) can be compactly generalized to higher multiplicity by means of BRST invariant combinations of anomalous superfields such as

$$\begin{aligned} \Gamma_{1|2,3,4,5,6} &\equiv M_1 \mathcal{Y}_{2,3,4,5,6}, & (5.27) \\ \Gamma_{1|23,4,5,6,7} &\equiv M_1 \mathcal{Y}_{23,4,5,6,7} + M_{12} \mathcal{Y}_{3,4,5,6,7} - M_{13} \mathcal{Y}_{2,4,5,6,7} \\ \Gamma_{1|2,3,4,5,6,7}^m &\equiv M_1 \mathcal{Y}_{2,3,4,5,6,7}^m + [k_2^m M_{12} \mathcal{Y}_{3,4,5,6,7} + (2 \leftrightarrow 3, 4, 5, 6, 7)] \\ \Gamma_{1|2|3,4,5,6,7,8} &\equiv M_1 \mathcal{Y}_{2|3,4,5,6,7,8} + M_{12} k_2^m \mathcal{Y}_{3,4,5,6,7,8}^m + [s_{23} M_{123} \mathcal{Y}_{4,5,6,7,8} + (3 \leftrightarrow 4, \dots, 8)]. \end{aligned}$$

The combinatorics of these expressions is identical to ghost-number three superfields $C_{1|2,3,4}$, $C_{1|23,4,5}$, $C_{1|2,3,4,5}^m$ and $P_{1|2|3,4,5,6}$ in (5.20), (5.21) and (5.24). Accordingly, the expansion of refined and tensorial invariants $\Gamma_{1|A_1, \dots, A_d|B_1, \dots, B_{d+r+5}}^{m_1 \dots m_r}$ in terms of $M_A \mathcal{Y}_{\dots}^{m_1 \dots}$ can be inferred from the analogous $P_{1|A_1, \dots, A_d|B_1, \dots, B_{d+r+3}}^{m_1 \dots m_r}$ after obvious adjustments in the number of unrefined slots²¹. These ghost-number four superfields capture the most general BRST variation

$$\begin{aligned} QC_{1|A_1, \dots, A_{r+3}}^{m_1 \dots m_r} &= -\delta^{(m_1 m_2} \Gamma_{1|A_1, \dots, A_{r+3}}^{m_3 \dots m_r)}, & (5.28) \\ QP_{1|A_1, \dots, A_d|B_1, \dots, B_{d+r+3}}^{m_1 \dots m_r} &= -\delta^{(m_1 m_2} \Gamma_{1|A_1, \dots, A_d|B_1, \dots, B_{d+r+3}}^{m_3 \dots m_r} \\ &\quad - \Gamma_{1|A_2, \dots, A_d|A_1, B_1, \dots, B_{d+r+3}}^{m_1 \dots m_r} + (A_1 \leftrightarrow A_2, \dots, A_d) \end{aligned}$$

of the above ghost-number three pseudo-invariants.

²¹ The mismatches in the numbers of slots, say between $C_{1|2,3,4,5}^m$ and the counterpart $\Gamma_{1|2,3,4,5,6,7}^m$, are accounted for by the more rigorous definition of $\Gamma_{1|A, \dots}$ in section 8 of [24], or in the appendix A.2.

5.2.4. Trace relations

The trace relations (4.45) to (4.47) of the local building blocks straightforwardly generalize under the Berends–Giele map. Moreover, the (pseudo-)invariants inherit the trace relations of their leading term, e.g.

$$\delta_{np} C_{1|B_1, B_2, \dots, B_{r+5}}^{n p m_1 \dots m_r} = 2 [P_{1|B_1|B_2, \dots, B_{r+5}}^{m_1 \dots m_r} + (B_1 \leftrightarrow B_2, \dots, B_{r+5})], \quad (5.29)$$

and at generic degree of refinement,

$$\delta_{np} P_{1|A_1, \dots, A_d|B_1, \dots, B_{d+r+5}}^{n p m_1 \dots m_r} = 2 [P_{1|A_1, \dots, A_d, B_1|B_2, \dots, B_{d+r+5}}^{m_1 \dots m_r} + (B_1 \leftrightarrow B_2, \dots, B_{d+r+5})]. \quad (5.30)$$

These relations will play a key role for the modular properties of the correlators after integration over ℓ .

5.3. Anomaly counterparts of BRST invariants

We shall now review an interesting class of anomaly superfields $\Delta_{1|\dots}$ that enter relations between the above (pseudo-)invariants of different tensor ranks. Similar to the formal operation $M_{B,C,\dots}^{m,\dots} \rightarrow \mathcal{Y}_{B,C,\dots}^{m,\dots}$ that translates the $C_{1|\dots}$ into the ghost-number four objects $\Gamma_{1|\dots}$ in section 5.2.3, one can generate anomaly building blocks of ghost-number three via

$$M_A M_{B,C,\dots}^{m,\dots} \rightarrow \mathcal{Y}_{A,B,C,\dots}^{m,\dots}, \quad M_A \mathcal{J}_{B|C,\dots}^{m,\dots} \rightarrow \mathcal{Y}_{B|A,C,\dots}^{m,\dots}, \quad (5.31)$$

by adjusting the number of slots in the obvious manner (also see the alternative algorithm in appendix A.2). Then, the Berends–Giele expansions of the simplest instances of the scalars (5.20) and vectors (5.21) translate into,

$$\begin{aligned} \Delta_{1|2,3,4,5} &\equiv \mathcal{Y}_{1,2,3,4,5} \\ \Delta_{1|23,4,5,6} &\equiv \mathcal{Y}_{1,23,4,5,6} + \mathcal{Y}_{12,3,4,5,6} - \mathcal{Y}_{13,2,4,5,6} \\ \Delta_{1|2,3,4,5,6}^m &\equiv \mathcal{Y}_{1,2,3,4,5,6}^m + [k_2^m \mathcal{Y}_{12,3,4,5,6} + (2 \leftrightarrow 3, 4, 5, 6)]. \end{aligned} \quad (5.32)$$

As discussed in [24], it turns out that all the unrefined superfields $\Delta_{1|A,\dots}$ are BRST exact after using momentum conservation,

$$k_1 + k_{B_1} + \dots + k_{B_{r+4}} = 0 \quad \Rightarrow \quad \langle \Delta_{1|B_1, B_2, \dots, B_{r+4}}^{m_1 m_2 \dots m_r} \rangle = 0. \quad (5.33)$$

However, their refined counterparts are non-zero in the BRST cohomology, i.e.

$$\langle \Delta_{1|A_1, \dots, A_d|B_1, B_2, \dots, B_{d+r+4}}^{m_1 m_2 \dots m_r} \rangle \neq 0, \quad d \geq 1. \quad (5.34)$$

For example, the simplest refined anomaly superfield is given by

$$\Delta_{1|2|3,4,5,6,7} = \mathcal{Y}_{2|1,3,4,5,6,7} + k_2^m \mathcal{Y}_{12,3,4,5,6,7}^m + [s_{23} \mathcal{Y}_{123,4,5,6,7} + (3 \leftrightarrow 4, 5, 6, 7)], \quad (5.35)$$

and follows the combinatorics of $P_{1|2|3,4,5,6}$ in (5.24) according to the translation (5.31). Its BRST variation is easily verified to be

$$\begin{aligned} Q\Delta_{1|2|3,4,5,6,7} &= V_1 k_2^m Y_{2,3,\dots,7}^m - V_{12} Y_{3,4,\dots,7} + [V_1 Y_{23,4,5,6,7} + (3 \leftrightarrow 4, \dots, 7)] \\ &= k_2^m \Gamma_{1|2,3,\dots,7}^m + [s_{23} \Gamma_{1|23,4,5,6,7} + (3 \leftrightarrow 4, \dots, 7)], \end{aligned} \quad (5.36)$$

see section 5.2.3 for the ghost-number four invariants $\Gamma_{1|\dots}^{\dots}$.

5.3.1. Locality of $\langle \Delta_{1|2|3,4,5,6,7} \rangle$

In contrast to the naive expectation from its slot structure, the component expansion of (5.35) is a *local* expression. The poles in s_{12} and s_{12j} with $j = 3, \dots, 7$ in the Berends–Giele currents of (5.35) do not propagate to the components for the following reasons:

- The components $\langle \mathcal{Y}_{123,4,5,6,7} \rangle$ cannot have a pole in s_{123}^{-1} since this would conflict with the vanishing of $\langle \Delta_{1|234,5,6,7} \rangle = \langle \mathcal{Y}_{123,4,5,6,7} + \frac{1}{2} \mathcal{Y}_{12,34,5,6,7} \rangle + \text{cyc}(1, 2, 3, 4)$ by (5.34): The Berends–Giele currents in the cyclic permutations do not introduce any additional pole in s_{123} to the right-hand side.
- The trace relations among tensorial $\Delta_{1|\dots}$ in section 10 of [24] include

$$\frac{1}{2} \langle \Delta_{1|2,3,4,5,6,7}^{mm} \rangle = \langle Y_{1|2,3,4,5,6,7} + [\Delta_{1|2|3,4,5,6,7} + (2 \leftrightarrow 3, 4, 5, 6, 7)] \rangle. \quad (5.37)$$

Since the left-hand side vanishes by (5.34) and each pole s_{1j}^{-1} can occur in no term other than $\langle \Delta_{1|j|\dots} \rangle$ on the right-hand side, the respective residues must be zero.

This superspace argument is confirmed by the bosonic components

$$\langle \Delta_{1|2|3,4,5,6,7} \rangle = -\frac{1}{2} (e_1 \cdot e_2) \epsilon_{10}^{k_3 k_4 k_5 k_6 k_7 e_3 e_4 e_5 e_6 e_7} + \text{fermions}, \quad (5.38)$$

that have been obtained from an automated calculation using [62]. We are using the schoonschip convention of writing the contracted vectors as $\epsilon_{10}^{\dots m \dots} k_m \equiv \epsilon_{10}^{\dots k \dots}$. The locality of $\langle \Delta_{1|2|3,4,5,6,7} \rangle$ will be important in section II.4.4.4 for asserting that the seven-point correlator comprising $\Delta_{1|2|3,4,5,6,7}$ (and permutations) is local.

5.3.2. Eight-point examples

Similarly, the eight-point topologies of $\Delta_{1|A|B,C,D,E,F}$ are obtained by applying the rules (5.31) to the expressions of the refined superfields $P_{1|A|B,C,D,E}$ at seven points, see (C.1) for the expansion of $\Delta_{1|23|4,5,6,7,8}$, $\Delta_{1|2|34,5,6,7,8}$ and $\Delta_{1|2|3,4,5,6,7,8}^m$ in terms of Berends–Giele currents. It is straightforward to show that their BRST variations are given by [24]

$$\begin{aligned}
Q\Delta_{1|23|4,5,6,7,8} &= k_{23}^m \Gamma_{1|23,4,\dots,8}^m + \Gamma_{1|3|2,4,\dots,8} - \Gamma_{1|2|3,4,\dots,8} \\
&\quad + [s_{34}\Gamma_{1|234,5,6,7,8} - s_{24}\Gamma_{1|243,5,6,7,8} + (4 \leftrightarrow 5, 6, 7, 8)] \\
Q\Delta_{1|2|34,5,6,7,8} &= k_2^m \Gamma_{1|2,34,5,\dots,8}^m + s_{23}\Gamma_{1|234,5,\dots,8} - s_{24}\Gamma_{1|243,5,\dots,8} \\
&\quad + [s_{25}\Gamma_{1|34,25,6,7,8} + (5 \leftrightarrow 6, 7, 8)] \\
Q\Delta_{1|2|3,4,5,6,7,8}^m &= k_2^p \Gamma_{1|2,3,4,\dots,8}^{pm} - k_2^m \Gamma_{1|2|3,4,\dots,8} + [s_{23}\Gamma_{1|23,4,\dots,8}^m + (3 \leftrightarrow 4, \dots, 8)]
\end{aligned} \tag{5.39}$$

in the momentum phase space with $k_{12\dots 8} = 0$. Since these superfields will play a role in the construction of the eight-point correlator later on, it is important to know about the kinematic poles in their component expansions. The first hint on non-locality comes from rewriting the BRST variations above after plugging in the definitions of the superfields Γ from [24]. Unlike the local expression (5.36) at seven points, some kinematic poles survive, e.g. s_{12}^{-1} , s_{13}^{-1} , s_{23}^{-1} in $Q\Delta_{1|23|4,5,6,7,8}$, see (C.2) for the full expressions.

Given the correspondence between BRST variations in pure-spinor superspace and gauge variations in components discussed in the appendix B of [24], the above relations indicate that the gauge variations of the eight-point anomaly superfields contain s_{ij}^{-1} poles. Therefore, the component expansions of $\langle \Delta_{1|A|B,C,D,E,F} \rangle$ for the eight-point topologies are not local, see their explicit expansions in the appendix C.

5.4. BRST cohomology identities

We have seen that the family of anomaly building blocks $\Delta_{1|A,\dots}$ is obtained by redistributing the slots of $M_A M_{B,C,\dots}$ in the Berends–Giele expansion of (pseudo-)invariants. The same procedure can be applied to derive BRST generators at ghost-number two whose Q -variation gives rise to relations between (pseudo-)invariants, see section 8 to 10 of [24]. For instance, translating the combinatorics of $\Delta_{1|2|3,4,5,6,7}$ in (5.35) into a non-anomalous context (and truncating the number of slots in the obvious manner) yields

$$D_{1|2|3,4,5} = \mathcal{J}_{2|1,3,4,5} + k_2^m M_{12,3,4,5}^m + [s_{23}M_{123,4,5} + (3 \leftrightarrow 4, 5)], \tag{5.40}$$

which generates the kinematic Jacobi identity [19]

$$QD_{1|2|3,4,5} = k_2^m C_{1|2,3,4,5}^m + [s_{23}C_{1|23,4,5} + (3 \leftrightarrow 4, 5)] + \Delta_{1|2,3,4,5}. \quad (5.41)$$

In a five-point momentum phase space, the last term $\Delta_{1|2,3,4,5}$ drops out from the cohomology, cf. (5.33). Of course, one could have obtained (5.40) directly by applying specializations like $M_A M_{B,C,\dots}^{m_1,\dots} \rightarrow M_{A,B,C,\dots}^{m_1,\dots}$ to each term in the Berends–Giele expansion (5.24) of $P_{1|2|3,4,5,6}$. In a similar manner, one can infer generalizations of $D_{1|2|3,4,5}$ by redistributing the slots of appropriate $P_{1|A_1,\dots,A_d|B_1,\dots}^{m_1,\dots}$, see section 8 of [24] for details. The resulting Jacobi identities include the six-point cases

$$QD_{1|4|23,5,6} = k_4^m C_{1|23,4,5,6}^m + s_{24}C_{1|423,5,6} - s_{34}C_{1|432,5,6} \quad (5.42)$$

$$+ s_{45}C_{1|23,45,6} + s_{46}C_{1|23,46,5} + \Delta_{1|23,4,5,6},$$

$$QD_{1|23|4,5,6} = k_{23}^m C_{1|23,4,5,6}^m + [s_{34}C_{1|234,5,6} - s_{24}C_{1|324,5,6} + (4 \leftrightarrow 5, 6)]$$

$$+ P_{1|3|2,4,5,6} - P_{1|2|3,4,5,6} + \Delta_{1|23,4,5,6},$$

$$QD_{1|2|3,4,5,6}^n = k_2^m C_{1|2,3,4,5,6}^{mn} + [s_{23}C_{1|23,4,5,6}^n + (3 \leftrightarrow 4, 5, 6)] - k_2^n P_{1|2|3,4,5,6} + \Delta_{1|2,3,4,5,6}^n,$$

and closed formulae for arbitrary multiplicity and tensor rank can be written down using the S-map (5.13), e.g.

$$QD_{1|A|B_1,\dots,B_{r+3}}^{m_1 m_2 \dots m_r} = k_A^p C_{1|A,B_1,\dots,B_{r+3}}^{p m_1 \dots m_r} + [C_{1|S[A,B_1],B_2,\dots,B_{r+3}}^{m_1 \dots m_r} + (B_1 \leftrightarrow B_2, \dots, B_{r+3})]$$

$$+ (k_{1AB_1 \dots B_{r+3}}^{(m_1)} - k_A^{(m_1)}) P_{1|A|B_1,\dots,B_{r+3}}^{m_2 \dots m_r} + \delta^{(m_1 m_2)} \Delta_{1|A|B_1,\dots,B_{r+3}}^{m_3 \dots m_r} \quad (5.43)$$

$$+ \Delta_{1|A,B_1,\dots,B_{r+3}}^{m_1 m_2 \dots m_r} + \sum_{XY=A} (P_{1|Y|X,B_1,\dots,B_{r+3}}^{m_1 \dots m_r} - P_{1|X|Y,B_1,\dots,B_{r+3}}^{m_1 \dots m_r}).$$

One can also derive kinematic Jacobi identities for refined (pseudo-)invariants

$$Q(\dots) = k_3^m P_{1|2|3,4,5,6,7}^m - s_{23}P_{1|23|4,5,6,7} + [s_{34}P_{1|2|34,5,6,7} + (4 \leftrightarrow 5, 6, 7)] + \Delta_{1|3|2,4,5,6,7}$$

$$Q(\dots) = k_4^m P_{1|23|4,\dots,8}^m + s_{24}P_{1|324|5,6,7,8} - s_{34}P_{1|234|5,6,7,8}$$

$$+ [s_{45}P_{1|23|45,6,7,8} + (5 \leftrightarrow 6, 7, 8)] + \Delta_{1|4|23,5,6,7,8}$$

$$Q(\dots) = k_{23}^m P_{1|4|23,5,6,7,8}^m + s_{34}P_{1|234|5,6,7,8} - s_{24}P_{1|324|5,6,7,8} - P_{1|2,4|3,5,6,7,8} \quad (5.44)$$

$$+ P_{1|3,4|2,5,6,7,8} + [s_{35}P_{1|4|235,6,7,8} - s_{25}P_{1|4|325,6,7,8} + (5 \leftrightarrow 6, 7, 8)] + \Delta_{1|23|4,\dots,8}$$

$$Q(\dots) = k_5^m P_{1|4|23,5,6,7,8}^m - s_{45}P_{1|45|23,6,7,8} + s_{25}P_{1|4|325,6,7,8} - s_{35}P_{1|4|235,6,7,8}$$

$$+ [s_{56}P_{1|4|23,56,7,8} + (6 \leftrightarrow 7, 8)] + \Delta_{1|5|23,4,6,7,8}$$

$$Q(\dots) = k_3^m P_{1|2|3,\dots,8}^{mn} - s_{23}P_{1|23|4,\dots,8}^n + [s_{34}P_{1|2|34,5,6,7,8}^n + (4 \leftrightarrow 5, 6, 7, 8)]$$

$$- k_3^n P_{1|2,3|4,\dots,8} + \Delta_{1|3|2,4,\dots,8}^n$$

by taking suitable combinations of $D_{1|A_1, \dots, A_d|B_1, \dots}^{m_1 \dots}$ at various degrees d of refinement as a BRST ancestor on the left-hand side.

Finally, the generalization of the Jacobi identity (5.43) to (pseudo-)invariants of arbitrary degree of refinement reads [24]

$$\begin{aligned}
QD_{1|A_1, \dots, A_d|B_1, \dots, B_{r+d+2}}^{m_1 \dots m_r} &= [\Delta_{1|A_2, \dots, A_d|A_1, B_1, \dots, B_{r+d+2}}^{m_1 \dots m_r} + (A_1 \leftrightarrow A_2, \dots, A_d)] \\
&+ \delta^{(m_1 m_2} \Delta_{1|A_1, \dots, A_d|B_1, \dots, B_{r+d+2}}^{m_3 \dots m_r)} - k_{A_1 A_2 \dots A_d}^{(m_1} P_{1|A_1, \dots, A_d|B_1, \dots, B_{r+d+2}}^{m_2 \dots m_r)} \\
&+ \left(k_{A_1}^p P_{1|A_2, \dots, A_d|A_1, B_1, \dots, B_{r+d+2}}^{m_1 \dots m_r} + [P_{1|A_2, \dots, A_d|S[A_1, B_1], B_2, \dots, B_{r+d+2}}^{m_1 \dots m_r} + (B_1 \leftrightarrow B_2, \dots, B_{r+d+2})] \right) \\
&- \sum_{XY=A_1} \left(P_{1|X, A_2, \dots, A_d|Y, B_1, \dots, B_{r+d+2}}^{m_1 \dots m_r} - (X \leftrightarrow Y) \right) + (A_1 \leftrightarrow A_2, \dots, A_d),
\end{aligned} \tag{5.45}$$

and the simplest example at degree of refinement $d = 2$ is

$$\begin{aligned}
QD_{1|2,3|4,5,6,7} &= \Delta_{1|2|3,4,5,6,7} + \Delta_{1|3|2,4,5,6,7} + k_3^m P_{1|2|3,4,5,6,7}^m + k_2^m P_{1|3|2,4,5,6,7}^m \\
&+ [s_{34} P_{1|2|34,5,6,7} + s_{24} P_{1|3|24,5,6,7} + (4 \leftrightarrow 5, 6, 7)].
\end{aligned} \tag{5.46}$$

6. Conclusions

In this first part of the series of papers [1] towards the derivation of one-loop correlators in string theory, several aspects related to the description of the massless string states via superfield kinematics have been thoroughly discussed.

The whole setup starts with the standard superfields describing super-Yang–Mills states in ten-dimensions [49] contained in the massless vertex operators of the pure-spinor superstring. We then used a combination of OPEs, zero-mode integration rules and covariance under BRST transformations to derive several compositions of superfields with the correct properties to describe higher-point amplitudes in the pure-spinor formalism, in both the anomalous and non-anomalous sectors. The comprehensive description of the local representatives is both new and relevant to the derivation of local n -point one-loop correlators in part III of this series²². The emphasis on these local superfield building blocks is warranted because the one-loop correlators are local objects prior to integration by parts, where only the OPE contractions and zero-mode integrations are performed.

²² Some of their properties were implicit in previous works (most notably in [24]) while in [19] it was realized that the one-loop amplitudes (up to six points) in field theory could be written using subsets of the definitions in this work.

We then reviewed their non-local representatives from [24] with special emphasis on the multitude of relations valid in the cohomology of the pure-spinor BRST charge. In this first part of the series, these relations represent cohomological identities among *superfields*. In the sequel part II, it will be shown that these same identities are realized by a set of objects completely different in nature: *functions* on the genus-one worldsheet! The pursue of this unexpected connection dubbed “duality between worldsheet functions and kinematics” will lead us to a detailed study of so-called “generalized elliptic integrands” (GEIs), which were briefly introduced in [17].

Finally, in part III the numerous definitions as well as surprising relationships and identities uncovered in parts I and II will pave the way to the assembly of one-loop correlators in many different representations.

Acknowledgements: We are indebted to the IAS Princeton and to Nima Arkani-Hamed for kind hospitality during an inspiring visit which initiated this project. This research was supported by the Munich Institute for Astro- and Particle Physics (MIAPP) of the DFG cluster of excellence “Origin and Structure of the Universe”, and we are grateful to the organizers for creating a stimulating atmosphere. CRM is supported by a University Research Fellowship from the Royal Society. The research of OS was supported in part by Perimeter Institute for Theoretical Physics. Research at Perimeter Institute is supported by the Government of Canada through the Department of Innovation, Science and Economic Development Canada and by the Province of Ontario through the Ministry of Research, Innovation and Science.

Appendix A. A new algorithm for the combinatorics of shuffle invariants

In this appendix, a multi-word generalization of the standard map $\rho(A)$ in (3.9) is introduced. It will lead to a systematic unification of the combinatorics of all C -like pseudo-invariants as well as their canonicalization identities. Among other things, this simplifies the discussion of the master recursion from section 8 of [24] and allows to write down a closed formula for the scalar functions $\mathcal{Z}_{A,B,C,D}$ of part II, see appendix II.A.

A.1. Multi-word generalization of the rho map

A.1.1. Scalar multi-word rhomap

Let us define a recursive two-word version of (3.9) by

$$\rho(A|iBj) \equiv (Ai|Bj) - (Aj|iB) + \rho(Ai|Bj) - \rho(Aj|iB), \quad \rho(A|i) \equiv 0. \quad (\text{A.1})$$

For example,

$$\rho(342|56) = (3425|6) - (3426|5) \quad (\text{A.2})$$

$$\begin{aligned} \rho(342|567) &= (3425|67) - (3427|56) + \rho(3425|67) - \rho(3427|56) \\ &= (3425|67) - (3427|56) + (34256|7) - (34257|6) - (34275|6) + (34276|5). \end{aligned}$$

The asymmetry $\rho(A|B) \neq \rho(B|A)$ motivates the vertical bar notation (rather than a comma) for separating the two words, in accordance with the convention for building blocks. In addition, the definition (A.1) will be generalized for an arbitrary number of words using the following recursion (with $\rho(A|B) \otimes_\rho \emptyset \equiv \rho(A|B)$)

$$\rho(A|B, C, \dots) \equiv \rho(A|B) \otimes_\rho C \otimes_\rho \dots + (B \leftrightarrow C, \dots) \quad (\text{A.3})$$

$$(A|B) \otimes_\rho C \otimes_\rho D \dots \equiv (A|B, C, D, \dots) + [\rho(A|C) \otimes_\rho B \otimes_\rho D \otimes_\rho \dots + (C \leftrightarrow D, \dots)] \quad (\text{A.4})$$

Note that the word B is excluded from the permutations in (A.4). It is also important to notice that the recursion (A.4) eventually stops due to the condition $\rho(A|i) = 0$. To illustrate this last point, consider $(234|56) \otimes_\rho 78 = (234|56, 78) + \rho(234|78) \otimes_\rho 56$ where the second term by itself requires further usage of (A.4):

$$\begin{aligned} \rho(234|78) \otimes_\rho 56 &= (2347|8) \otimes_\rho 56 - (2348|7) \otimes_\rho 56 \quad (\text{A.5}) \\ &= (2347|8, 56) + \rho(2347|56) \otimes_\rho 8 - (2348|7, 56) - \rho(2348|56) \otimes_\rho 7. \end{aligned}$$

Since the words attached to \otimes_ρ end up becoming letters, the recursion eventually stops due to $\rho(A|i) \equiv 0$.

A.1.2. Tensorial multi-word rhomap

In order to upgrade the recursions above to a tensorial setting we modify the end point of the recursive definition (A.1) to

$$\rho^m(A|i) \equiv k_i^m(Ai|\emptyset), \quad \rho^{m_1 m_2 m_3 \dots}(A|i) \equiv k_i^{(m_1}(Ai|\emptyset)^{m_2 m_3 \dots)}. \quad (\text{A.6})$$

All other definitions are kept unchanged except for having extra vector indices; e.g.,

$$\rho^m(A|iBj) \equiv (Ai|Bj)^m - (Aj|iB)^m + \rho^m(Ai|Bj) - \rho^m(Aj|iB), \quad (\text{A.7})$$

$$\rho^m(A|B, C, \dots) \equiv \rho^m(A|B) \otimes_\rho C \otimes_\rho \dots + (B \leftrightarrow C, \dots)$$

$$(A|B)^m \otimes_\rho C \otimes_\rho D \dots \equiv (A|B, C, D, \dots)^m + [\rho^m(A|C) \otimes_\rho B \otimes_\rho D \otimes_\rho \dots + (C \leftrightarrow D, \dots)],$$

and the generalization to multiple vector indices is straightforward.

A.1.3. Tensorial word-invariant maps

The multi-word generalization of the rhomap can be used to define the following *word invariants*,

$$\begin{aligned}\mathcal{I}(A|B, C, \dots) &\equiv (A|B, C, \dots) + \rho(A|B, C, \dots) \\ \mathcal{I}^{m_1 m_2 \dots}(A|B, C, \dots) &\equiv (A|B, C, \dots)^{m_1 m_2 \dots} + \rho^{m_1 m_2 \dots}(A|B, C, \dots).\end{aligned}\tag{A.8}$$

The reason for dubbing them “word invariants” will become clear shortly.

For an example application of the scalar three-word-invariant, let us consider

$$\begin{aligned}\mathcal{I}(2|34, 56) &= (2|34, 56) + \rho(2|34, 56) \\ &= (2|34, 56) + \rho(2|34) \otimes_\rho 56 + \rho(2|56) \otimes_\rho 34.\end{aligned}\tag{A.9}$$

Applying the recursion (A.4) in the right-hand side gives

$$\begin{aligned}\rho(2|34) \otimes_\rho 56 &= (23|4) \otimes_\rho 56 - (24|3) \otimes_\rho 56 \\ &= (23|56, 4) + \rho(23|56) \otimes_\rho 4 - (24|56, 3) - \rho(24|56) \otimes_\rho 3 \\ &= (23|56, 4) + (235|6, 4) - (236|5, 4) - (24|56, 3) - (245|6, 3) + (246|5, 3), \\ \rho(2|56) \otimes_\rho 34 &= (25|34, 6) + (253|4, 6) - (254|3, 6) - (26|34, 5) - (263|4, 5) + (264|3, 5).\end{aligned}\tag{A.10}$$

And finally collecting everything from (A.9) yields the final result:

$$\begin{aligned}\mathcal{I}(2|34, 56) &= (2|34, 56) + (23|56, 4) + (235|6, 4) - (236|5, 4) - (24|56, 3) \\ &\quad - (245|6, 3) + (246|5, 3) + (25|34, 6) + (253|4, 6) - (254|3, 6) \\ &\quad - (26|34, 5) - (263|4, 5) + (264|3, 5).\end{aligned}\tag{A.11}$$

A.2. Unifying all C -like building blocks

The word invariants (A.8) can be used to provide an alternative derivation of the Berends–Giele expansions of tensorial C -like BRST invariants defined in [24]. More explicitly, the observation is

$$\begin{aligned}C_{1|A_1, \dots, A_3} &= \mathcal{I}_C(1|A_1, \dots, A_3), & \Delta_{1|A_1, \dots, A_4} &= \mathcal{I}_\Delta(1|A_1, \dots, A_4), \\ D_{1|A_1, \dots, A_2} &= \mathcal{I}_D(1|A_1, \dots, A_2), & \Lambda_{1|A_1, \dots, A_6} &= \mathcal{I}_\Lambda(1|A_1, \dots, A_6), \\ L_{1|A_1, \dots, A_4} &= \mathcal{I}_L(1|A_1, \dots, A_4), & \Gamma_{1|A_1, \dots, A_5} &= \mathcal{I}_\Gamma(1|A_1, \dots, A_5),\end{aligned}\tag{A.12}$$

where

$$\begin{aligned}
\mathcal{I}_C(A_1|\dots, A_4) &\equiv \mathcal{I}(A_1|\dots, A_4) \text{ with } (P_1|\dots, P_4) \rightarrow M_{P_1}M_{P_2, P_3, P_4}, & (\text{A.13}) \\
\mathcal{I}_D(A_1|\dots, A_3) &\equiv \mathcal{I}(A_1|\dots, A_3) \text{ with } (P_1|\dots, P_3) \rightarrow M_{P_1, P_2, P_3}, \\
\mathcal{I}_L(A_1|\dots, A_5) &\equiv \mathcal{I}(A_1|\dots, A_5) \text{ with } (P_1|\dots, P_5) \rightarrow \mathcal{J}_{P_1|P_2, \dots, P_5}, \\
\mathcal{I}_\Delta(A_1|\dots, A_5) &\equiv \mathcal{I}(A_1|\dots, A_5) \text{ with } (P_1|\dots, P_5) \rightarrow \mathcal{Y}_{P_1, \dots, P_5}, \\
\mathcal{I}_\Lambda(A_1|\dots, A_7) &\equiv \mathcal{I}(A_1|\dots, A_7) \text{ with } (P_1|\dots, P_7) \rightarrow \mathcal{Y}_{P_1|P_2, \dots, P_7}, \\
\mathcal{I}_\Gamma(A_1|\dots, A_6) &\equiv \mathcal{I}(A_1|\dots, A_6) \text{ with } (P_1|\dots, P_6) \rightarrow M_{P_1}\mathcal{Y}_{P_2, \dots, P_6},
\end{aligned}$$

with obvious tensorial generalizations,

$$C_{1|A, B, C, D}^m = \mathcal{I}_C^m(1|A, B, C, D), \quad C_{1|A_1, \dots, A_{r+3}}^{m_1 \dots m_r} = \mathcal{I}_C^{m_1 \dots m_r}(1|A_1, \dots, A_{r+3}), \quad (\text{A.14})$$

where $(A|B, C, D, E)^m \rightarrow M_A M_{B, C, D, E}^m$ and $(A|B_1, \dots, B_{r+3})^{m_1 \dots m_r} \rightarrow M_A M_{B_1, \dots, B_{r+3}}^{m_1 \dots m_r}$. For the simplest example, consider:

$$\mathcal{I}(1|23, 4) = (1|23, 4) + \rho(1|23) \otimes_\rho 4 = (1|23, 4) + (12|3, 4) - (13|2, 4). \quad (\text{A.15})$$

Hence, we get $D_{1|23, 4} = \mathcal{I}_D(1|23, 4) = M_{1, 23, 4} + M_{12, 3, 4} - M_{13, 2, 4}$ from the second definition in (A.13), in accordance with equation (8.16) from [24]. Also note that the expansion of unrefined GEIs $E_{1|\dots}$ in terms of \mathcal{Z} -functions (cf. part II) can be obtained from

$$\begin{aligned}
E_{1|A_1, \dots, A_3} &\equiv \mathcal{I}(1|A_1, \dots, A_3) \text{ with } (P_1|\dots, P_4) \rightarrow \mathcal{Z}_{P_1, \dots, P_4} & (\text{A.16}) \\
E_{1|A_1, \dots, A_{r+3}}^{m_1 \dots m_r} &\equiv \mathcal{I}^{m_1 \dots m_r}(1, A_1, \dots, A_{r+3}) \text{ with } (P_1|\dots, P_{r+4})^{m_1 \dots m_r} \rightarrow \mathcal{Z}_{P_1, \dots, P_{r+4}}^{m_1 \dots m_r}.
\end{aligned}$$

A.3. Change-of-basis identities

The word invariants also give rise to a simple algorithm to obtain various identities for the change of basis in BRST invariants.

A.3.1. Scalar BRST invariants

Using the shuffle symmetries within the words A, B and C of $C_{j|A, B, C}$, one can always rewrite an arbitrary $C_{j|P_i Q, R, S}$ as a linear combination of the form $C_{j|iA, B, C}$ (with a given label i in the first position). So, without loss of generality, the change of basis of scalar

BRST invariants can be restricted to the case $C_{j|iA,B,C}$. The change of basis from j to i follows from

$$C_{j|iA,B,C} = C_{i|\mathcal{I}(Aj|B,C)}, \quad (\text{A.17})$$

which is equivalent to $\mathcal{I}_C(j|iA, B, C) = \mathcal{I}_C(i|\mathcal{I}(Aj|B, C))$. For example, from $C_{2|1,34,56} = C_{1|\mathcal{I}(2|34,56)}$ and (A.11) we get

$$\begin{aligned} C_{2|1,34,56} &= C_{1|2,34,56} + C_{1|23,56,4} + C_{1|235,6,4} - C_{1|236,5,4} - C_{1|24,56,3} - C_{1|26,34,5} - C_{1|263,4,5} \\ &\quad - C_{1|245,6,3} + C_{1|246,5,3} + C_{1|25,34,6} + C_{1|253,4,6} - C_{1|254,3,6} + C_{1|264,3,5}, \end{aligned} \quad (\text{A.18})$$

which reproduces the expression (F.1) from [24].

A.3.2. Tensor BRST invariants

A straightforward generalization of the algorithm for the change of basis of scalar BRST invariants leads to the vectorial identity

$$C_{j|iA,B,C,D}^m = C_{i|\mathcal{I}^m(Aj|B,C,D)}, \quad (\text{A.19})$$

where the vectorial word invariant map was defined in (A.8). However, the tensorial identity requires trace corrections proportional to \mathcal{I}_Δ from (A.13),

$$C_{j|iA,B,C,D,E}^{mn} = C_{i|\rho^{mn}(Aj|B,C,D,E)} + \delta^{mn}\mathcal{I}_\Delta(iAj|B, C, D, E). \quad (\text{A.20})$$

In summary, the word invariant map can be applied to any number of words (slots) and it unifies all C -like building block recursions from [24], as well as their change-of-basis identities.

As a final example of the unifying power of the word invariants, consider the monodromy derivation of $\mathcal{Z}_{12,34,56,78}$ to be defined by (II.4.7). We will need to change the GEI $E_{2|1,34,56,78}$ in (II.4.7) to a basis of $E_{1|A,B,C,D}$. Since with the exception of the ‘‘basis’’ letters 1 and 2 all words are multiparticle, there is no instance of an analogous identity for $C_{2|A,B,C}$ that can be slot extended. However, it is easy to use the word invariant map to obtain

$$\begin{aligned} E_{2|1,34,56,78} &= \mathcal{I}_E(1|\mathcal{I}(2|34, 56, 78)) \\ &= E_{1|23,56,78,4} + E_{1|235,78,4,6} + E_{1|2357,4,6,8} - E_{1|2358,4,6,7} - E_{1|236,78,4,5} \\ &\quad - E_{1|2367,4,5,8} + E_{1|2368,4,5,7} + E_{1|237,56,4,8} + E_{1|2375,4,6,8} - E_{1|2376,4,5,8} - E_{1|238,56,4,7} \end{aligned} \quad (\text{A.21})$$

$$\begin{aligned}
& - E_{1|2385,4,6,7} + E_{1|2386,4,5,7} - E_{1|24,56,78,3} - E_{1|245,78,3,6} - E_{1|2457,3,6,8} + E_{1|2458,3,6,7} \\
& + E_{1|246,78,3,5} + E_{1|2467,3,5,8} - E_{1|2468,3,5,7} - E_{1|247,56,3,8} - E_{1|2475,3,6,8} + E_{1|2476,3,5,8} \\
& + E_{1|248,56,3,7} + E_{1|2485,3,6,7} - E_{1|2486,3,5,7} + E_{1|25,34,78,6} + E_{1|253,78,4,6} + E_{1|2537,4,6,8} \\
& - E_{1|2538,4,6,7} - E_{1|254,78,3,6} - E_{1|2547,3,6,8} + E_{1|2548,3,6,7} + E_{1|257,34,6,8} + E_{1|2573,4,6,8} \\
& - E_{1|2574,3,6,8} - E_{1|258,34,6,7} - E_{1|2583,4,6,7} + E_{1|2584,3,6,7} - E_{1|26,34,78,5} - E_{1|263,78,4,5} \\
& - E_{1|2637,4,5,8} + E_{1|2638,4,5,7} + E_{1|264,78,3,5} + E_{1|2647,3,5,8} - E_{1|2648,3,5,7} - E_{1|267,34,5,8} \\
& - E_{1|2673,4,5,8} + E_{1|2674,3,5,8} + E_{1|268,34,5,7} + E_{1|2683,4,5,7} - E_{1|2684,3,5,7} + E_{1|27,34,56,8} \\
& + E_{1|273,56,4,8} + E_{1|2735,4,6,8} - E_{1|2736,4,5,8} - E_{1|274,56,3,8} - E_{1|2745,3,6,8} + E_{1|2746,3,5,8} \\
& + E_{1|275,34,6,8} + E_{1|2753,4,6,8} - E_{1|2754,3,6,8} - E_{1|276,34,5,8} - E_{1|2763,4,5,8} + E_{1|2764,3,5,8} \\
& - E_{1|28,34,56,7} - E_{1|283,56,4,7} - E_{1|2835,4,6,7} + E_{1|2836,4,5,7} + E_{1|284,56,3,7} + E_{1|2845,3,6,7} \\
& - E_{1|2846,3,5,7} - E_{1|285,34,6,7} - E_{1|2853,4,6,7} + E_{1|2854,3,6,7} + E_{1|286,34,5,7} + E_{1|2863,4,5,7} \\
& - E_{1|2864,3,5,7} + E_{1|34,56,78,2},
\end{aligned}$$

where \mathcal{I}_E is defined in analogy with (A.13). This example demonstrates that the large number of terms in such identities can be generated by a simple set of combinatorial rules.

Appendix B. Empty BRST cohomology for manifestly local expressions

Using the alphabet of local building blocks presented in section 4, in this appendix we will demonstrate that their most general BRST-closed linear combinations are also BRST-exact when more than four and up to eight particles are involved; i.e., the BRST cohomology is empty for manifestly local expressions of multiplicities five to eight.

At four points, the expression $V_1 T_{2,3,4}$ is clearly local, and it is not BRST-exact in the four-point phase where $s_{ijk} = 0$ (its components reproduce the one-loop four-point amplitude [63]). Consequently, there is a local cohomology at four points. In the following we study the higher-point generalizations and find that the manifestly local BRST cohomology is empty²³ for up to eight points among the combinations of building blocks in section 4.

²³ In absence of five-point momentum conservation $\langle Y_{1,2,3,4,5} \rangle \sim \epsilon_{10}(f_1, f_2, \dots, f_5)$ [53], and one could argue that it constitutes a local element of the BRST cohomology. But this is no longer true once we invoke $k_{12345}=0$ since $Y_{1,2,3,4,5}$ becomes BRST exact in this case (see section 9.1 of [24]).

B.1. BRST-closed expressions without momentum conservation

In the first series of checks, we treat all the momenta k_1, k_2, \dots, k_n in an n -point superspace expression as independent, i.e. temporarily relax momentum conservation. An automated brute-force scan of all possible linear combinations of building blocks from section 4 using FORM [64] led to the following unique BRST-closed expressions²⁴

$$\text{loc}_1^{5\text{pt}} = k_1^m V_1 T_{2,3,4,5}^m + [V_{12} T_{3,4,5} + (2 \leftrightarrow 3, 4, 5)] + Y_{1,2,3,4,5}, \quad (\text{B.1})$$

$$\begin{aligned} \text{loc}_1^{6\text{pt}} &= k_1^m Y_{1,2,3,4,5,6}^m + [Y_{12,3,4,5,6} + (2 \leftrightarrow 3, 4, 5, 6)] + k_1^m k_1^n V_1 T_{2,3,4,5,6}^{mn} \\ &+ [(2k_1^m + k_2^m) V_{12} T_{3,4,5,6}^m + s_{12} V_1 J_{2|3,4,5,6} + (2 \leftrightarrow 3, 4, 5, 6)] \\ &+ [(V_{123} + V_{132}) T_{4,5,6} + (2, 3|2, 3, 4, 5, 6)], \end{aligned} \quad (\text{B.2})$$

$$\begin{aligned} \text{loc}_1^{7\text{pt}} &= k_1 \cdot k_{234567} Y_{1|2,3,4,5,6,7} + k_1^m k_1^n Y_{1,2,3,4,5,6,7}^{mn} \\ &+ [(2k_1^m + k_2^m) Y_{12,3,4,5,6,7}^m + s_{12} Y_{2|1,3,4,5,6,7} + (2 \leftrightarrow 3, 4, 5, 6, 7)] \\ &+ [Y_{123,4,5,6,7} + Y_{132,4,5,6,7} + (2, 3|2, 3, 4, 5, 6, 7)] \\ &+ k_1^m k_1^n k_1^p V_1 T_{2,\dots,7}^{mnp} + [(3k_1^m k_1^n + 3k_1^m k_2^n + k_2^m k_2^n) V_{12} T_{3,\dots,7}^{mn} + (2 \leftrightarrow 3, \dots, 7)] \\ &+ [((3k_1^m + 2k_2^m + k_3^m) V_{123} + (3k_1^m + 2k_3^m + k_2^m) V_{132}) T_{4,\dots,7}^m + (2, 3|2, \dots, 7)] \\ &+ [(V_{1234} + \text{perm}(2, 3, 4)) T_{5,6,7} + (2, 3, 4|2, \dots, 7)] \\ &+ [s_{12} V_1 (3k_1^m + k_2^m) J_{2|3,4,5,6,7}^m + (2 \leftrightarrow 3, 4, 5, 6, 7)] \\ &+ [(3s_{13} + s_{23}) V_{12} J_{3|4,5,6,7} + (3s_{12} + s_{23}) V_{13} J_{2|4,5,6,7} \\ &\quad + (s_{12} - s_{13}) V_1 J_{23|4,5,6,7} + (2, 3|2, \dots, 7)], \end{aligned} \quad (\text{B.3})$$

as well as a much bigger expression at eight points,

$$\begin{aligned} \text{loc}_1^{8\text{pt}} &= Y_{1|2,3,\dots,8}^m \left(3k_1^m (k_1 \cdot k_{23\dots 8}) + [s_{12} k_2^m + (2 \leftrightarrow 3, 4, \dots, 8)] \right) \\ &+ [(s_{12} - s_{13}) Y_{1|23,4,\dots,8} + (2, 3|2, 3, \dots, 8)] \\ &+ [Y_{12|3,4,\dots,8} (3(k_1 \cdot k_{23\dots 8}) + (k_2 \cdot k_{34\dots 8})) + (2 \leftrightarrow 3, 4, \dots, 8)] \\ &+ [s_{12} (3k_1^m + k_2^m) Y_{2|1,3,4,\dots,8}^m + (2 \leftrightarrow 3, 4, \dots, 8)] \\ &+ [(3s_{12} + s_{23}) Y_{2|13,4,\dots,8} + (3s_{13} + s_{23}) Y_{3|12,4,\dots,8} \\ &\quad + (s_{12} - s_{13}) Y_{23|1,4,5,\dots,8} + (2, 3|2, 3, 4, \dots, 8)] \\ &+ k_1^m k_1^n k_1^p Y_{1,2,\dots,8}^{mnp} + [Y_{12,3,\dots,8}^{mn} (3k_1^m k_1^n + 3k_1^m k_2^n + k_2^m k_2^n) + (2 \leftrightarrow 3, 4, \dots, 8)] \end{aligned} \quad (\text{B.4})$$

²⁴ Note that $QY_{1,2,3,4,5} = 0$ is BRST closed by itself but we add it to $\text{loc}_1^{5\text{pt}}$ for convenience.

$$\begin{aligned}
& + [Y_{123,4,\dots,8}^m(3k_1^m + 2k_2^m + k_3^m) + Y_{132,4,\dots,8}^m(3k_1^m + 2k_3^m + k_2^m) + (2, 3|2, 3, \dots, 8)] \\
& + [(Y_{1234,5,6,7,8} + \text{perm}(2, 3, 4)) + (2, 3, 4|2, 3, 4, \dots, 8)] \\
& + k_1^m k_1^n k_1^p k_1^q V_1 T_{2,3,\dots,8}^{mnpq} + [s_{12} V_1 J_{2|3,4,\dots,8}^{mn}(6k_1^m k_1^n + 4k_1^m k_2^n + k_2^m k_2^n) + (2 \leftrightarrow 3, 4, \dots, 8)] \\
& + [(6s_{12}s_{13} + s_{12}s_{23} + s_{13}s_{23})V_1 J_{2,3|4,5,\dots,8} + (2, 3|2, 3, 4, \dots, 8)] \\
& + [V_1 J_{23|4,5,\dots,8}^m(4k_1^m(s_{12}-s_{13}) + k_2^m(2s_{12}-s_{13}) + k_3^m(s_{12}-2s_{13})) + (2, 3|2, 3, 4, \dots, 8)] \\
& + [(s_{12}-2s_{13}+s_{14})V_1 J_{234|5,6,7,8} + (s_{12}+s_{13}-2s_{14})V_1 J_{243|5,6,7,8} + (2, 3, 4|2, 3, 4, \dots, 8)] \\
& + [V_{12} T_{3,4,\dots,8}^{mnp}(4k_1^m k_1^n k_1^p + 6k_1^m k_1^n k_2^p + 4k_1^m k_2^n k_2^p + k_2^m k_2^n k_2^p) + (2 \leftrightarrow 3, 4, \dots, 8)] \\
& + [V_{12} J_{3|4,5,6,7,8}^m(k_3^m s_{23} + 4k_3^m s_{13} + 3k_2^m s_{23} + 6k_2^m s_{13} + 4k_1^m s_{23} + 12k_1^m s_{13}) \\
& \quad + V_{13} J_{2|4,5,6,7,8}^m(k_2^m s_{23} + 4k_2^m s_{12} + 3k_3^m s_{23} + 6k_3^m s_{12} + 4k_1^m s_{23} + 12k_1^m s_{12}) + (2, 3|2, 3, 4, \dots, 8)] \\
& + [V_{12} J_{34|5,6,7,8}(4s_{13} - 4s_{14} + s_{23} - s_{24}) + V_{13} J_{24|5,6,7,8}(4s_{12} - 4s_{14} + s_{23} - s_{34}) \\
& \quad + V_{14} J_{23|5,6,7,8}(4s_{12} - 4s_{13} + s_{24} - s_{34}) + (2, 3, 4|2, 3, 4, 5, 6, 7, 8)] \\
& + [V_{123}(6k_1^m k_1^n + 8k_1^m k_2^n + 4k_1^m k_3^n + 3k_2^m k_2^n + 3k_2^m k_3^n + k_3^m k_3^n)T_{4,5,6,7,8}^{mn} \\
& \quad + V_{132}(6k_1^m k_1^n + 8k_1^m k_3^n + 4k_1^m k_2^n + 3k_3^m k_3^n + 3k_2^m k_3^n + k_2^m k_2^n)T_{4,5,6,7,8}^{mn} + (2, 3|2, 3, \dots, 8)] \\
& + [(V_{123}(s_{34} + 3s_{24} + 6s_{14})J_{4|5,6,7,8} + \text{perm}(2, 3, 4)) + (2, 3, 4|2, 3, 4, 5, 6, 7, 8)] \\
& + [(V_{1234}(k_4^m + 2k_3^m + 3k_2^m + 4k_1^m)T_{5,6,7,8}^m + \text{perm}(2, 3, 4)) + (2, 3, 4|2, 3, 4, 5, 6, 7, 8)] \\
& + [(V_{12345}T_{6,7,8} + \text{perm}(2, 3, 4, 5)) + (2, 3, 4, 5|2, 3, 4, 5, 6, 7, 8)] ,
\end{aligned}$$

and these are the only BRST-closed expressions with manifestly local building blocks up to eight points. But, as alluded to in the beginning of this appendix, the above BRST-closed expressions are also BRST-exact, and the local BRST cohomology is trivial in absence of momentum conservation. In fact,

$$\text{loc}_1^{5\text{pt}} = QL_{1|2,3,4,5}, \quad (\text{B.5})$$

$$\text{loc}_1^{6\text{pt}} = Q(k_1^m L_{1|2,3,4,5,6}^m), \quad (\text{B.6})$$

$$\text{loc}_1^{7\text{pt}} = Q(k_1^m k_1^n L_{1|2,3,4,5,6,7}^{mn} + [s_{12} L_{1|2|3,4,5,6,7} + (2 \leftrightarrow 3, 4, 5, 6, 7)]), \quad (\text{B.7})$$

$$\begin{aligned}
\text{loc}_1^{8\text{pt}} \equiv & Q(k_1^m k_1^n k_1^p L_{1|2,3,\dots,8}^{mnp} + [s_{12}(3k_1^m + k_2^m)L_{1|2|3,4,\dots,8}^m + (2 \leftrightarrow 3, \dots, 8)] \\
& + [(s_{12} - s_{13})s_{23}L_{1|23|4,\dots,8} + (2, 3|2, 3, 4, \dots, 8)]), \quad (\text{B.8})
\end{aligned}$$

where the building blocks L were defined in [24], e.g. $L_{1|2,3,4,5} = \mathcal{J}_{1|2,3,4,5}$ as well as $L_{1|2,\dots,6}^m = \mathcal{J}_{1|2,\dots,6}^m + [k_2^m \mathcal{J}_{12|3,\dots,6} + (2 \leftrightarrow 3, \dots, 6)]$. Their unrefined instances $L_{1|A_1,\dots,A_{r+4}}^{m_1 \dots m_r}$ can also be explicitly obtained using the alternative algorithm in (A.12).

Although straightforward to check, the relations (B.5) to (B.8) contain a hidden systematics noteworthy of uncovering. To this effect, we rewrite the BRST-exactness solution of $\text{loc}_1^{5\text{pt}}$ in terms of BRST invariants (see equation (10.1) of [24]):

$$\text{loc}_1^{5\text{pt}} = QL_{1|2,3,4,5} = \Delta_{1|2,3,4,5} + k_1^m C_{1|2,3,4,5}^m, \quad (\text{B.9})$$

which manifests BRST invariance rather than locality. We can interpret (B.9) as the statement that the non-localities in the Berends–Giele expansion of $k_1^m C_{1|2,3,4,5}^m$ in (5.21) are in fact spurious (which is easy to verify). Since the combinatorics of the Berends–Giele expansion of the building blocks L and C is the same (see (A.12)), the conclusion is that the ghost-number-two expression $k_m^1 L_{1|2,3,4,5,6}^m$ is also a *local* expression. Even though it was not guaranteed to be the case, computing its BRST variation leads to the manifestly local expression $\text{loc}_1^{6\text{pt}}$.

Similarly, the BRST-closed expression $\text{loc}_1^{6\text{pt}}$ can be rewritten in terms of non-local BRST pseudo-invariants

$$Q(k_1^m L_{1|2,3,4,5,6}^m) = k_1^m \Delta_{1|2,3,4,5,6}^m + k_1^m k_1^n C_{1|2,3,4,5,6}^{mn} + [s_{12} P_{1|2|3,4,5,6} + (2 \leftrightarrow 3, 4, 5, 6)], \quad (\text{B.10})$$

where again the locality of the left-hand side (i.e., $\text{loc}_1^{6\text{pt}}$) is obscured by the representation with BRST (pseudo-)invariants on the right-hand side.

However, the same logic can be applied again: when promoting the BRST descendant on the right-hand side of (B.10) to a BRST generator via $C_{1|A,B,C,D,E}^{mn\dots} \rightarrow L_{1|A,B,C,D,E}^{mn\dots}$ and $\Delta_{1|A,B,C,D,E}^{mn\dots} \rightarrow 0$, locality follows from the equivalence of the respective expansions in terms of Berends–Giele currents. We therefore obtain the ghost-number-two terms $k_1^m k_1^n L_{1|2,3,4,5,6,7}^{mn} + [s_{12} L_{1|2|3,4,5,6,7} + (2 \leftrightarrow 3, 4, 5, 6, 7)]$ that are guaranteed to generate a local ghost-number-three expression upon BRST variation. The fact that it exactly reproduces the unique expression for $\text{loc}_1^{7\text{pt}}$ obtained by a brute-force search demonstrates that the manifestly local BRST cohomology is empty at seven points (using the set of building blocks from section 4).

Similarly, we use the promotion $C \rightarrow L$ and $\Delta \rightarrow 0$ in the BRST variation of the seven-point identity (see [24] or (A.12) for the anomaly superfields $\Lambda_{1|\dots}$)

$$\begin{aligned} Q(k_1^m k_1^n L_{1|2,3,4,5,6,7}^{mn} + [s_{12} L_{1|2|3,4,5,6,7} + (2 \leftrightarrow 3, 4, 5, 6, 7)]) = & \quad (\text{B.11}) \\ & + k_1 \cdot k_{1234567} \Lambda_{1|2,3,4,5,6,7} + k_1^m k_1^n \Delta_{1|2,3,4,5,6,7}^{mn} \\ & + [s_{12} \Delta_{1|2|3,4,5,6,7} + (2 \leftrightarrow 3, 4, 5, 6, 7)] \\ & + k_1^m k_1^n k_1^p C_{1|2,\dots,7}^{mnp} + [s_{12} (3k_1^m + k_2^m) P_{1|2|3,\dots,7}^m + (2 \leftrightarrow 3, \dots, 7)] \\ & + [(s_{12} - s_{13}) s_{23} P_{1|23|4,5,6,7} + (2, 3|2, 3, 4, 5, 6, 7)] \end{aligned}$$

to obtain a local eight-point expression. Surprisingly, as already stated in (B.8), the outcome of its BRST variation exactly matches the expression (B.4) and demonstrates that the manifestly local BRST cohomology is empty at eight points.

For completeness, the local expression (B.4) for $\text{loc}_1^{\text{8pt}}$ takes the following form in terms of (pseudo-)invariants

$$\begin{aligned}
\text{loc}_1^{\text{8pt}} = & \Lambda_{1|2,3,\dots,8}^m (3k_1^m (k_1 \cdot k_{12\dots 8}) + [k_2^m s_{12} + (2 \leftrightarrow 3, 4, \dots, 8)]) \tag{B.12} \\
& + [(s_{12} - s_{13})s_{23}\Lambda_{1|23,4,\dots,8} + (2, 3|2, 3, 4, \dots, 8)] \\
& + k_1^m k_1^n k_1^p \Delta_{1|2,3,\dots,8}^{mnp} + [s_{12}(3k_1^m + k_2^m)\Delta_{1|2|3,4,\dots,8}^m + (2 \leftrightarrow 3, 4, \dots, 8)] \\
& + [(s_{12} - s_{13})s_{23}\Delta_{1|23|4,\dots,8} + (2, 3|2, 3, 4, \dots, 8)] \\
& + k_1^m k_1^n k_1^p k_1^q C_{1|2,3,\dots,8}^{mnpq} + [s_{12}(6k_1^m k_1^n + 4k_1^m k_2^n + k_2^m k_2^n)P_{1|2|3,4,\dots,8}^{mn} + (2 \leftrightarrow 3, 4, \dots, 8)] \\
& + [(6s_{12}s_{13} + s_{12}s_{23} + s_{13}s_{23})P_{1|2,3|4,\dots,8} + (2, 3|2, 3, 4, \dots, 8)] \\
& + [s_{23}(4(s_{12}-s_{13})k_1^m + (2s_{12}-s_{13})k_2^m + (s_{12}-2s_{13})k_3^m)P_{1|23|4,5,6,7,8}^m + (2, 3|2, 3, \dots, 8)] \\
& + [(s_{12} - 2s_{13} + s_{14})s_{23}s_{34}P_{1|234|5,6,7,8} + (s_{12} - 2s_{14} + s_{13})s_{24}s_{34}P_{1|243|5,6,7,8} \\
& \quad + (s_{13} - 2s_{12} + s_{14})s_{23}s_{24}P_{1|324|5,6,7,8} + (2, 3, 4|2, 3, 4, 5, 6, 7, 8)],
\end{aligned}$$

and it can be used to obtain the (tentatively unique) BRST-closed manifestly local combination at nine points.

B.2. BRST-closed expressions using momentum conservation

We shall now repeat the above analysis in presence of momentum conservation and count the number of manifestly local BRST invariants in an n -particle phase space. At five points, $k_{12345} = 0$ gives rise to four independent local BRST invariants obtained from permutations of

$$\begin{aligned}
QD_{1|2|3,4,5} &= \Delta_{1|2,3,4,5} + k_2^m C_{1|2,3,4,5}^m + [s_{23}C_{1|23,4,5} + (3 \leftrightarrow 4, 5)] \tag{B.13} \\
&= Y_{1,2,3,4,5} + k_2^m V_1 T_{2,3,4,5}^m - V_{12}T_{3,4,5} + [V_1 T_{23,4,5} + (3 \leftrightarrow 4, 5)]
\end{aligned}$$

in $2 \leftrightarrow 3, 4, 5$. The earlier solution (B.5) then follows from a sum over the $2 \leftrightarrow 3, 4, 5$ permutations of (B.13) via momentum conservation.

The non-obvious locality of $k_2^m C_{1|2,3,4,5}^m + [s_{23} C_{1|23,4,5} + (3 \leftrightarrow 4, 5)]$ is not altered when promoting $C_{1|\dots} \rightarrow L_{1|\dots}$, and we can identify a six-point BRST generator from the first line of (B.13),

$$\begin{aligned}
& Q(k_2^m L_{1|2,3,4,5,6}^m + [s_{23} L_{1|23,4,5,6} + (3 \leftrightarrow 4, 5, 6)]) \\
&= k_2^m \Delta_{1|2,3,4,5,6}^m + [s_{23} \Delta_{1|23,4,5,6} + (3 \leftrightarrow 4, 5, 6)] \\
&\quad + k_1^m k_2^n C_{1|2,3,4,5,6}^{mn} - s_{12} P_{1|2|3,4,5,6} + [s_{23} k_1^m C_{1|23,4,5,6}^m + (3 \leftrightarrow 4, 5, 6)] \\
&= k_2^m Y_{1,2,\dots,6}^m - Y_{12,3,4,5,6} + [Y_{23,1,4,5,6} + (3 \leftrightarrow 4, 5, 6)] \\
&\quad + V_1 k_1^m k_2^n T_{2,3,4,5,6}^{mn} - s_{12} V_1 J_{2|3,4,5,6} - V_{12} k_1^m T_{3,4,5,6}^m \\
&\quad + [V_1 k_1^m T_{23,4,5,6}^m + V_{13} k_2^m T_{2,4,5,6}^m + V_{312} T_{4,5,6} + (3 \leftrightarrow 4, 5, 6)] \\
&\quad + [V_{13} T_{24,5,6} + V_{14} T_{23,5,6} + (3, 4|3, 4, 5, 6)] .
\end{aligned} \tag{B.14}$$

The locality of the right-hand side is clear from the locality of the BRST generator. Note that the first line $k_2^m Y_{1,2,\dots,6}^m - Y_{12,3,4,5,6} + [Y_{23,1,4,5,6} + (3 \leftrightarrow 4, 5, 6)]$ is separately BRST closed under $k_2 \cdot k_{13456} = 0$. They are expressible in terms of the BRST exact $\Delta_{1|\dots}$ and can be viewed as the anomaly analogue of the second line of (B.13). The five permutations of the anomalous and the non-anomalous terms on the right-hand side of (B.14) under $2 \leftrightarrow 3, 4, 5, 6$ exhaust the 5 + 5 manifestly local BRST invariants at six points as obtained via a brute-force search with FORM.

At seven points, the BRST descendant in the third line of (B.14) can be promoted to a BRST generator whose locality is guaranteed by the last three lines of (B.14),

$$\begin{aligned}
& Q(k_1^m k_2^n L_{1|2,3,4,5,6,7}^{mn} - s_{12} L_{1|2|3,4,5,6,7} + [s_{23} k_1^m L_{1|23,4,5,6,7}^m + (3 \leftrightarrow 4, 5, 6, 7)]) \\
&= k_1^m k_2^n \Delta_{1|2,3,4,5,6,7}^{mn} - s_{12} \Delta_{1|2|3,4,5,6,7} + [s_{23} k_1^m \Delta_{1|23,4,5,6,7}^m + (3 \leftrightarrow 4, 5, 6, 7)] \\
&\quad + k_1^m k_1^n k_2^p C_{1|2,3,\dots,7}^{mnp} + [s_{23} k_1^m k_1^n C_{1|23,4,5,6,7}^{mn} + (3 \leftrightarrow 4, 5, 6, 7)] - 2s_{12} k_1^m P_{1|2|3,4,5,6,7}^m \\
&\quad + [s_{13} k_2^m P_{1|3|2,4,5,6,7}^m + s_{13} s_{23} P_{1|23|4,5,6,7} + (3 \leftrightarrow 4, 5, 6, 7)] \\
&\quad + [s_{14} s_{23} P_{1|4|23,5,6,7} + s_{13} s_{24} P_{1|3|24,5,6,7} + (3, 4|3, 4, 5, 6, 7)] .
\end{aligned} \tag{B.15}$$

Hence, the right-hand side has to be local as well, as can be verified by explicit computation,

$$\begin{aligned}
& Q(k_1^m k_2^n L_{1|2,3,4,5,6,7}^{mn} - s_{12} L_{1|2|3,4,5,6,7} + [s_{23} k_1^m L_{1|23,4,5,6,7}^m + (3 \leftrightarrow 4, 5, 6, 7)]) \\
&= k_1^m k_2^n Y_{1,2,3,4,5,6,7}^{mn} - s_{12} Y_{2|1,3,4,5,6,7} - k_1^m Y_{12,3,4,5,6,7}^m \\
&\quad + [k_1^m Y_{1,23,4,5,6,7}^m + k_2^m Y_{13,2,4,5,6,7}^m + Y_{312,4,5,6,7} + (3 \leftrightarrow 4, 5, 6, 7)]
\end{aligned}$$

$$\begin{aligned}
& + [Y_{13,24,5,6,7} + Y_{14,23,5,6,7} + (3, 4|3, 4, 5, 6, 7)] \\
& + k_1^m k_1^n k_2^p V_1 T_{2,3,4,5,6,7}^{mnp} - 2s_{12} V_1 k_1^m J_{2|3,4,5,6,7}^m + V_1 k_2^m [s_{13} J_{3|2,4,5,6,7}^m + (3 \leftrightarrow 4, 5, 6, 7)] \\
& + [V_{13} k_2^m (2k_1^n + k_3^n) T_{2,4,5,6,7}^{mn} + k_1^m k_1^n V_1 T_{23,4,5,6,7}^{mn} - 2V_{13} s_{12} J_{2|4,5,6,7} + (3 \leftrightarrow 4, 5, 6, 7)] \\
& - V_{12} k_1^m k_1^n T_{3,4,5,6,7}^{mn} + [s_{13} (V_1 J_{23|4,5,6,7} - V_{12} J_{3|4,5,6,7}) - s_{23} V_{13} J_{2|4,5,6,7} + (3 \leftrightarrow 4, 5, 6, 7)] \\
& + [V_1 (s_{13} J_{3|24,5,6,7} + (3 \leftrightarrow 4)) + (3, 4|3, 4, 5, 6, 7)] - [V_{132} (2k_1^m + k_3^m) T_{4,5,6,7}^m + (3 \leftrightarrow 4, 5, 6, 7)] \\
& + [V_{14} (2k_1^m + k_4^m) T_{23,5,6,7}^m + V_{13} (2k_1^m + k_3^m) T_{24,5,6,7}^m + (3, 4|3, 4, 5, 6, 7)] \tag{B.16} \\
& + [(V_{134} + V_{143}) k_2^m T_{2,5,6,7}^m - (V_{1342} + V_{1432}) T_{5,6,7} + (3, 4|3, 4, 5, 6, 7)] \\
& + [(V_{145} + V_{154}) T_{23,6,7} + (V_{135} + V_{153}) T_{24,6,7} + (V_{134} + V_{143}) T_{25,6,7} + (3, 4, 5|3, 4, 5, 6, 7)] .
\end{aligned}$$

In contrast to the six-point analogue (B.14), the anomalous terms

$$\begin{aligned}
Y_{1|2|3,4,5,6,7}^{(7)} & \equiv k_1^m k_2^n Y_{1,2,3,4,5,6,7}^{mn} - s_{12} Y_{2|1,3,4,5,6,7} - k_1^m Y_{12,3,4,5,6,7}^m \\
& + [k_1^m Y_{1,23,4,5,6,7}^m + k_2^m Y_{13,2,4,5,6,7}^m + Y_{312,4,5,6,7} + (3 \leftrightarrow 4, 5, 6, 7)] \\
& + [Y_{13,24,5,6,7} + Y_{14,23,5,6,7} + (3, 4|3, 4, 5, 6, 7)] \tag{B.17} .
\end{aligned}$$

in the first three lines on the right-hand side of (B.16) are not BRST invariant by themselves. Instead, we have

$$\begin{aligned}
QY_{1|2|3,4,5,6,7}^{(7)} & = -s_{12} \left(V_1 k_2^m Y_{2,3,4,5,6,7}^m - V_{12} Y_{3,4,5,6,7} + [V_1 Y_{23,4,5,6,7} + (3 \leftrightarrow 4, 5, 6, 7)] \right) \\
& = -s_{12} Q\Delta_{1|2|3,4,5,6,7} , \tag{B.18}
\end{aligned}$$

consistent with (B.15) and the fact that $Q\Delta_{1|2,3,\dots,7}^{mn} = Q\Delta_{1|23,\dots,7}^m = 0$. Accordingly, $\langle Y_{1|2|3,4,5,6,7}^{(7)} \rangle = -s_{12} \langle \Delta_{1|2|3,4,5,6,7} \rangle$ in the cohomology. The six permutations of (B.16) under $2 \leftrightarrow 3, \dots, 7$ are the only manifestly local BRST invariants at seven points which can be built from our alphabet of building blocks. By the availability of the BRST generator on the left-hand side of (B.16), all of them are excluded from the cohomology.

At eight points, a brute-force search with FORM identified seven BRST invariant linear combinations of manifestly local building blocks, namely the $2 \leftrightarrow 3, \dots, 8$ permutations of

$$\begin{aligned}
Q & \left(k_1^m k_1^n k_2^p L_{1|2,3,\dots,8}^{mnp} + [s_{23} k_1^m k_1^n L_{1|23,4,\dots,8}^{mn} + (3 \leftrightarrow 4, 5, 6, 7, 8)] - 2s_{12} k_1^m L_{1|2|3,4,\dots,8}^m \right. \\
& + [s_{13} k_2^m L_{1|3|2,4,\dots,8}^m + s_{13} s_{23} L_{1|23|4,\dots,8} + (3 \leftrightarrow 4, \dots, 8)] \tag{B.19} \\
& \left. + [s_{14} s_{23} L_{1|4|23,5,6,7,8} + s_{13} s_{24} L_{1|3|24,5,6,7,8} + (3, 4|3, 4, 5, 6, 7, 8)] \right) \\
& = -s_{12} (k_2^m \Lambda_{1|2,3,\dots,8}^m + [s_{23} \Lambda_{1|23,4,\dots,8} + (3 \leftrightarrow 4, \dots, 8)])
\end{aligned}$$

$$\begin{aligned}
& + k_1^m k_1^n k_2^p \Delta_{1|2,3,\dots,8}^{mnp} + [s_{23} k_1^m k_1^n \Delta_{1|23,4,\dots,8}^{mn} + (3 \leftrightarrow 4, 5, 6, 7, 8)] - 2s_{12} k_1^m \Delta_{1|2|3,4,\dots,8}^m \\
& + [s_{13} k_2^m \Delta_{1|3|2,4,\dots,8}^m + s_{13} s_{23} \Delta_{1|23|4,\dots,8} + (3 \leftrightarrow 4, \dots, 8)] \\
& + [s_{14} s_{23} \Delta_{1|4|23,5,6,7,8} + s_{13} s_{24} \Delta_{1|3|24,5,6,7,8} + (3, 4|3, 4, 5, 6, 7, 8)] \\
& + k_1^m k_1^n k_1^p k_2^q C_{1|2,3,\dots,8}^{mnpq} + [s_{23} k_1^m k_1^n k_1^p C_{1|23,4,\dots,8}^{mnp} + (3 \leftrightarrow 4, \dots, 8)] \\
& - 3s_{12} k_1^m k_1^n P_{1|2|3,4,\dots,8}^{mn} + [s_{13} k_2^m (3k_1^n + k_3^n) P_{1|3|2,4,\dots,8}^{mn} + (3 \leftrightarrow 4, \dots, 8)] \\
& + [s_{13} s_{23} (3k_1^m + k_3^m) P_{1|23|4,5,\dots,8}^m - s_{13} (3s_{12} + s_{23}) P_{1|2,3|4,5,6,7,8} + (3 \leftrightarrow 4, \dots, 8)] \\
& + [(s_{13} - s_{14}) (s_{34} k_2^m P_{1|34|2,5,6,7,8}^m + s_{23} s_{34} P_{1|234|5,6,7,8} - s_{24} s_{34} P_{1|243|5,6,7,8}) + (3, 4|3, 4, \dots, 8)] \\
& + [s_{14} s_{23} (3k_1^m + k_4^m) P_{1|4|23,5,6,7,8}^m + s_{13} s_{24} (3k_1^m + k_3^m) P_{1|3|24,5,6,7,8}^m + (3, 4|3, 4, \dots, 8)] \\
& + [s_{23} s_{45} (s_{14} - s_{15}) P_{1|45|23,6,7,8} + s_{24} s_{35} (s_{13} - s_{15}) P_{1|35|24,6,7,8} \\
& \quad + s_{25} s_{34} (s_{13} - s_{14}) P_{1|34|25,6,7,8} + (3, 4, 5|3, 4, 5, 6, 7, 8)] .
\end{aligned}$$

The BRST generator inherits its locality from the seven-point analogue (B.15). Again, the anomalous terms on the right-hand side of (B.19) are not by themselves BRST invariant. Hence, by the BRST generator on the left-hand side of (B.19), none of the seven local BRST invariants at eight points can belong to the cohomology. As long as we do not consider superspace combinations beyond the building blocks of section 4, the manifestly local BRST cohomology is therefore demonstrated to be empty at five to eight points.

Appendix C. Eight-point anomalous building blocks

This appendix is dedicated to the refined anomalous building blocks at eight points that enter the discussion of the eight-point correlator via (III.3.90) and do not vanish in the BRST cohomology,

$$\begin{aligned}
\Delta_{1|23|4,5,6,7,8} &= \mathcal{Y}_{23|1,4,5,6,7,8} + \mathcal{Y}_{3|12|4,5,6,7,8} - \mathcal{Y}_{2|13,4,5,6,7,8} \tag{C.1} \\
&+ \left(\mathcal{Y}_{123,4,5,6,7,8}^m k_3^m + [s_{34} \mathcal{Y}_{1234,5,6,7,8} + (4 \leftrightarrow 5, 6, 7, 8)] - (2 \leftrightarrow 3) \right) \\
\Delta_{1|2|34,5,6,7,8} &= \mathcal{Y}_{2|1,34,5,6,7,8} + \mathcal{Y}_{2|13,4,5,6,7,8} - \mathcal{Y}_{2|14,3,5,6,7,8} \\
&+ k_2^m \mathcal{Y}_{12,34,5,6,7,8}^m + [s_{25} \mathcal{Y}_{125,34,6,7,8} + (5 \leftrightarrow 6, 7, 8)] \\
&+ \left(k_2^m \mathcal{Y}_{124,3,5,6,7,8}^m + s_{24} (\mathcal{Y}_{1234,5,6,7,8} + \mathcal{Y}_{1324,5,6,7,8}) \right. \\
&\quad \left. + [s_{25} \mathcal{Y}_{4125,3,6,7,8} + (5 \leftrightarrow 6, 7, 8)] - (3 \leftrightarrow 4) \right) \\
\Delta_{1|2|3,4,5,6,7,8}^m &= \mathcal{Y}_{2|1,3,4,5,6,7,8}^m + k_2^p \mathcal{Y}_{12,3,4,5,6,7,8}^{pm} \\
&+ [s_{23} \mathcal{Y}_{123,4,5,6,7,8}^m + k_3^m (\mathcal{Y}_{2|13,4,5,6,7,8} - k_2^p \mathcal{Y}_{213,4,5,6,7,8}^p) \\
&\quad - [k_4^m s_{23} \mathcal{Y}_{4123,5,6,7,8} + (4 \leftrightarrow 5, 6, 7, 8)] + (3 \leftrightarrow 4, 5, 6, 7, 8)] .
\end{aligned}$$

Their BRST variations (5.39) can be rewritten such as to expose the kinematic poles

$$\begin{aligned}
Q\Delta_{1|23|4,5,6,7,8} &= \frac{k_2^m V_1 Y_{23,4,\dots,8}^m + V_{231} Y_{4,5,6,7,8} + [V_1 Y_{234,5,6,7,8} + (4 \leftrightarrow 5, 6, 7, 8)]}{s_{23}} \\
&+ \frac{k_3^m V_{12} Y_{3,4,\dots,8}^m - V_{123} Y_{4,5,6,7,8} + [V_{12} Y_{34,5,6,7,8} + (4 \leftrightarrow 5, 6, 7, 8)]}{s_{12}} \\
&- \frac{k_2^m V_{13} Y_{2,4,\dots,8}^m - V_{132} Y_{4,5,6,7,8} + [V_{13} Y_{24,5,6,7,8} + (4 \leftrightarrow 5, 6, 7, 8)]}{s_{13}} \\
&- V_1 Y_{2|3,4,\dots,8} + V_1 Y_{3|2,4,\dots,8} \\
Q\Delta_{1|2|34,5,6,7,8} &= \frac{k_2^m V_1 Y_{2,34,5,\dots,8}^m - V_{12} Y_{34,5,6,7,8} - V_1 Y_{342,5,6,7,8} + [V_1 Y_{34,25,6,7,8} + (5 \leftrightarrow 6, 7, 8)]}{s_{34}} \\
&+ \frac{k_2^m V_{13} Y_{2,4,\dots,8}^m - V_{132} Y_{4,5,6,7,8} + [V_{13} Y_{24,5,6,7,8} + (4 \leftrightarrow 5, 6, 7, 8)]}{s_{13}} \\
&- \frac{k_2^m V_{14} Y_{2,3,\dots,8}^m - V_{142} Y_{3,5,6,7,8} + [V_{14} Y_{23,5,6,7,8} + (3 \leftrightarrow 5, 6, 7, 8)]}{s_{14}} \\
Q\Delta_{1|2|3,4,5,6,7,8}^m &= V_1 k_2^p Y_{2,3,\dots,8}^{mp} - V_{12} Y_{3,4,\dots,8}^m - k_2^m V_1 Y_{2|3,4,\dots,8} + [V_1 Y_{23,4,\dots,8}^m + (3 \leftrightarrow 4, \dots, 8)] \\
&+ \left[k_3^m \frac{k_2^p V_{13} Y_{2,4,\dots,8}^p - V_{132} Y_{4,5,6,7,8} + [V_{13} Y_{24,5,6,7,8} + (4 \leftrightarrow 5, 6, 7, 8)]}{s_{13}} \right. \\
&\quad \left. + (3 \leftrightarrow 4, \dots, 8) \right]
\end{aligned} \tag{C.2}$$

which should therefore be present in the components of the $\Delta_{1|\dots}$ superfields themselves.

Indeed, we shall now write down the bosonic components of the eight-point topology $\Delta_{1|2|34,5,6,7,8}$ in terms of Berends–Giele currents in the BCJ gauge, see [23] for more details.

$$\begin{aligned}
\langle \Delta_{1|2|34,5,6,7,8} \rangle &= +\frac{1}{2} (\mathcal{A}_{12}^{k^2} \mathcal{F}_{34}^{mn} + \mathcal{A}_{34}^{k^2} \mathcal{F}_{12}^{mn}) \epsilon^{k^5 k^6 k^7 k^8 e^5 e^6 e^7 e^8 mn} + \frac{1}{2} \mathcal{A}_{34}^{e^2} \epsilon^{k^1 k^5 k^6 k^7 k^8 e^1 e^5 e^6 e^7 e^8} \\
&- \frac{1}{4} [(\mathcal{F}_{12}^{mn} \mathcal{F}_{34}^{pq} (k^2 \cdot e^5) + \mathcal{F}_{125}^{mn} \mathcal{F}_{34}^{pq} s_{25}) \epsilon^{k^6 k^7 k^8 e^6 e^7 e^8 mnpq} + (5 \leftrightarrow 6, 7, 8)] \\
&+ \frac{1}{2} [(\mathcal{F}_{214}^{mn} (k^2 \cdot e^5) + \mathcal{F}_{4125}^{mn} s_{25}) \epsilon^{k^3 k^6 k^7 k^8 e^3 e^6 e^7 e^8 mn} + (5 \leftrightarrow 6, 7, 8)] - (3 \leftrightarrow 4) \\
&+ \frac{1}{2} [(\mathcal{F}_{214}^{mn} (k^2 \cdot e^3) + \mathcal{F}_{1324}^{mn} s_{24} + \mathcal{F}_{1234}^{mn} s_{24}) \epsilon^{k^5 k^6 k^7 k^8 e^5 e^6 e^7 e^8 mn} - (3 \leftrightarrow 4)] \\
&+ \frac{1}{2} [(\mathcal{A}_{13}^{e^2} - 2\mathcal{A}_{213}^{k^2}) \epsilon^{k^4 k^5 k^6 k^7 k^8 e^4 e^5 e^6 e^7 e^8} - (3 \leftrightarrow 4)] \\
&+ \frac{1}{12} \left[\frac{1}{s_{13}} (((k^1 \cdot e^2) - (k^3 \cdot e^2))(e^1 \cdot e^3) + ((k^2 \cdot e^3) - (k^1 \cdot e^3))(e^1 \cdot e^2) \right. \\
&\quad \left. + ((k^3 \cdot e^1) - (k^2 \cdot e^1))(e^2 \cdot e^3)) \epsilon^{k^4 k^5 k^6 k^7 k^8 e^4 e^5 e^6 e^7 e^8} - (3 \leftrightarrow 4) \right] \\
&+ \frac{1}{12 s_{34}} \left[[(k^3 \cdot e^2)(e^3 \cdot e^4) + ((k^2 \cdot e^4) - (k^3 \cdot e^4))(e^2 \cdot e^3)] \epsilon^{k^1 k^5 k^6 k^7 k^8 e^1 e^5 e^6 e^7 e^8} \right. \\
&\quad \left. - (3 \leftrightarrow 4) \right]
\end{aligned}$$

$$\begin{aligned}
& + \frac{1}{4} [\mathcal{F}_{13}^{mn}(e^2 \cdot e^5) \epsilon^{k^4 k^6 k^7 k^8 e^4 e^6 e^7 e^8 mn} + (5 \leftrightarrow 6, 7, 8)] - (3 \leftrightarrow 4) \\
& + \frac{1}{4} [\mathcal{F}_{34}^{mn}(e^2 \cdot e^5) \epsilon^{k^1 k^6 k^7 k^8 e^1 e^6 e^7 e^8 mn} + (5 \leftrightarrow 6, 7, 8)] \\
& + \frac{1}{4} [\mathcal{F}_{13}^{mn}(e^2 \cdot e^4) \epsilon^{k^5 k^6 k^7 k^8 e^5 e^6 e^7 e^8 mn} - (3 \leftrightarrow 4)] \\
& + \frac{1}{4} \mathcal{F}_{34}^{mn}(e^1 \cdot e^2) \epsilon^{k^5 k^6 k^7 k^8 e^5 e^6 e^7 e^8 mn} \Big|_{\theta=0} + \text{fermions} . \tag{C.3}
\end{aligned}$$

Given that θ and fermionic wavefunctions are suppressed on the right-hand side, the superfields \mathcal{A}_P^m and \mathcal{F}_P^{mn} reduce to the bosonic Berends–Giele currents in the BCJ gauge [23]. The other topologies in (C.1) have similar expansions but were omitted and can be found as computer-readable files attached to the arXiv submission. Although it is not manifest in the form presented above, all four-channel s_{ijkl} poles turn out to be absent²⁵ in every eight-point topology of Δ . The three-channel poles s_{ijk} are however present (despite being absent in the BRST variation).

Appendix D. The BRST variations of local building blocks

In this appendix we list the BRST variations of every local building block appearing in the one-loop correlators up to eight points that have not appeared as examples in the preceding sections. The list below can be generated from the formulas given in the main text but are presented here for convenience. Together with the examples from the main text, these equations allow one to verify all claims related to BRST variations in this series of papers.

D.1. Scalar $T_{A,B,C}$

Similarly, the BRST variations given in (4.3) are appended by

$$\begin{aligned}
QT_{1234,5,6} &= (k_1 \cdot k_2) [V_1 T_{234,5,6} + V_{13} T_{24,5,6} + V_{134} T_{2,5,6} + V_{14} T_{23,5,6} - (1 \leftrightarrow 2)] \tag{D.1} \\
&+ (k_{12} \cdot k_3) [V_{12} T_{34,5,6} + V_{124} T_{3,5,6} - (12 \leftrightarrow 3)] \\
&+ (k_{123} \cdot k_4) [V_{123} T_{4,5,6} - (123 \leftrightarrow 4)] , \\
QT_{123,45,6} &= (k_1 \cdot k_2) [V_1 T_{23,45,6} + V_{13} T_{2,45,6} - V_{23} T_{1,45,6} - V_2 T_{13,45,6}]
\end{aligned}$$

²⁵ This can be seen by inserting $\mathcal{F}_{1234}^{mn} = k_{1234}^m \mathcal{A}_{1234}^n - \mathcal{A}_{123}^m \mathcal{A}_4^n - \mathcal{A}_{12}^m \mathcal{A}_{34}^n - \mathcal{A}_1^m \mathcal{A}_{234}^n - (m \leftrightarrow n)$ into $\mathcal{F}_{1234}^{mn} \epsilon^{k^5 k^6 k^7 k^8 e^5 e^6 e^7 e^8 mn}$ and noting the consequence $k_{1234}^m \mathcal{A}_{1234}^n \epsilon^{k^5 k^6 k^7 k^8 e^5 e^6 e^7 e^8 mn} = 0$ of momentum conservation.

$$\begin{aligned}
& + (k_{12} \cdot k_3) [V_{12}T_{3,45,6} - V_3T_{12,45,6}] \\
& + (k_4 \cdot k_5) [V_4T_{123,5,6} - V_5T_{123,4,6}], \\
QT_{12,34,56} & = (k_1 \cdot k_2) [V_1T_{2,34,56} - V_2T_{1,34,56}] + (12 \leftrightarrow 34, 56), \\
QT_{12345,6,7} & = (k_1 \cdot k_2) [V_1T_{2345,6,7} + V_{13}T_{245,6,7} + V_{134}T_{25,6,7} + V_{1345}T_{2,6,7} \\
& \quad + V_{135}T_{24,6,7} + V_{14}T_{235,6,7} + V_{145}T_{23,6,7} + V_{15}T_{234,6,7} - (1 \leftrightarrow 2)] \\
& + (k_{12} \cdot k_3) [V_{12}T_{345,6,7} + V_{124}T_{35,6,7} + V_{1245}T_{3,6,7} + V_{125}T_{34,6,7} - (12 \leftrightarrow 3)] \\
& + (k_{123} \cdot k_4) [V_{123}T_{45,6,7} + V_{1235}T_{4,6,7} - (123 \leftrightarrow 4)] \\
& + (k_{1234} \cdot k_5) [V_{1234}T_{5,6,7} - (1234 \leftrightarrow 5)], \\
QT_{1234,56,7} & = (k_1 \cdot k_2) [V_1T_{234,56,7} + V_{13}T_{24,56,7} + V_{134}T_{56,2,7} + V_{14}T_{23,56,7} - (1 \leftrightarrow 2)] \\
& + (k_{12} \cdot k_3) [V_{12}T_{34,56,7} + V_{124}T_{56,3,7} - (12 \leftrightarrow 3)] \\
& + (k_{123} \cdot k_4) [V_{123}T_{56,4,7} - (123 \leftrightarrow 4)] \\
& + (k_5 \cdot k_6) [V_5T_{1234,6,7} - (5 \leftrightarrow 6)] \\
QT_{123,456,7} & = (k_1 \cdot k_2) [V_1T_{456,23,7} + V_{13}T_{456,2,7} - (1 \leftrightarrow 2)] \\
& + (k_{12} \cdot k_3) [V_{12}T_{456,3,7} - (12 \leftrightarrow 3)] \\
& + (k_4 \cdot k_5) [V_4T_{123,56,7} + V_{46}T_{123,5,7} - (4 \leftrightarrow 5)] \\
& + (k_{45} \cdot k_6) [V_{45}T_{123,6,7} - (45 \leftrightarrow 6)] \\
QT_{123,45,67} & = (k_1 \cdot k_2) [V_1T_{23,45,67} + V_{13}T_{45,67,2} - (1 \leftrightarrow 2)] \\
& + (k_{12} \cdot k_3) [V_{12}T_{45,67,3} - (12 \leftrightarrow 3)] \\
& + (k_4 \cdot k_5) [V_4T_{123,67,5} - (4 \leftrightarrow 5)] \\
& + (k_6 \cdot k_7) [V_6T_{123,45,7} - (6 \leftrightarrow 7)].
\end{aligned}$$

D.2. Tensorial $T_{A,B,C,\dots}^m$

The vectorial BRST variations in (4.8) generalize to

$$\begin{aligned}
QT_{1234,5,6,7}^m & = [k_{1234}^m V_{1234}T_{5,6,7} + (1234 \leftrightarrow 5, 6, 7)] \tag{D.2} \\
& + (k_1 \cdot k_2) [V_1T_{234,5,6,7}^m + V_{13}T_{24,5,6,7}^m + V_{14}T_{23,5,6,7}^m + V_{134}T_{2,5,6,7}^m - (1 \leftrightarrow 2)] \\
& + (k_{12} \cdot k_3) [V_{12}T_{34,5,6,7}^m + V_{124}T_{3,5,6,7}^m - (12 \leftrightarrow 3)] \\
& + (k_{123} \cdot k_4) [V_{123}T_{4,5,6,7}^m - (123 \leftrightarrow 4)] \\
QT_{123,45,6,7}^m & = [k_{123}^m V_{123}T_{45,6,7} + (123 \leftrightarrow 45, 6, 7)] \\
& + (k_1 \cdot k_2) [V_1T_{23,45,6,7}^m + V_{13}T_{2,45,6,7}^m - (1 \leftrightarrow 2)]
\end{aligned}$$

$$\begin{aligned}
& + (k_{12} \cdot k_3) [V_{12} T_{3,45,6,7}^m - (12 \leftrightarrow 3)] \\
& + (k_4 \cdot k_5) [V_4 T_{123,5,6,7}^m - (4 \leftrightarrow 5)] \\
QT_{12,34,56,7}^m & = [k_{12}^m V_{12} T_{34,56,7} + (12 \leftrightarrow 34, 56, 7)] \\
& + [(k_1 \cdot k_2)(V_1 T_{2,34,56,7}^m - (1 \leftrightarrow 2)) + (12 \leftrightarrow 34, 56)],
\end{aligned}$$

while the higher-multiplicity analogues of the tensorial BRST variations (4.17) read

$$\begin{aligned}
QT_{123,4,5,6,7}^{mn} & = \delta^{mn} Y_{123,4,5,6,7} + [k_{123}^{(m)} V_{123} T_{4,5,6,7}^n + (123 \leftrightarrow 4, 5, 6, 7)] \quad (D.3) \\
& + (k_1 \cdot k_2) [V_1 T_{23,4,5,6,7}^{mn} + V_{13} T_{2,4,5,6,7}^{mn} - (1 \leftrightarrow 2)] \\
& + (k_{12} \cdot k_3) [V_{12} T_{3,4,5,6,7}^{mn} - (12 \leftrightarrow 3)] \\
QT_{12,34,5,6,7}^{mn} & = \delta^{mn} Y_{12,34,5,6,7} + [k_{12}^{(m)} V_{12} T_{34,5,6,7}^n + (12 \leftrightarrow 34, 5, 6, 7)] \\
& + (k_1 \cdot k_2) [V_1 T_{2,34,5,6,7}^{mn} - (1 \leftrightarrow 2)] \\
& + (k_3 \cdot k_4) [V_3 T_{12,4,5,6,7}^{mn} - (3 \leftrightarrow 4)] \\
QT_{12,3,4,5,6,7}^{mnp} & = \delta^{(mn} Y_{12,3,4,5,6,7}^{p)} + [k_{12}^{(m)} V_{12} T_{3,4,5,6,7}^{np} + (12 \leftrightarrow 34, 5, 6, 7)] \\
& + (k_1 \cdot k_2) [V_1 T_{2,3,4,5,6,7}^{mnp} - (1 \leftrightarrow 2)] \\
QT_{1,2,3,4,5,6,7}^{mnpq} & = \delta^{(mn} Y_{1,2,3,4,5,6,7}^{pq)} + [k_1^{(m)} V_1 T_{2,3,4,5,6,7}^{npq} + (1 \leftrightarrow 2, 3, 4, 5, 6, 7)].
\end{aligned}$$

D.3. Refined $J_{A|B,C,\dots}^{m\dots}$

The BRST variations of the refined building blocks $J_{A|B,C,D,E}$ that appear in the eight-point correlator read,

$$\begin{aligned}
QJ_{123|4,5,6,7} & = k_{123}^m V_{123} T_{4,5,6,7}^m + [V_{[123,4]} T_{5,6,7} + (4 \leftrightarrow 5, 6, 7)] + Y_{123,4,5,6,7} \quad (D.4) \\
& + (k^1 \cdot k^2) [V_1 J_{23|4,5,6,7} + V_{13} J_{2|4,5,6,7} - (1 \leftrightarrow 2)] \\
& + (k^{12} \cdot k^3) [V_{12} J_{3|4,5,6,7} - (12 \leftrightarrow 3)] \\
QJ_{12|34,5,6,7} & = k_{12}^m V_{12} T_{34,5,6,7}^m + [V_{[12,34]} T_{5,6,7} + (34 \leftrightarrow 5, 6, 7)] + Y_{12,34,5,6,7} \\
& + (k^1 \cdot k^2) [V_1 J_{2|34,5,6,7} - (1 \leftrightarrow 2)] + (k^3 \cdot k^4) [V_3 J_{12|4,5,6,7} - (3 \leftrightarrow 4)] \\
QJ_{1|23,45,6,7} & = k_1^m V_1 T_{23,45,6,7}^m + [V_{[1,23]} T_{45,6,7} + (23 \leftrightarrow 45, 6, 7)] + Y_{1,23,45,6,7} \\
& + (k^2 \cdot k^3) [V_2 J_{1|3,45,6,7} - (2 \leftrightarrow 3)] + (k^4 \cdot k^5) [V_4 J_{1|23,5,6,7} - (4 \leftrightarrow 5)] \\
QJ_{1|234,5,6,7} & = k_1^m V_1 T_{234,5,6,7}^m + [V_{[1,234]} T_{5,6,7} + (234 \leftrightarrow 5, 6, 7)] + Y_{1,234,5,6,7} \\
& + (k^2 \cdot k^3) [V_2 J_{1|34,5,6,7} + V_{24} J_{1|3,5,6,7} - (2 \leftrightarrow 3)] \\
& + (k^{23} \cdot k^4) [V_{23} J_{1|4,5,6,7} - (23 \leftrightarrow 4)],
\end{aligned}$$

see (4.26) for examples at lower multiplicity. The BRST variation of every tensorial $J^{m\dots}$ relevant to one-loop correlators up to eight points has been spelled out in (4.31).

D.4. Anomaly building blocks $Y_{A,\bar{B},C,\dots}^{m,\dots}$

The BRST variations (4.21) of anomaly building blocks generalize as follows to multiplicity seven and eight,

$$\begin{aligned}
QY_{123,4,5,6,7} &= (k_1 \cdot k_2) [V_1 Y_{23,4,5,6,7} + V_{13} Y_{2,4,5,6,7} - (1 \leftrightarrow 2)] \\
&\quad + (k_{12} \cdot k_3) [V_{12} Y_{3,4,5,6,7} - (12 \leftrightarrow 3)] \\
QY_{12,34,5,6,7} &= (k_1 \cdot k_2) [V_1 Y_{2,34,5,6,7} - (1 \leftrightarrow 2)] + (12 \leftrightarrow 34) \\
QY_{1234,5,6,7,8} &= (k_1 \cdot k_2) [V_1 Y_{234,5,6,7,8} + V_{13} Y_{24,5,6,7,8} + V_{14} Y_{23,5,6,7,8} + V_{134} Y_{2,5,6,7,8} - (1 \leftrightarrow 2)] \\
&\quad + (k_{12} \cdot k_3) [V_{12} Y_{34,5,6,7,8} + V_{124} Y_{3,5,6,7,8} - (12 \leftrightarrow 3)] \\
&\quad + (k_{123} \cdot k_4) [V_{123} Y_{4,5,6,7,8} - (123 \leftrightarrow 4)] \\
QY_{123,45,6,7,8} &= (k_1 \cdot k_2) [V_1 Y_{23,45,6,7,8} + V_{13} Y_{2,45,6,7,8} - (1 \leftrightarrow 2)] \\
&\quad + (k_{12} \cdot k_3) [V_{12} Y_{3,45,6,7,8} - (12 \leftrightarrow 3)] \\
&\quad + (k_4 \cdot k_5) [V_4 Y_{123,5,6,7,8} - (4 \leftrightarrow 5)] \\
QY_{12,34,56,7,8} &= (k_1 \cdot k_2) [V_1 Y_{2,34,56,7,8} - (1 \leftrightarrow 2)] + (12 \leftrightarrow 34, 56) \\
QY_{12,3,4,5,6,7}^m &= [k_{12}^m V_{12} Y_{3,4,5,6,7} + (12 \leftrightarrow 3, 4, 5, 6, 7, 8)] \\
&\quad + (k_1 \cdot k_2) [V_1 Y_{2,3,4,5,6,7}^m - (1 \leftrightarrow 2)] \\
QY_{123,4,5,6,7,8}^m &= [k_{123}^m V_{123} Y_{4,5,6,7,8} + (123 \leftrightarrow 4, 5, 6, 7, 8)] \\
&\quad + (k_1 \cdot k_2) [V_1 Y_{23,4,5,6,7,8}^m + V_{13} Y_{2,4,5,6,7,8}^m - (1 \leftrightarrow 2)] \\
&\quad + (k_{12} \cdot k_3) [V_{12} Y_{3,4,5,6,7,8}^m - (12 \leftrightarrow 3)] \\
QY_{12,34,5,6,7,8}^m &= [k_{12}^m V_{12} Y_{34,5,6,7,8} + (12 \leftrightarrow 34, 5, 6, 7, 8)] \\
&\quad + (k_1 \cdot k_2) [V_1 Y_{2,34,5,6,7,8}^m - (1 \leftrightarrow 2)] + (k_3 \cdot k_4) [V_3 Y_{12,4,5,6,7,8}^m - (3 \leftrightarrow 4)] \\
QY_{1,2,3,4,5,6,7}^{mn} &= k_1^{(m)} V_1 Y_{2,3,4,5,6,7}^{(n)} + (1 \leftrightarrow 2, 3, 4, 5, 6, 7) \\
QY_{12,3,4,5,6,7,8}^{mn} &= [k_{12}^{(m)} V_{12} Y_{3,4,5,6,7,8}^{(n)} + (12 \leftrightarrow 3, 4, 5, 6, 7, 8)] + (k_1 \cdot k_2) [V_1 Y_{2,3,4,5,6,7,8}^{mn} - (1 \leftrightarrow 2)] \\
QY_{1,2,3,4,5,6,7,8}^{mnp} &= k_1^{(m)} V_1 Y_{2,3,4,5,6,7,8}^{(np)} + (1 \leftrightarrow 2, 3, 4, 5, 6, 7, 8).
\end{aligned} \tag{D.5}$$

The BRST variations of refined anomaly building blocks relevant to $(n \leq 8)$ -point correlators are spelled out in (4.34).

References

- [1] C.R. Mafra and O. Schlotterer, “Towards the n-point one-loop superstring amplitude I: Pure spinors and superfield kinematics”, [arXiv:1812.xxxxx [hep-th]];
C.R. Mafra and O. Schlotterer, “Towards the n-point one-loop superstring amplitude II: Worldsheet functions and their duality to kinematics”, [arXiv:1812.xxxxx [hep-th]];
C.R. Mafra and O. Schlotterer, “Towards the n-point one-loop superstring amplitude III: One-loop correlators and their double-copy structure”, [arXiv:1812.xxxxx [hep-th]]
- [2] N. Berkovits, “Super-Poincare covariant quantization of the superstring,” *JHEP* **0004**, 018 (2000) [arXiv:hep-th/0001035].
- [3] N. Berkovits, “Multiloop amplitudes and vanishing theorems using the pure spinor formalism for the superstring,” *JHEP* **0409**, 047 (2004). [hep-th/0406055].
- [4] N. Berkovits and B. C. Vallilo, “Consistency of superPoincare covariant superstring tree amplitudes,” *JHEP* **0007**, 015 (2000). [hep-th/0004171].
- [5] N. Berkovits, “Super-Poincare covariant two-loop superstring amplitudes,” *JHEP* **0601**, 005 (2006). [hep-th/0503197]. ;
N. Berkovits and C.R. Mafra, “Equivalence of two-loop superstring amplitudes in the pure spinor and RNS formalisms,” *Phys. Rev. Lett.* **96**, 011602 (2006). [hep-th/0509234]. ;
H. Gomez and C.R. Mafra, “The Overall Coefficient of the Two-loop Superstring Amplitude Using Pure Spinors,” *JHEP* **1005**, 017 (2010). [arXiv:1003.0678 [hep-th]].
- [6] N. Berkovits, “Cohomology in the pure spinor formalism for the superstring,” *JHEP* **0009**, 046 (2000). [hep-th/0006003]. ;
N. Berkovits and O. Chandia, “Lorentz invariance of the pure spinor BRST cohomology for the superstring,” *Phys. Lett. B* **514**, 394 (2001). [hep-th/0105149].
- [7] N. Berkovits, “Pure spinor formalism as an N=2 topological string,” *JHEP* **0510**, 089 (2005). [hep-th/0509120].
- [8] H. Gomez and C.R. Mafra, “The closed-string 3-loop amplitude and S-duality,” *JHEP* **1310**, 217 (2013). [arXiv:1308.6567 [hep-th]].
- [9] C.R. Mafra, O. Schlotterer and S. Stieberger, “Complete N-Point Superstring Disk Amplitude I. Pure Spinor Computation,” *Nucl. Phys. B* **873**, 419 (2013). [arXiv:1106.2645 [hep-th]].
- [10] C.R. Mafra, O. Schlotterer, S. Stieberger and D. Tsimpis, “A recursive method for SYM n-point tree amplitudes,” *Phys. Rev. D* **83**, 126012 (2011). [arXiv:1012.3981 [hep-th]].
- [11] Z. Bern, J. J. M. Carrasco and H. Johansson, “New Relations for Gauge-Theory Amplitudes,” *Phys. Rev. D* **78**, 085011 (2008). [arXiv:0805.3993 [hep-ph]].
- [12] C.R. Mafra, O. Schlotterer and S. Stieberger, “Explicit BCJ Numerators from Pure Spinors,” *JHEP* **1107**, 092 (2011). [arXiv:1104.5224 [hep-th]].

- [13] N. E. J. Bjerrum-Bohr, P. H. Damgaard and P. Vanhove, “Minimal Basis for Gauge Theory Amplitudes,” *Phys. Rev. Lett.* **103**, 161602 (2009). [arXiv:0907.1425 [hep-th]].
- [14] S. Stieberger, “Open & Closed vs. Pure Open String Disk Amplitudes,” [arXiv:0907.2211 [hep-th]].
- [15] J. Broedel, O. Schlotterer and S. Stieberger, “Polylogarithms, Multiple Zeta Values and Superstring Amplitudes,” *Fortsch. Phys.* **61**, 812 (2013). [arXiv:1304.7267 [hep-th]].
- [16] H. Kawai, D. C. Lewellen and S. H. H. Tye, “A Relation Between Tree Amplitudes of Closed and Open Strings,” *Nucl. Phys. B* **269**, 1 (1986).
- [17] C.R. Mafra and O. Schlotterer, “Double-Copy Structure of One-Loop Open-String Amplitudes,” *Phys. Rev. Lett.* **121**, no. 1, 011601 (2018). [arXiv:1711.09104 [hep-th]].
- [18] C.R. Mafra and O. Schlotterer, “Multiparticle SYM equations of motion and pure spinor BRST blocks,” *JHEP* **1407**, 153 (2014). [arXiv:1404.4986 [hep-th]].
- [19] C.R. Mafra and O. Schlotterer, “Towards one-loop SYM amplitudes from the pure spinor BRST cohomology,” *Fortsch. Phys.* **63**, no. 2, 105 (2015). [arXiv:1410.0668 [hep-th]].
- [20] R. Kleiss, H. Kuijf, “Multi - Gluon Cross-sections And Five Jet Production At Hadron Colliders,” *Nucl. Phys.* **B312**, 616 (1989).
- [21] R. Ree, “Lie elements and an algebra associated with shuffles”, *Ann. Math.* **62**, No. 2 (1958), 210–220.
- [22] C. Reutenauer, “Free Lie Algebras”, London Mathematical Society Monographs, 1993.
- [23] S. Lee, C.R. Mafra and O. Schlotterer, “Non-linear gauge transformations in $D = 10$ SYM theory and the BCJ duality,” *JHEP* **1603**, 090 (2016). [arXiv:1510.08843 [hep-th]].
- [24] C.R. Mafra and O. Schlotterer, “Cohomology foundations of one-loop amplitudes in pure spinor superspace,” [arXiv:1408.3605 [hep-th]].
- [25] E.P. Verlinde and H.L. Verlinde, “Chiral Bosonization, Determinants and the String Partition Function,” *Nucl. Phys. B* **288**, 357 (1987).
- [26] E. D’Hoker and D. H. Phong, “The Geometry of String Perturbation Theory,” *Rev. Mod. Phys.* **60**, 917 (1988).
- [27] E. D’Hoker and D. H. Phong, “Conformal Scalar Fields and Chiral Splitting on Superriemann Surfaces,” *Commun. Math. Phys.* **125**, 469 (1989).
- [28] Z. Bern, J. J. M. Carrasco and H. Johansson, “Perturbative Quantum Gravity as a Double Copy of Gauge Theory,” *Phys. Rev. Lett.* **105**, 061602 (2010). [arXiv:1004.0476 [hep-th]].
- [29] Z. Bern, J. J. Carrasco, W. M. Chen, H. Johansson and R. Roiban, “Gravity Amplitudes as Generalized Double Copies of Gauge-Theory Amplitudes,” *Phys. Rev. Lett.* **118**, no. 18, 181602 (2017). [arXiv:1701.02519 [hep-th]].

- [30] M. B. Green, J. H. Schwarz and E. Witten, “Superstring Theory. Vol. 2: Loop Amplitudes, Anomalies And Phenomenology,” Cambridge University Press (1987).
- [31] C.R. Mafra and C. Stahn, “The One-loop Open Superstring Massless Five-point Amplitude with the Non-Minimal Pure Spinor Formalism,” *JHEP* **0903**, 126 (2009). [arXiv:0902.1539 [hep-th]].
- [32] C.R. Mafra and O. Schlotterer, “One-loop superstring six-point amplitudes and anomalies in pure spinor superspace,” *JHEP* **1604**, 148 (2016). [arXiv:1603.04790 [hep-th]].
- [33] M. B. Green and J. H. Schwarz, “Infinity Cancellations in SO(32) Superstring Theory,” *Phys. Lett.* **151B**, 21 (1985)..
- [34] M.B. Green and J.H. Schwarz, “The Hexagon Gauge Anomaly in Type I Superstring Theory,” *Nucl. Phys. B* **255** (1985) 93. ;
M.B. Green and J.H. Schwarz, “Anomaly Cancellation in Supersymmetric D=10 Gauge Theory and Superstring Theory,” *Phys. Lett. B* **149** (1984) 117.
- [35] I. Antoniadis, C. Bachas, C. Fabre, H. Partouche and T. R. Taylor, “Aspects of type I - type II - heterotic triality in four-dimensions,” *Nucl. Phys. B* **489**, 160 (1997). [hep-th/9608012].
- [36] J. Polchinski, “String theory. Vol. 1: An introduction to the bosonic string,” Cambridge University Press (2007).
- [37] W. Siegel, “Classical Superstring Mechanics,” *Nucl. Phys. B* **263**, 93 (1986).
- [38] M.A.A. van Leeuwen, A.M. Cohen and B. Lisser, “LiE, A Package for Lie Group Computations”, Computer Algebra Nederland, Amsterdam, ISBN 90-74116-02-7, 1992
- [39] N. Berkovits, “Explaining Pure Spinor Superspace,” [hep-th/0612021].
- [40] C.R. Mafra, “Pure Spinor Superspace Identities for Massless Four-point Kinematic Factors,” *JHEP* **0804**, 093 (2008). [arXiv:0801.0580 [hep-th]].
- [41] D. M. Richards, “The One-Loop Five-Graviton Amplitude and the Effective Action,” *JHEP* **0810**, 042 (2008). [arXiv:0807.2421 [hep-th]].
- [42] M. B. Green, C.R. Mafra and O. Schlotterer, “Multiparticle one-loop amplitudes and S-duality in closed superstring theory,” *JHEP* **1310**, 188 (2013). [arXiv:1307.3534 [hep-th]].
- [43] A. Gregori, E. Kiritsis, C. Kounnas, N. A. Obers, P. M. Petropoulos and B. Pioline, “R**2 corrections and nonperturbative dualities of N=4 string ground states,” *Nucl. Phys. B* **510**, 423 (1998). [hep-th/9708062].
- [44] M. Berg, I. Buchberger and O. Schlotterer, “From maximal to minimal supersymmetry in string loop amplitudes,” *JHEP* **1704**, 163 (2017). [arXiv:1603.05262 [hep-th]].
- [45] A. Ochirov and P. Tourkine, “BCJ duality and double copy in the closed string sector,” *JHEP* **1405**, 136 (2014). [arXiv:1312.1326 [hep-th]].
- [46] Lothaire, M., “Combinatorics on Words”, (Cambridge Mathematical Library), Cambridge University Press (1997).

- [47] F.A. Berends and W.T. Giele, “Multiple Soft Gluon Radiation in Parton Processes,” Nucl. Phys. B **313**, 595 (1989).
- [48] F.A. Berends, W.T. Giele, “Recursive Calculations for Processes with n Gluons,” Nucl. Phys. **B306**, 759 (1988).
- [49] E. Witten, “Twistor-Like Transform In Ten-Dimensions,” Nucl. Phys. B **266**, 245 (1986).
- [50] W. Siegel, “Superfields in Higher Dimensional Space-time,” Phys. Lett. B **80**, 220 (1979).
- [51] C.R. Mafra, “Simplifying the Tree-level Superstring Massless Five-point Amplitude,” JHEP **1001**, 007 (2010). [arXiv:0909.5206 [hep-th]].
- [52] J. Polchinski and Y. Cai, “Consistency of Open Superstring Theories,” Nucl. Phys. B **296**, 91 (1988).
- [53] N. Berkovits and C.R. Mafra, “Some Superstring Amplitude Computations with the Non-Minimal Pure Spinor Formalism,” JHEP **0611**, 079 (2006). [hep-th/0607187].
- [54] Z. Bern, L. J. Dixon, M. Perelstein and J. S. Rozowsky, “Multileg one loop gravity amplitudes from gauge theory,” Nucl. Phys. B **546**, 423 (1999). [hep-th/9811140].
- [55] N. E. J. Bjerrum-Bohr, P. H. Damgaard, T. Sondergaard and P. Vanhove, “The Momentum Kernel of Gauge and Gravity Theories,” JHEP **1101**, 001 (2011). [arXiv:1010.3933 [hep-th]].
- [56] F. Cachazo, S. He and E.Y. Yuan, “Scattering of Massless Particles: Scalars, Gluons and Gravitons,” JHEP **1407**, 033 (2014). [arXiv:1309.0885 [hep-th]].
- [57] J. J. M. Carrasco, C.R. Mafra and O. Schlotterer, “Abelian Z-theory: NLSM amplitudes and α' -corrections from the open string,” JHEP **1706**, 093 (2017). [arXiv:1608.02569 [hep-th]].
- [58] C.R. Mafra, “Berends-Giele recursion for double-color-ordered amplitudes,” JHEP **1607**, 080 (2016). [arXiv:1603.09731 [hep-th]].
- [59] M. Schocker, “Lie elements and Knuth relations,” Canad. J. Math. **56** (2004), 871-882. [math/0209327].
- [60] C.R. Mafra and O. Schlotterer, “Berends-Giele recursions and the BCJ duality in superspace and components,” JHEP **1603**, 097 (2016). [arXiv:1510.08846 [hep-th]].
- [61] I. C. Michos, “On twin and anti-twin words in the support of the free Lie algebra.” Journal of Algebraic Combinatorics **36.3** (2012) 355-388.
- [62] C.R. Mafra, “PSS: A FORM Program to Evaluate Pure Spinor Superspace Expressions,” [arXiv:1007.4999 [hep-th]].
- [63] C.R. Mafra, “Four-point one-loop amplitude computation in the pure spinor formalism,” JHEP **0601**, 075 (2006). [hep-th/0512052].
- [64] J.A.M. Vermaseren, “New features of FORM,” arXiv:math-ph/0010025. ;
M. Tentyukov and J.A.M. Vermaseren, “The multithreaded version of FORM,” arXiv:hep-ph/0702279.