DEUTSCHES ELEKTRONEN-SYNCHROTRON
Ein Forschungszentrum der Helmholtz-Gemeinschaft

DESY 21-223
CERN-TH-2021-224
HU-EP-21/59-RTG
arXiv:2112.11453
December 2021

The Seeds of EFT Double Copy

Q. Bonnefoy, C. S. Machado
Deutsches Elektronen-Synchrotron DESY, Hamburg

G. Durieux
CERN, Geneva, Switzerland

C. Grojean
Deutsches Elektronen-Synchrotron DESY, Hamburg
and
Institut für Physik, Humboldt-Universität, Berlin

J. R. Nepveu
Institut für Physik, Humboldt-Universität, Berlin

ISSN 0418-9833

NOTKESTRASSE 85 - 22607 HAMBURG

DESY-21-223
CERN-TH-2021-224
HU-EP-21/59-RTG

The seeds of EFT double copy

Quentin Bonnefoy,[a] Gauthier Durieux,[b] Christophe Grojean,[a,c] Camila S. Machado,[a] Jasper Roosmale Nepveu[c]

[a] *Deutsches Elektronen-Synchrotron DESY, Notkestr. 85, 22607 Hamburg, Germany*
[b] *CERN, Theoretical Physics Department, Geneva 23 CH-1211, Switzerland*
[c] *Humboldt-Universität zu Berlin, Institut für Physik, Newtonstr. 15, 12489 Berlin, Germany*

E-mail: quentin.bonnefoy@desy.de, gauthier.durieux@cern.ch, christophe.grojean@desy.de, camila.machado@desy.de, jasper.roosmalenepveu@physik.hu-berlin.de

Abstract: We explore the double copy of effective field theories (EFTs), in the recently proposed generalized color-kinematics and Kawai–Lewellen–Tye (KLT) approaches. In the former, we systematically construct scalar numerators satisfying the Jacobi identities from simpler numerator *seeds* with trace-like permutation properties. This construction has the advantage of being easily applicable to any multiplicity, which we exemplify up to 6-point. It employs the linear map between color factors formed by single traces of generators and by products of the structure constants, which also relates the generalized KLT and color-kinematics formalisms, allowing to produce KLT kernels at arbitrary order in the EFT expansion. At 4-point, we show that all EFT kernels are generated and that they only yield double-copy amplitudes which can also be obtained from the traditional KLT kernel. We perform initial checks suggesting that the same conclusions also hold at 5-point. We focus on single-trace massless scalar EFTs which however also control the higher-derivative corrections to gauge and gravity theories.

Contents

1 Introduction

Gravity amplitudes can be obtained from Yang–Mills amplitudes through a squaring operation known as the *double copy*. This relation between theories has been studied through different lenses, exhibiting a rich underlying mathematical structure. New elements of understanding are however still being collected. The double copy was first discovered by Kawai, Lewellen and Tye (KLT) from relations between tree-level open- and closed-string amplitudes [1]. In the KLT formalism, (dilaton-axion-)gravity amplitudes are obtained from products of the appropriate color-ordered on-shell amplitudes of a gauge theory, normalized by a scalar factor dubbed the KLT *kernel*, which depends on the kinematics of the process.

A different but equivalent product structure was discovered at the level of trivalent graphs in field theory by Bern, Carrasco and Johansson (BCJ) [2, 3]. In this approach, each Yang–Mills graph is expressed in terms of a color and a kinematic *numerator* multiplying propagators. It was found that both types of numerators can be chosen such that they obey identical algebraic relations under particle permutations, e.g. Jacobi-like identities. This remarkable fact is known as the color-kinematics (CK) duality. Replacing the color numerator by a second kinematic numerator then yields gravity amplitudes, in an economical way. The double-copy structure was also observed using scattering equations by Cachazo, He and Yuan [4–6], who identified the aforementioned KLT kernel with the inverse matrix of color-ordered amplitudes computed in a bi-adjoint scalar theory.

In recent years, it has become clear that double-copy relations exist beyond the (gravity) = (Yang–Mills)^2 example, and that theories different from Yang–Mills can be used as inputs, thereby double-copying to other theories than gravity. A web of theories connected through double-copy relations was identified and further explorations are ongoing about the space of theories it covers [6, 7]. It was for example found that gauge theories with massive (scalar or fermionic) matter in the fundamental representation can also be double-copied [8–14], and even theories with spontaneous symmetry breaking obey the CK duality [15]. In addition, the double copy of massive gauge bosons has been considered in [16–20]. Besides gauge theories, a double-copy structure has been observed in pure scalar theories such as the aforementioned bi-adjoint scalar theory and the non-linear sigma model [21, 22]. Furthermore, the double-copy product based on the CK duality has been found to extend to the loop level in various examples. For a comprehensive review on the double copy, see [23].

The above examples mostly correspond to renormalizable theories, however higher-dimensional operators are also expected to take part in some form of the double copy. Higher-derivative corrections to Yang–Mills theories were explicitly studied in [24] and more recently in [25], while scalar effective field theories (EFTs) were studied in e.g. [26–31]. Other heavy mass EFTs have also been considered in [32, 33].

Two more systematic approaches to study the range of higher-derivative operators that can be double-copied have recently been proposed by Carrasco, Rodina, Yin, Zekioğlu [34, 35] and Chi, Elvang, Herderschee, Jones, Paranjape [36]. In the former, the color-kinematics approach to the double copy is taken and extended by considering *generalized numerators*

that may simultaneously depend on both color and kinematics, while still satisfying Jacobi-like identities. A set of composition rules is defined to systematically build such numerators at a given order in the EFT expansion. This approach can be extended from 4- to 5-point amplitudes [35].

In the approach taken by [36], the KLT kernel is generalized. The initial observation is that the map between open- and closed-string amplitudes involves a kernel with higher-derivative corrections, which is the inverse matrix of the color-ordered amplitudes in a bi-adjoint EFT with operator coefficients correlated in a specific way [37]. This notion is generalized by allowing for more free parameters in the KLT kernel, arising from a more general bi-adjoint EFT. A set of bootstrap equations on the KLT kernel and input theories are proposed to guarantee a healthy analytical structure in the resulting double-copy amplitudes. Solving the bootstrap then yields a systematic study of the operators possibly involved in the double copy. Correct factorization properties for a theory with fixed particle content need to be imposed as an additional constraint on both the kernel and the input amplitudes.

In this paper, we aim to shed light on these two approaches to the double copy of EFTs and on their relation, by introducing a new method to construct generalized numerators at any multiplicity. The traditional color numerators, consisting of products of Lie-group structure constants, can be written as linear combinations of the single traces of products of group generators. In the same way, we show that all generalized numerators can be constructed from simple *numerator seeds*, which satisfy the same permutation properties as the single traces of generators.

The construction of numerators has previously been investigated from different perspectives. They have for instance been extracted from known amplitudes and the KLT kernel [38, 39]. *Dual trace factors*, analogous to our numerator seeds but involving momenta and polarization vectors, have also previously been identified in Yang–Mills amplitudes [40–44]. In contrast, we use numerators to construct EFT amplitudes, and seeds built out of color and momenta to construct generalized scalar numerators. Based on the kinematic algebra, vector numerators for Yang–Mills and heavy-quark effective theories have also recently been constructed from simpler "pre-numerators" [32, 45].

A further advantage of numerator seeds is that they can be directly related to KLT kernels. This enables the study of the operators involved in the KLT double copy, through a method that is alternative to the bootstrap of [36]. Numerator seeds thus provide further insight into the relation between the double copy approaches of [34, 35] and [36], and into the structure of the generalized KLT kernel. In particular, at 4-point, we show that the double-copy amplitudes obtained with a generalized KLT kernel can equivalently be achieved by the traditional kernel, multiplying healthy local input amplitudes including higher-derivative corrections. As emphasized in [36], the generalized kernel does however allow for more general EFT inputs to the double copy. We also report on various results which indicate that this observation extends to higher multiplicities.

The structure of this paper is as follows. To be self-contained and set the notation, we first provide in Sec. 2 a detailed review of the two aforementioned approaches to the double copy of EFTs. The construction of generalized numerators from seeds at any multiplicity

is then presented in Sec. 3. In this section, we also show how this construction facilitates the reorganization of CK-dual representations of amplitudes in terms of color-ordered amplitudes, which are the building blocks of the KLT formalism. Moreover, for any input amplitude that can be double-copied with a generalized kernel, we identify new objects which yield the same double copy with the traditional kernel. This holds provided the generalized kernel can be constructed from numerator seeds, and provided one can ensure locality of the new objects in order to call them amplitudes. We discuss these two caveats at 4- and 5-point in the subsequent sections. Restricting to 4-point amplitudes, Sec. 4 and Sec. 5 illustrate our method and show that it generates all solutions to the KLT bootstrap. We also analyse the double-copy structure in the KLT formalism, and the factorization properties of the amplitudes involved. The two caveats above are successfully addressed in this 4-point case. Moving on to 5-point amplitudes in Sec. 6, we demonstrate that the lowest-order bootstrap solutions provided in [36], can be reproduced from our numerator construction. We also present partial results suggesting that no new double copies are generated by the generalized kernel at 5-point either.

2 The systematic double copy of EFTs

We start with a review on the generalized KLT method of [36] and the generalized numerator method of [34, 35].

2.1 The generalized KLT approach

The KLT formula for an amplitude with n external particles in the adjoint representation of SU(N) (or U(N)) symmetry groups is given by

$$\mathcal{M}_n = \sum_{\alpha,\beta}^{(n-3)!} A_n^L[\alpha]\, S_n[\alpha|\beta]\, A_n^R[\beta]\,. \tag{2.1}$$

Here, A_n^L and A_n^R are the color-ordered amplitudes of potentially different theories, called *single copies*, and $\mathcal{M}_n$ is the *double-copy* amplitude. Due to the Kleiss–Kuijf (KK) [46] and BCJ [2] relations, the number of independent color-ordered amplitudes forming a *BCJ basis* is $(n-3)!$. The indices α, β in Eq. (2.1) refer to the color-orderings of the single-copy amplitudes and the sums run over the elements of any two BCJ bases, while the KK and BCJ relations ensure that the double-copy amplitude does not depend on the chosen bases.

The multiplication of the single copies is governed by S_n, the KLT kernel, which is a scalar function of Lorentz invariants. Its form depends on the BCJ bases considered in the sum. The kernel plays a crucial role in ensuring that the resulting double-copy amplitude has a healthy analytical structure. It cancels poles that are present in both ordered amplitudes, to prevent double poles in $\mathcal{M}_n$, and provides missing poles so that all physical factorization channels are generated. This requires the KLT kernel to have a precise structure, which was found to be closely related to the amplitudes of the bi-adjoint scalar theory (BAS) [6].

The BAS Lagrangian is given by

$$\mathcal{L}_{\text{BAS}} = -\frac{1}{2}(\partial_\mu \phi^{a\tilde{a}})^2 - \frac{g_\phi}{6} f^{abc} \tilde{f}^{\tilde{a}\tilde{b}\tilde{c}} \phi^{a\tilde{a}} \phi^{b\tilde{b}} \phi^{c\tilde{c}} \,, \tag{2.2}$$

where the scalar field has two adjoint color-group indices.[1] We normalize the adjoint generators and structure constants such that $[T^a, T^b] = f^{abc} T^c$ and $\text{Tr}\big(T^a T^b\big) = \delta^{ab}$. Using the decomposition of the structure constants, the full bi-adjoint n-point amplitude can be written in terms of linearly independent traces of the generators,

$$\mathcal{A}_n^{\text{BAS}} = \sum_{\alpha,\beta \in S_{n-1}} \text{Tr}(T^{a_{\alpha_1}} T^{a_{\alpha_2}} \cdots T^{a_{\alpha_n}})\, m_n[\alpha|\beta]\, \text{Tr}(\tilde{T}^{\tilde{a}_{\beta_1}} \tilde{T}^{\tilde{a}_{\beta_2}} \cdots \tilde{T}^{\tilde{a}_{\beta_n}}) \,. \tag{2.3}$$

The objects $m_n[\alpha|\beta]$ are called doubly color-ordered amplitudes. They can be computed by summing over the trivalent graphs that contribute to the color orderings of both arguments, with appropriate relative signs. For example, at 3-point we have $m_3[123|123] = g_\phi$.

At 4-point, it is useful to define the vector of color factors,

$$c_0 = \Big((1234), (1243), (1324), (1342), (1423), (1432)\Big)^{\mathsf{T}} \,, \tag{2.4}$$

where $(1234) \equiv \text{Tr}(T^{a_1} T^{a_2} T^{a_3} T^{a_4})$, etc. We define $\tilde{c}_0$ similarly as c_0 for the color factors with tilded indices. The full BAS amplitude is then written compactly in matrix form as,[2]

$$\mathcal{A}_4^{\text{BAS}} = c_0 \cdot \mathbf{m}_4 \cdot \tilde{c}_0 \,, \tag{2.5}$$

with

$$\mathbf{m}_4 = \begin{pmatrix} m_4[1234|1234] & m_4[1234|1243] & \cdots & m_4[1234|1432] \\ m_4[1243|1234] & m_4[1243|1243] & \cdots & m_4[1243|1432] \\ \vdots & \vdots & \ddots & \vdots \\ m_4[1432|1234] & m_4[1432|1243] & \cdots & m_4[1432|1432] \end{pmatrix} = g_\phi^2 \begin{pmatrix} \frac{1}{s}+\frac{1}{u} & -\frac{1}{s} & \cdots & \frac{1}{s}+\frac{1}{u} \\ -\frac{1}{s} & \frac{1}{s}+\frac{1}{t} & \cdots & -\frac{1}{s} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{s}+\frac{1}{u} & -\frac{1}{s} & \cdots & \frac{1}{s}+\frac{1}{u} \end{pmatrix} .$$

where we use the conventions $s = s_{12}$, $t = s_{13}$ and $u = s_{14}$ with $s_{ab} = (p_a + p_b)^2$ and all momenta incoming.

Remarkably, the KLT kernel can be identified as the inverse matrix of doubly color-ordered amplitudes [6],

$$S_n[\alpha|\beta] = (m_n[\alpha|\beta])^{-1} \,, \tag{2.6}$$

where the indices α and β should be restricted to any two BCJ bases (in which color-ordered amplitudes are independent), such that the $m_n[\alpha|\beta]$ sub-matrix is of full rank and can be inverted. The rows and columns of $\mathbf{m}_n$ all satisfy the KK and BCJ relations, and the double copies of BAS amplitudes are trivial:

$$m_n[\alpha|\delta] = \sum_{\beta,\gamma}^{(n-3)!} m_n[\alpha|\beta]\, S_n[\beta|\gamma]\, m_n[\gamma|\delta] \,, \tag{2.7}$$

[1] We refer to these indices as color ones although they are associated to global symmetries.

[2] We will often omit vector arrows and transposes, writing for instance $c_0 \cdot \mathbf{m}_4 \cdot \tilde{c}_0$ instead of $\vec{c}_0^{\,\mathsf{T}} \cdot \mathbf{m}_4 \cdot \vec{\tilde{c}}_0$.

where the uncontracted α, δ indices correspond to the color orderings that are left untouched in this relation.

The BAS theory can also be double-copied with another single-copy theory, in which case the KLT product encodes the KK and BCJ relations. This is simple to illustrate at 4-point, where $m_4[\alpha|\beta]$ has rank 1 and the kernel is simply the inverse of a number for fixed α, β. In this case, Eq. (2.6) implies that there are choices of BCJ bases that render the KLT relation trivial, such as in

$$A_4[\alpha] = m_4[\alpha|\beta]\, S_4[\beta|\gamma]\, A_4[\gamma] \quad \text{or} \quad A_4[\alpha] = A_4[\gamma]\, S_4[\gamma|\beta]\, m_4[\beta|\alpha]\,, \tag{2.8}$$

when $\alpha = \gamma$. However, because of the KK and BCJ relations, the left-hand side does not depend on the β, γ bases chosen in the product. For example, (note $\alpha \neq \gamma$, but arbitrary β)

$$A_4[1234] = m_4[1234|\beta]\, S_4[\beta|1243]\, A_4[1243] = \frac{t}{u}\, A_4[1243]\,, \tag{2.9}$$

is exactly a BCJ relation, while one of the KK relations is given by

$$A_4[1234] = m_4[1234|\beta]\, S_4[\beta|1432]\, A_4[1432] = A_4[1432]\,. \tag{2.10}$$

It was emphasized in [36] that the BAS behaves like an identity element in the KLT product, which is why it is also referred to as the *zeroth copy*.

The identification of the BAS as the identity element of the KLT product leads to generalizations of this product associated to modifications of the BAS theory. This is exactly the case for the field theory form of the KLT relation between the open- and closed- string amplitudes [37]. However, not all modifications of the BAS theory result in acceptable KLT kernels. It was found that these corrections should preserve the rank of the matrix of doubly color-ordered amplitudes, which is $(n-3)!$ [36]. This is called the *minimal rank condition*. In this paper, we will focus on higher-derivative (h.d.) corrections suppressed by powers of an EFT cutoff scale Λ, i.e. $\mathbf{m}_n^{\text{h.d.}} = \mathbf{m}_n + \mathcal{O}(1/\Lambda)$. In the decoupling limit, $\Lambda \to \infty$, one therefore recovers the traditional KLT product.

In the generalized KLT formalism, the kernel $S_n^{\text{h.d.}}$ is the inverse of a full-rank submatrix of $\mathbf{m}_n^{\text{h.d.}}$ and satisfies

$$m_n^{\text{h.d.}}[\alpha|\delta] = \sum_{\beta,\gamma} m_n^{\text{h.d.}}[\alpha|\beta]\, S_n^{\text{h.d.}}[\beta|\gamma]\, m_n^{\text{h.d.}}[\gamma|\delta]\,. \tag{2.11}$$

The consistency conditions that follow from products of the identity element (BAS) with another theory are called the generalized KKBCJ relations. For single-copy color-ordered amplitudes, $A'_{n,\text{L/R}}$, they have the form

$$\begin{aligned} A'_{n,\text{R}}[\alpha] &= \sum_{\beta,\gamma} m_n^{\text{h.d.}}[\alpha|\beta]\, S_n^{\text{h.d.}}[\beta|\gamma]\, A'_{n,\text{R}}[\gamma]\,, \\ A'_{n,\text{L}}[\alpha] &= \sum_{\beta,\gamma} A'_{n,\text{L}}[\gamma]\, S_n^{\text{h.d.}}[\gamma|\beta]\, m_n^{\text{h.d.}}[\beta|\alpha]\,, \end{aligned} \tag{2.12}$$

and allow to bootstrap the single-copy amplitudes $A'_{n,\mathrm{L/R}}$ that can take part in the double copy,

$$\mathcal{M}'_n = \sum_{\alpha,\beta} A'_{n,\mathrm{L}}[\alpha]\; S_n^{\mathrm{h.d.}}[\alpha|\beta]\; A'_{n,\mathrm{R}}[\beta]\,. \tag{2.13}$$

Depending on the form of $m_n^{\mathrm{h.d.}}$ and $S_n^{\mathrm{h.d.}}$, the generalized KKBCJ relations for $A'_{n,\mathrm{R}}$ and $A'_{n,\mathrm{L}}$ may be different. We emphasize that both $A_{n,\mathrm{L/R}}$ and $A'_{n,\mathrm{L/R}}$ may in principle contain higher-derivative corrections. The prime indicates that the amplitudes satisfy generalized KKBCJ relations, which allow for more operators.

That Eq. (2.13) produces a healthy amplitude when $\mathbf{m}_n^{\mathrm{h.d.}}$ has minimal rank is a non-trivial empirical result [36]. The generalized KLT formalism allows for a systematic study of the space of theories that can appear as input and output of the double-copy procedure. Although the bootstrap equations strongly constrain the higher-derivative corrections that are allowed in the input (single-copy) amplitudes, the generalized KLT formalism allows for more independent operators in the single copies than its traditional version. However, up to the orders checked explicitly in [36], it was found that the space of generalized output (double-copy) amplitudes $\mathcal{M}'_n$ is the same as $\mathcal{M}_n$. At 4-point, where the double-copy relation contains a single term, 'similarity transformations' were proposed in [36] to explain this fact (see also [47]). We aim to shed further light on this observation in Sec. 3 and Sec. 5.

As an example, the 4-point amplitude of the BAS theory with higher-derivative corrections is

$$\mathcal{A}_4^{\mathrm{BAS+h.d.}} = c_0 \cdot \mathbf{m}_4^{\mathrm{h.d.}} \cdot \tilde{c}_0\,. \tag{2.14}$$

Solving the minimal rank condition, the $\mathbf{m}_4^{\mathrm{h.d.}}$ matrix of doubly color-ordered amplitudes corresponding to BAS+h.d. can be written as [36]

$$\mathbf{m}_4^{\mathrm{h.d.}} = \begin{pmatrix} f_1(s,t) & f_2(s,t) & f_2(u,t) & f_2(s,t) & f_2(u,t) & f_1(s,t) \\ f_2(s,u) & f_1(s,u) & f_2(t,u) & f_1(s,u) & f_2(t,u) & f_2(s,u) \\ f_2(u,s) & f_2(t,s) & f_1(t,s) & f_2(t,s) & f_1(t,s) & f_2(u,s) \\ f_2(s,u) & f_1(t,u) & f_2(t,u) & f_1(t,u) & f_2(t,u) & f_2(s,u) \\ f_2(u,s) & f_2(t,s) & f_1(u,s) & f_2(t,s) & f_1(u,s) & f_2(u,s) \\ f_1(s,t) & f_2(s,t) & f_2(u,t) & f_2(s,t) & f_2(u,t) & f_1(s,t) \end{pmatrix}, \tag{2.15}$$

where

$$f_1(s,t) = f_1(u,t) \equiv \frac{f_2(s,t)f_2(u,s)}{f_2(t,s)}\,, \tag{2.16}$$

and f_2 satisfies the bootstrap equation

$$f_2(s,t)f_2(t,u)f_2(u,s) = f_2(t,s)f_2(s,u)f_2(u,t)\,. \tag{2.17}$$

Furthermore, the aforementioned assumption $\mathbf{m}_4^{\mathrm{h.d.}} = \mathbf{m}_4 + \mathcal{O}(g_\phi^2/\Lambda^2)$ requires $f_2(s,t) = -g_\phi^2/s + \mathcal{O}(g_\phi^2/\Lambda^2)$. Additional constraints on f_2 arise if we forbid extra particles in the BAS+h.d. theory.

For later reference, we note that the alternative ordering of the single traces, $c_0 = \big((1324),(1234),(1243),(1423),(1432),(1342)\big)^{\mathsf{T}}$, exposes a block matrix structure,

$$\mathbf{m}_4^{\text{h.d.}} = \begin{pmatrix} \tilde{\mathbf{m}}_4^{\text{h.d.}} & \tilde{\mathbf{m}}_4^{\text{h.d.}} \\ \tilde{\mathbf{m}}_4^{\text{h.d.}} & \tilde{\mathbf{m}}_4^{\text{h.d.}} \end{pmatrix}, \quad \text{with} \quad \tilde{\mathbf{m}}_4^{\text{h.d.}} \equiv \begin{pmatrix} f_1(t,s) & f_2(u,s) & f_2(t,s) \\ f_2(u,t) & f_1(s,t) & f_2(s,t) \\ f_2(t,u) & f_2(s,u) & f_1(s,u) \end{pmatrix}. \tag{2.18}$$

This structure also exists if we turn off the higher-derivative corrections, with

$$\tilde{\mathbf{m}}_4 = g_\phi^2 \begin{pmatrix} \frac{1}{t}+\frac{1}{u} & -\frac{1}{u} & -\frac{1}{t} \\ -\frac{1}{u} & \frac{1}{s}+\frac{1}{u} & -\frac{1}{s} \\ -\frac{1}{t} & -\frac{1}{s} & \frac{1}{s}+\frac{1}{t} \end{pmatrix}. \tag{2.19}$$

Restricting to the smaller $\tilde{\mathbf{m}}_4^{\text{h.d.}}$ and $\tilde{\mathbf{m}}_4$ matrices will prove useful in the following sections. We will actually only use these hereafter and drop the tildes for convenience. However, such a block matrix structure generally only exists at lowest derivative order for $n > 4$ particles. The 4-point case is special because the kinematics is invariant under reversal of the particle labels: $\tilde{f}(1,2,3,4) \equiv f(s_{12}, s_{13}) = f(s_{43}, s_{42}) \equiv \tilde{f}(4,3,2,1)$, for any function f of the Mandelstam invariants.

2.2 The generalized numerators approach

Another approach to the double copy is based on the color-kinematics (CK) duality [2, 3]. The basic idea is to use the decomposition of an on-shell n-point amplitude $\mathcal{A}_n$ on trivalent graphs g,

$$\mathcal{A}_n = \sum_g \frac{c_g n_g}{d_g}, \tag{2.20}$$

where d_g is the product of the (inverse) propagators it involves; c_g traditionally correspond to color factors associated to that same graph (i.e. combinations of generators of the gauge algebra); while n_g are the kinematic numerators that depend on Lorentz invariants and possibly on polarization vectors. Given an amplitude $\mathcal{A}_n$, the numerators n_g are not unique. CK duality is then a property of amplitudes for which there exists a choice of numerators n_g which verify the same algebraic relations as those of the color factors c_g, inherited from the gauge algebra. In certain theories, such as Yang–Mills, all tree-level amplitudes satisfy the CK duality. In particular, the SU(N) color factors of 4-point amplitudes obey Jacobi identities of the form[3]

$$c_{g_a} + c_{g_b} + c_{g_c} = 0, \tag{2.21}$$

and antisymmetry relations upon interchanging two legs on one vertex in the g_a, g_b, g_c graphs. Any numerator satisfying these *adjoint algebraic relations* will be called an *adjoint numerator*.

In gauge theories, the color relations ensure the gauge invariance of the amplitude under shifts of the polarization vectors contained in n_g, $\mathcal{A}|_{\epsilon_i \to p_i} = 0$, for any particle label

[3] The signs in the Jacobi relation are fixed by the c_s, c_t, c_u conventions. We use $c_s = f^{abe} f^{ecd}$, $c_t = f^{ace} f^{edb}$ and $c_u = f^{ade} f^{ebc}$.

i. This implies, in turn, that c_g can be replaced by any expression which satisfies the same relations without spoiling gauge invariance, and the latter applies in particular to another copy of a CK-dual n_g. The BCJ double-copy procedure [2, 3] is thus schematically

$$\mathcal{A}_n = \sum_g \frac{c_g n_g}{d_g}\,, \quad \tilde{\mathcal{A}}_n = \sum_g \frac{c_g \tilde{n}_g}{d_g} \quad \longrightarrow \quad \mathcal{M}_n = \sum_g \frac{\tilde{n}_g n_g}{d_g}\,, \tag{2.22}$$

where $\mathcal{M}_n$ is the double-copy amplitude. The numerator $\tilde{n}_g$ is not necessarily the same as n_g, but both are CK-dual to the same c_g color factors.[4] For two copies of Yang–Mills, the resulting amplitude describes the scattering of gravitons (as well as scalars).

Let us illustrate this approach at 4-point with the well-known case of Yang–Mills theories. There are three trivalent diagrams associated to the s, t, u channels, and the amplitude reads

$$\mathcal{A}_4^{\rm YM}(1_a^+ 2_b^+ 3_c^- 4_d^-) = g_{\rm YM}^2\, [12]^2\, \langle 34\rangle^2 \left(\frac{f^{ace} f^{edb}}{st} - \frac{f^{ade} f^{ebc}}{su} \right) = \frac{c_s n_s}{s} + \frac{c_t n_t}{t} + \frac{c_u n_u}{u}\,, \tag{2.23}$$

where we defined

$$c_{\rm adj} \equiv \begin{pmatrix} c_s \\ c_t \\ c_u \end{pmatrix} \equiv \begin{pmatrix} f^{abe} f^{ecd} \\ f^{ace} f^{edb} \\ f^{ade} f^{ebc} \end{pmatrix}\,, \quad n \equiv \begin{pmatrix} n_s \\ n_t \\ n_u \end{pmatrix} \equiv g_{\rm YM}^2\, [12]^2\, \langle 34\rangle^2 \begin{pmatrix} \frac{1-2\alpha}{u} - \frac{\alpha}{t} \\ \frac{\alpha}{s} - \frac{1-2\alpha}{u} \\ \frac{\alpha}{t} - \frac{\alpha}{s} \end{pmatrix}\,, \tag{2.24}$$

with an arbitrary function α. Square and angle brackets denote spinor-helicity variables (for pedagogical introductions see e.g. [48, 49]). The Jacobi identity reads $c_s + c_t + c_u = 0$ and the explicit formulæ above allow to check that $n_s + n_t + n_u = 0$ for any α. The color factors also satisfy antisymmetry relations such as $c_s|_{1\leftrightarrow 2} = c_s|_{a\leftrightarrow b} = -c_s$. For $\alpha = 1/3$, the numerators are antisymmetric too: e.g. $n_s|_{1\leftrightarrow 2} = n_s|_{t\leftrightarrow u} = -n_s$. This shows the CK duality of the 4-point YM amplitude. The double-copy method then leads to a diffeomorphism-invariant four-gravitons amplitude,

$$\mathcal{M}_4^{\rm GR}(1^{+2} 2^{+2} 3^{-2} 4^{-2}) = \frac{n_s^2}{s} + \frac{n_t^2}{t} + \frac{n_u^2}{u}\,. \tag{2.25}$$

The bi-adjoint scalar theory considered in the previous subsection also plays the role of a zeroth copy in the CK approach. It has a (color-color) dual structure,

$$\mathcal{A}_n^{\rm BAS} = g_\phi^{n-2} \sum_g \frac{c_g \tilde{c}_g}{d_g}\,, \tag{2.26}$$

where the two color groups lead to c_g and $\tilde{c}_g$. Being both color factors in the adjoint representation, they verify Jacobi and antisymmetry relations and any of the two can be treated as a CK-dual numerator n_g. Moreover, replacing c_g in Eq. (2.22) by one of the BAS color factors has a trivial effect.

[4] The CK duality has to be manifest in at least one numerator. Notice, however, that a manifest CK duality can be achieved through so-called generalized gauge transformations, which are not always trivial to perform.

Higher-derivative effects in the BAS theory can be included, while preserving the dual CK structure, by building generalized numerators, $c^{\text{h.d.}}(c, s_{ab})$ [34, 35] (see also [31]). These verify the same adjoint algebraic relations as the c_g, but may depend on both the Mandelstam invariants s_{ab} and color factors c. The color factors are themselves not necessarily of adjoint type, but more generally built from products of traces of group generators. Only their combination with Mandelstams is required to satisfy the adjoint algebraic relations. In this paper, we focus on single traces and do for instance not consider generalizations in the form of double traces (see e.g. [31]).

Choosing $c^{\text{h.d.}} = c_g + \mathcal{O}(s_{ab}/\Lambda^2)$, the generalized numerators result in EFT amplitudes for the bi-adjoint scalar theory that retains the CK-duality,

$$\mathcal{A}_n^{\text{BAS+h.d.}} = g_\phi^{n-2} \sum_g \frac{c_g^{\text{h.d.}}\, \tilde{c}_g^{\text{h.d.}}}{d_g} \ . \tag{2.27}$$

Similarly, higher-order corrections in a gauge theory that preserve the CK-duality can be obtained by replacing the color factors by generalized numerators,

$$\mathcal{A}_n^{\text{YM+h.d.}} = \sum_g \frac{c_g^{\text{h.d.}} n_g}{d_g} \ , \tag{2.28}$$

for unmodified n_g. This can also be interpreted as the double copy between $\mathcal{A}_n^{\text{YM}}$ and $\mathcal{A}_n^{\text{BAS+h.d.}}$ amplitudes. The gauge invariance of the amplitude is maintained because the $c_g^{\text{h.d.}}$ satisfy the same relations as the c_g. This is the approach taken in [34, 35] to extend the CK double-copy method to higher-derivative EFT amplitudes.

To find the most general allowed numerators, one could construct an ansatz for $c^{\text{h.d.}}$ at each order in the Mandelstams and impose the adjoint algebraic relations. Exploiting the underlying structure instead, [34] demonstrated that all purely kinematic adjoint 4-point numerators can be built using a simple composition rule acting on existing lower-order adjoint kinematic numerators j and k (vectors),

$$n(j,\, k) = \begin{pmatrix} j_t k_t - j_u k_u \\ j_u k_u - j_s k_s \\ j_s k_s - j_t k_t \end{pmatrix} \ . \tag{2.29}$$

This requires only one building block made out of kinematic invariants,

$$n^{(ss)} = \begin{pmatrix} t - u \\ u - s \\ s - t \end{pmatrix} \ . \tag{2.30}$$

Furthermore, all adjoint 4-point numerators involving one factor of color can be generated using Eq. (2.29) with one of the two input numerators containing color, and the additional rule

$$c\,(j,\, d) = d\, j \ , \tag{2.31}$$

where d is the color factor that is fully symmetric under external particle permutations,

$$d^{(abcd)} \equiv \frac{1}{3!} \sum_{\sigma \in S_3} \mathrm{Tr}(T^a T^{\sigma(b)} T^{\sigma(c)} T^{\sigma(d)}) \,. \tag{2.32}$$

The only other color structures required as primary building blocks are the adjoint ones, $c_s = f^{abe} f^{ecd}$, etc. encountered in Eq. (2.24). All possible adjoint structures at 4-point can then be obtained by successive applications of these composition rules, and linear combinations of these yield the most general $c^{\text{h.d.}}$ [34]. At 5-point, the situation is complicated by the presence of more composition rules and algebraic structures [35].

In the next section, instead of constructing adjoint numerators from lower-order adjoint numerators and applying composition rules, we will build all of them from simpler *non-adjoint* objects. This alternative construction procedure can be extended to higher multiplicities without much complication.

3 Numerator construction from seeds at any multiplicity

We now propose an alternative method to construct generalized adjoint numerators of the SU(N) (or U(N)) unitary group. At this stage, we are only interested in the adjoint numerators themselves, without regard to the factorization properties and particle content of amplitudes in which they enter. Section 5 discusses which extra constraints are imposed by such considerations. The current section applies to any number of particles, while the construction is repeated explicitly in the next section at 4-point.

3.1 Numerator seeds

The main observation is that the adjoint color factor, c_{adj}, consisting of products of the structure constants, can be written in terms of linear combinations of single traces of the group generators (see also [44]),

$$c_{\text{adj}} = \mathbf{J} \cdot c_0 \,, \tag{3.1}$$

where

$$c_0 = \big((123...n), \quad \text{all permutations of } \{2, 3, ..., n\}\big)^{\mathsf{T}} \,, \tag{3.2}$$

The matrix $\mathbf{J}$ contains only $\{\pm 1, 0\}$ entries and will play a central role in our construction. Its entries are determined by decomposing the structure constants in terms of traces through

$$f^{abc} = (abc) - (acb) \,. \tag{3.3}$$

Products of traces can then be combined using the SU(N) completeness relation,

$$\sum_a T^a_{ij} T^a_{kl} = \delta_{il}\delta_{kj} - \frac{1}{N}\delta_{ij}\delta_{kl} \,, \tag{3.4}$$

and one can show that the $1/N$ terms cancel for products of structure constants.

In this way, the matrix $\mathbf{J}$ relates the simple algebraic structure of single traces (contained in the c_0 vector) to the more involved adjoint algebraic properties. It encodes

the Jacobi identities, and the antisymmetry relations which follow from the permutation properties of the single traces. The matrix $\mathbf{J}$ can be decomposed as follows:

$$\mathbf{J} = \mathbf{A} \cdot \mathbf{B} \tag{3.5}$$

where $\mathbf{B}$ is the $(n-2)! \times (n-1)!$ matrix of rank $(n-2)!$ that relates c_0 to the color factors in a Del Duca–Dixon–Maltoni (DDM) basis [50], and $\mathbf{A}$ is the $(2n-5)!! \times (n-2)!$ matrix of rank $(n-2)!$ that relates the DDM basis to c_{adj}. Conventions can also be chosen such that both $\mathbf{A}$ and $\mathbf{B}$ are sub-matrices of $\mathbf{J}$, see [44].

Therefore, any object n_0 that satisfies the same algebraic properties as the single traces will be mapped to an adjoint numerator under multiplication by $\mathbf{J}$. The vector of single traces has linearly independent entries that are given in terms of permutations of one functional form, which is cyclically invariant in its arguments,

$$(ab...c) = (b...ca)\,. \tag{3.6}$$

In matrix notation, the algebraic properties can be summarized as follows. Under a relabeling σ of the particles, c_0 transforms as

$$c_0 \underset{\sigma}{\rightarrow} M_{c_0,\sigma} \cdot c_0\,, \tag{3.7}$$

where $M_{c_0,\sigma}$ is a permutation matrix. We shall call objects that obey the algebraic properties of the single traces,

$$n_0 \underset{\sigma}{\rightarrow} M_{c_0,\sigma} \cdot n_0\,, \tag{3.8}$$

numerator seeds, or *seeds* for short, as they can be used to generate adjoint numerators. We shall use both n_0 and $c_0^{\text{h.d.}}$ to refer to numerator seeds. The latter notation emphasizes that the cyclically invariant functional form of the seeds depends on both Mandelstam invariants and color factors. Such seeds generate the generalized adjoint numerators discussed in Sec. 2.2.

In App. A, we prove that *any* adjoint numerator can be constructed from a numerator seed, i.e.

$$n_{\text{adj}} = \mathbf{J} \cdot n_0\,. \tag{3.9}$$

Since the seeds are straightforwardly constructed, this provides an efficient way to explore the space of possible (generalized) adjoint numerators. A similar result was proven for kinematic numerators in renormalizable Yang–Mills theory using a different method [42].

A set of linearly independent numerator seeds generally maps to a redundant set of adjoint numerators, namely $\mathbf{J} \cdot n_0 = \mathbf{J} \cdot n_0'$ could happen even for $n_0 \neq n_0'$. Therefore, identifying a set of seeds which generates independent adjoint numerators requires an extra step of reduction. This can be done by directly inspecting the general expression of n_0, or by relying on a construction which removes redundancies. We give explicit examples of the former below, while the latter can be achieved using the Moore–Penrose pseudo-inverse of $\mathbf{J}$, called $\mathbf{J}^+$ (see App. A): it also follows from the argument of App. A that $\mathbf{J}^+ \cdot \mathbf{J} \cdot n_0 = \mathbf{J}^+ \cdot \mathbf{J} \cdot n_0'$ is a valid numerator seed.[5] A complete and independent set of

k	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
4-pt	1	1	1	2	2	2	3	3	3	4	4	4	5	5	5
5-pt	0	0	1	2	5	8	14	21	32	45	63	84	112	144	185
6-pt	1	3	9	23	54	120	243	469	861	1509	2546	4158	$\cdots$	$\cdots$	$\cdots$

Table 1. Number of independent scalar kinematic numerators at $\mathcal{O}(1/\Lambda^{2k})$ in the EFT expansion. The Gram determinant constraints relevant in 4 spacetime dimensions have been accounted for. The counting at 4- and 5-point was achieved in [34] while the 6-point one is provided here for the first time. Although straightforward in principle, the numerically intensive reduction of the overcomplete set of numerators was not pushed beyond $k = 12$ at 6-point.

numerator seeds can thus be obtained by projecting with $\mathbf{J}^+ \cdot \mathbf{J}$ on all cyclically invariant functions.

As an example of redundancies, permutation invariant functional forms result in valid numerator seeds, but they are mapped to zero and thus do not give rise to independent adjoint numerators. In addition, $\mathbf{J}$ always combines seed entries with reversed ordering of particle labels, since $c_{\text{adj}} \to (-1)^n c_{\text{adj}}$ under reversal, at n-point. The entries in a numerator seed can thus be ordered such that $\mathbf{J}$ has a block matrix structure, schematically: $\mathbf{J}_{a\times b} = \big(\mathbf{J}_{a\times(b/2)}\,,\ (-1)^n \mathbf{J}_{a\times(b/2)}\big)$. Therefore, the general numerator seed

$$n_0 = \Big(n_0(1,2,...,n),\ \ldots\ , n_0(n,...,2,1),\ \ldots\ \Big)\,, \tag{3.10}$$

and the seed on which we impose (anti)symmetry under reversal on the functional form,

$$\begin{aligned}\bar{n}_0 = \frac{1}{2}\Big(&n_0(1,2,...,n) + (-1)^n\, n_0(n,...,2,1),\ \ldots\ ,\\ &n_0(n,...,2,1) + (-1)^n\, n_0(1,2,...,n),\ \ldots\ \Big)\,,\end{aligned} \tag{3.11}$$

result in the same adjoint numerator. At 4-point, these are the only sources of redundancy in the construction of adjoint numerators. There are further redundancies in the construction of CK-dual amplitudes, called generalized gauge transformations that will be addressed in section Sec. 4.3. At higher multiplicity, redundancies can take a more complicated form, to be exemplified at 5-point in Sec. 6.

Even without identifying the specific algebraic origin of the redundancies, it is straightforward to just build the overcomplete set of seeds and identify numerically a basis of independent adjoint numerators. We provide the counting of the latter in Table 1, for numerators taking the form of polynomials of Mandelstam invariants. This table can be compared with Table 2 in App. B, listing the number of independent seeds, which grows faster with n than the number of independent numerators. The construction of adjoint numerators at 4-point agrees with the observation made in [34], namely that higher-order

[5]It is also interesting to note that the projection by $\mathbf{J}^+ \cdot \mathbf{J}$ is equivalent to imposing the KK relations on the seeds. As shown in [44], a vector $\vec{v}$ satisfies the KK relations if and only if $\vec{r}_i \cdot \vec{v} = 0$, where r_i are the right null-vectors of $\mathbf{J}$. Since the right null-space of $\mathbf{J}$ is captured by $\mathbb{1} - \mathbf{J}^+ \cdot \mathbf{J}$, and $(\mathbb{1} - \mathbf{J}^+ \cdot \mathbf{J}) \cdot \mathbf{J}^+ \cdot \mathbf{J} \cdot n_0 = 0$, the seed $\mathbf{J}^+ \cdot \mathbf{J} \cdot n_0$ satisfies the KK relations.

adjoint numerators can be obtained from lower order ones by multiplication with a permutation invariant function. The results at 5-point also agree with the number of independent numerators listed in Table 2 of [35] and we have constructed all the adjoint kinematic scalar numerators up to 6-point and $\mathcal{O}(1/\Lambda^{24})$.[6] We also provide explicit examples of numerator seeds and adjoint numerators of lowest orders in App. B.

3.2 Double copy and color-ordered amplitudes from numerators

The matrix $\mathbf{J}$ is also useful to obtain color-ordered amplitudes from adjoint numerators. This approach was previously taken in [7], with a similar definition for $\mathbf{J}$.[7] For instance, writing the full bi-adjoint scalar amplitude of Eq. (2.26) in matrix form and using the fact that c_{adj} and c_0 are related by $\mathbf{J}$ through Eq. (3.1) yields

$$\mathcal{A}^{\text{BAS}} = c_{\text{adj}} \cdot \mathbf{P} \cdot \tilde{c}_{\text{adj}} = c_0 \cdot \mathbf{J}^{\mathsf{T}} \cdot \mathbf{P} \cdot \mathbf{J} \cdot \tilde{c}_0 \,, \tag{3.12}$$

where $\mathbf{P}$ contains the propagators (and coupling constants) of the trivalent graphs on its diagonal. By definition, this produces the BAS matrix of doubly ordered amplitudes (see also [44]),

$$\mathbf{m} = \mathbf{J}^{\mathsf{T}} \cdot \mathbf{P} \cdot \mathbf{J} \,. \tag{3.13}$$

Similarly, one can write the single-copy (color-ordered) amplitudes in terms of a numerator seed (replacing $\tilde{c}_{\text{adj}}$ by $n_{\text{adj}} = \mathbf{J} \cdot n_{0,\text{R}}$),

$$A_{\text{R}} = \mathbf{m} \cdot n_{0,\text{R}} \,. \tag{3.14}$$

These single-copy amplitudes satisfy the traditional KK and BCJ relations as a consequence of the explicit factor of $\mathbf{m}$ and of the relation in Eq. (2.7). The double-copy amplitude can be obtained in the CK way,

$$\mathcal{M} = n_{\text{adj,L}} \cdot \mathbf{P} \cdot n_{\text{adj,R}} = n_{0,\text{L}} \cdot \mathbf{m} \cdot n_{0,\text{R}} \,, \tag{3.15}$$

or equivalently through the KLT relations,

$$\begin{aligned} \mathcal{M} &= \sum_{\alpha,\beta}^{(n-3)!} A_{\text{L}}[\alpha]\, S[\alpha|\beta]\, A_{\text{R}}[\beta] \\ &= \sum_{\alpha,\delta}^{(n-1)!} \sum_{\beta,\gamma}^{(n-3)!} n_{0,\text{L}}[\alpha]\, m[\alpha|\beta]\, S[\beta|\gamma]\, m[\gamma|\delta]\, n_{0,\text{R}}[\delta] \\ &= n_{0,\text{L}} \cdot \mathbf{m} \cdot n_{0,\text{R}} \,. \end{aligned} \tag{3.16}$$

This exposes the special role played by the BAS matrix of color-ordered amplitudes to ensure the correct propagator structure of the double-copy amplitude.

[6] We thank the authors of [51] for private communications which lead us to correct our enumeration of 6-point numerators.

[7] For us, $\mathbf{J}$ relates the adjoint color factors to the single traces while, in [7], they are instead related to a subset of adjoint color factors (which are independent under the Jacobi identities).

The same method can be applied to obtain color-ordered amplitudes from generalized numerators. Defining a matrix $\mathbf{H}^{\text{h.d.}}$, which depends only on Lorentz invariants, one can decompose the numerator seeds (which, for simplicity, we build using only single traces) as $c_0^{\text{h.d.}} = \mathbf{H}^{\text{h.d.}} \cdot c_0$.[8] This allows Eq. (2.27) to be rewritten as

$$\mathcal{A}^{\text{BAS+h.d.}} = c_0^{\text{h.d.}} \cdot \mathbf{m} \cdot \tilde{c}_0^{\text{h.d.}} = c_0 \cdot \mathbf{H}_{\text{R}}^{\text{h.d.}} \cdot \mathbf{m} \cdot \mathbf{H}_{\text{L}}^{\text{h.d.}} \cdot \tilde{c}_0 \,. \tag{3.17}$$

It follows that the higher-derivative color-ordered amplitudes can be constructed by left- and right-multiplication of the lowest order matrix $\mathbf{m}$,

$$\mathbf{m}^{\text{h.d.}} = \mathbf{H}_{\text{R}}^{\text{h.d.}} \cdot \mathbf{m} \cdot \mathbf{H}_{\text{L}}^{\text{h.d.}} \,. \tag{3.18}$$

Similarly, starting from Eq. (2.28), one can write the higher-derivative single-copy (full) amplitude as

$$\mathcal{A}'_{\text{R}} = c_0^{\text{h.d.}} \cdot \mathbf{J}^{\mathsf{T}} \cdot \mathbf{P} \cdot \mathbf{J} \cdot n_{0,\text{R}} \,, \tag{3.19}$$

and color-ordered amplitudes as

$$A'_{\text{R}} = \mathbf{H}_{\text{R}}^{\text{h.d.}} \cdot \mathbf{J}^{\mathsf{T}} \cdot \mathbf{P} \cdot \mathbf{J} \cdot n_{0,\text{R}} \,, \tag{3.20}$$

where we defined $n_{0,\text{R}}$ such that $n_{\text{adj}} = \mathbf{J} \cdot n_{0,\text{R}}$. Assuming that $\mathbf{H}_{\text{L}}^{\text{h.d.}}$ is of full rank, which always holds for an EFT expansion of the form $\mathbf{H}_{\text{L}}^{\text{h.d.}} = \mathbb{1} + \mathcal{O}(s_{ab}/\Lambda^2)$, we can further define $n'_{0,\text{R}}$ such that $n_{0,\text{R}} \equiv \mathbf{H}_{\text{L}}^{\text{h.d.}} \cdot n'_{0,\text{R}}$, and we obtain

$$A'_{\text{R}} = \mathbf{H}_{\text{R}}^{\text{h.d.}} \cdot \mathbf{m} \cdot \mathbf{H}_{\text{L}}^{\text{h.d.}} \cdot n'_{0,\text{R}} = \mathbf{m}^{\text{h.d.}} \cdot n'_{0,\text{R}} \,. \tag{3.21}$$

Note that $n'_{0,\text{R}}$ also satisfies the properties of numerator seeds, and A'_{R} satisfies the generalized KKBCJ relations.

The naive procedure for obtaining a double-copy amplitude in the generalized CK formalism is to replace $c_0^{\text{h.d.}}$ in Eq. (3.19) by a kinematic seed $n_{0,\text{L}}$ (or equivalently replace $c_{\text{adj}}^{\text{h.d.}} = \mathbf{J} \cdot c_0^{\text{h.d.}}$ by $n_{\text{adj,L}}$). This results in the same double-copy amplitude $\mathcal{M}$ as one could have been obtained without generalized numerators (Eq. (3.15)). The same occurs in the generalized KLT formalism:

$$\begin{aligned}
\mathcal{M}' &= \sum_{\alpha,\beta}^{(n-3)!} A'_{\text{L}}[\alpha]\, S^{\text{h.d.}}[\alpha|\beta]\, A'_{\text{R}}[\beta] \\
&= \sum_{\alpha,\delta}^{(n-1)!} \sum_{\beta,\gamma}^{(n-3)!} n'_{0,\text{R}}[\alpha]\, m^{\text{h.d.}}[\alpha|\beta]\, S^{\text{h.d.}}[\beta|\gamma]\, m^{\text{h.d.}}[\gamma|\delta]\, n'_{0,\text{R}}[\delta] \\
&= n'_{0,\text{L}} \cdot \mathbf{m}^{\text{h.d.}} \cdot n'_{0,\text{R}} = n_{0,\text{L}} \cdot \mathbf{m} \cdot n_{0,\text{R}} = \mathcal{M} \,,
\end{aligned} \tag{3.22}$$

where we distinguish sums over all $(n-1)!$ color factors and sums over BCJ bases.

We have thus derived that the double-copy amplitudes obtained by the generalized KLT relations can equivalently be obtained through the traditional KLT double copy. However,

[8] At 4-point, the form of the matrix $\mathbf{H}^{\text{h.d.}}$ will be derived in Sec. 4, where it will be shown to be diagonal. This is not generally true at higher multiplicity.

there are two caveats to this statement. First, we derived this statement assuming that $\mathbf{m}^{\text{h.d.}}$ is constructed via generalized numerators. It is unclear whether it then reproduces all solutions to the KLT bootstrap. In Sec. 5, we prove that this is the case at 4-point, and we have performed initial checks at 5-point presented in Sec. 6. Second, while the double-copy amplitude obtained by $A'_{\text{L/R}}$ and a generalized kernel is the same as the one obtained by $A_{\text{L/R}} = (\mathbf{H}^{\text{h.d.}}_{\text{L/R}})^{-1} \cdot A'_{\text{L/R}}$ and a traditional kernel, it is unclear whether $A_{\text{L/R}}$ are physical amplitudes and what is their particle content. In Sec. 5, we show at 4-point that the assumption that $\mathbf{H}^{\text{h.d.}}_{\text{L/R}}$ does not affect the BAS particle content implies that this is also the case for $A_{\text{L/R}}$.

4 Seeds and generalized numerators at 4-point

In this section, we will work out the 4-point construction of scalar adjoint numerators from their seeds. This serves as an illustration of the method, and prepares for a comparison with the generalized KLT formalism in the next section.

A product of structure constants can be written in terms of single traces as follows,

$$f^{abx} f^{xcd} = \big[(abcd) + (dcba)\big] - \big[(abdc) + (cdba)\big]. \tag{4.1}$$

As noted before, the traces appear together with their reversed ordering in this relation. The adjoint color numerator can thus be written compactly in terms of traces as

$$c_{\text{adj}} = \begin{pmatrix} f^{12x} f^{x34} \\ f^{13x} f^{x42} \\ f^{14x} f^{x23} \end{pmatrix} = \begin{pmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix} \cdot \begin{pmatrix} (1324) + (4231) \\ (1234) + (4321) \\ (1243) + (3421) \end{pmatrix} = \mathbf{J}_4 \cdot c_0 \,. \tag{4.2}$$

This defines the matrix $\mathbf{J}_4$ with $\{\pm 1, 0\}$ entries, which encodes the Jacobi identity as the vanishing sum of its rows. Notice that we redefined the vector of traces c_0 shown in Eq. (2.4), combining traces and their reversed orderings such that

$$c_0 = \big(c_0(1,3,2,4),\ c_0(1,2,3,4),\ c_0(1,2,4,3)\big)^{\mathsf{T}}, \quad c_0(a,b,c,d) = (abcd) + (dcba)\,. \tag{4.3}$$

We have conventionally chosen the ordering in the arguments of c_0 entries such that the s, t, u Mandelstams are invariant under the cyclic permutation of the (1,3,2,4), (1,2,3,4), (1,2,4,3), respectively. For instance, $s|_{1\to 3\to 2\to 4} = s$.

The 4-point BAS amplitude can now be rewritten as

$$\begin{aligned} \mathcal{A}_4^{\text{BAS}} &= g_\phi^2 \left(\frac{c_s \tilde{c}_s}{s} + \frac{c_t \tilde{c}_t}{t} + \frac{c_u \tilde{c}_u}{u} \right) = c_{\text{adj}} \cdot \mathbf{P}_4 \cdot \tilde{c}_{\text{adj}} \\ &= c_0 \cdot \mathbf{J}_4^{\mathsf{T}} \cdot \mathbf{P}_4 \cdot \mathbf{J}_4 \cdot \tilde{c}_0 = c_0 \cdot \mathbf{m}_4 \cdot \tilde{c}_0 \,, \end{aligned} \tag{4.4}$$

with

$$\mathbf{P}_4 \equiv g_\phi^2 \begin{pmatrix} \frac{1}{s} & 0 & 0 \\ 0 & \frac{1}{t} & 0 \\ 0 & 0 & \frac{1}{u} \end{pmatrix}, \tag{4.5}$$

and $\tilde{\mathbf{m}}_4$ defined in Eq. (2.19).

It is clear that $\mathbf{J}_4$ multiplying any numerator seed that satisfies the same algebraic relations[9] as c_0 results in an adjoint numerator. However, it is not immediately clear that *all* adjoint numerators can be constructed in this way. In the following, we prove that indeed the construction via numerators seeds leads to the complete set of adjoint numerators. For the general proof at any multiplicity, see App. A. First, notice that $\mathbf{J}_4^\mathsf{T}/3$ satisfies $\mathbf{J}_4 \cdot \mathbf{J}_4^\mathsf{T}/3 \cdot \mathbf{J}_4 = \mathbf{J}_4$, and thus (since $\mathbf{J}_4$ encodes the Jacobi identities as the sum of its rows) $\mathbf{J}_4 \cdot \mathbf{J}_4^\mathsf{T}/3 \cdot n_{\text{adj}} = n_{\text{adj}}$ for any vector n_{adj} that satisfies the Jacobi identities. Therefore, $\mathbf{J}_4^\mathsf{T} \cdot n_{\text{adj}}/3$ is the pre-image of any n_{adj}. Importantly $\mathbf{J}_4^\mathsf{T} \cdot n_{\text{adj}}/3$ is also a numerator seed:

$$[\mathbf{J}_4^\mathsf{T} \cdot n_{\text{adj}}](1,3,2,4) = -n_{\text{adj},t} + n_{\text{adj},u} \tag{4.6}$$

is invariant under cyclic permutations of its arguments and the reversal of their order, thanks to the algebraic properties of n_{adj}. This completes the proof. As an example, consider multiplying the adjoint color factor by $\mathbf{J}_4^\mathsf{T}$,

$$[\mathbf{J}_4^\mathsf{T} \cdot c_{\text{adj}}](1,3,2,4) = -f^{13x} f^{x42} + f^{14x} f^{x23} = 3\left[(1324) + (4231)\right] - \sum_{\sigma \in S_3} (1\sigma(234))\,. \tag{4.7}$$

The resulting seed is equivalent to the usual c_0 up to the addition of a fully permutation-invariant quantity (mapped to zero by $\mathbf{J}_4$), as it should since they generate the same adjoint color numerator.

It is illustrative to compare the construction via numerator seeds with the composition method to construct adjoint numerators [34], reviewed in Sec. 2.2. Both the composition rule of Eq. (2.29) and the basic kinematic building block of Eq. (2.30) can be rewritten as a numerator seed multiplied by $\mathbf{J}_4$,

$$n(j,k) = \mathbf{J}_4 \cdot \begin{pmatrix} j_s\, k_s \\ j_t\, k_t \\ j_u\, k_u \end{pmatrix}, \qquad n^{(ss)} = \mathbf{J}_4 \cdot \begin{pmatrix} s \\ t \\ u \end{pmatrix}. \tag{4.8}$$

The composition rule of Eq. (2.31) can similarly be rewritten,

$$n(j,d) = \vec{n}(\mathbf{J}_4 \cdot j_0\,,\ d) = \mathbf{J}_4 \cdot (d\, j_0)\ , \tag{4.9}$$

where j_0 is in fact a numerator seed (vector). This last composition rule encodes the simple statement that a fully symmetric object, such as d, takes an adjoint numerator to another one. This means that, at 4-point, there is a direct correspondence between the construction of adjoint numerators from seeds and by composition. This is however not the case at 5-point, where the composition rules [35] are not equivalent to a simple multiplication by $\mathbf{J}_5$.

[9] Since the 4-point c_0 of Eq. (4.3) is invariant under cyclic permutations and the reversal of its arguments, numerator seeds for example transform as $n_0 \underset{\sigma}{\to} M_{c_0,\sigma} \cdot n_0$, with $M_{c_0,\sigma} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$ for $\sigma = \{1{\to}3{\to}2{\to}4{\to}1\}$ and $M_{c_0,\sigma} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ for $\sigma = \{1 \leftrightarrow 4, 2 \leftrightarrow 3\}$.

4.1 Kinematic numerator seeds

Analogous to Eq. (4.3), purely kinematic numerator seeds for a scalar theory have the form

$$n_0 = \big(n_0(1,3,2,4),\ n_0(1,2,3,4),\ n_0(1,2,4,3)\big)^{\mathsf{T}}, \qquad n_0(a,b,c,d) = g(s_{ac}, s_{ab})\,, \tag{4.10}$$

where g is a function of the Mandelstam invariants $s_{ab} \equiv (p_a + p_b)^2$. Invariance under reversal is automatic at 4-point, while cyclic invariance requires $g(s,t) = g(s,-s-t) = g(s,u)$. A general polynomial expansion of $g(s,t)$ can then be written as[10]

$$g(s,t) = \sum_{i,j=0} \frac{a_{i,j}\, s^i \left(t\,[-s-t]\right)^j}{\Lambda^{2i+4j}} = a_{0,0} + \frac{a_{1,0}s}{\Lambda^2} + \frac{a_{2,0}s^2 + a_{0,1}tu}{\Lambda^4} + \frac{a_{3,0}s^3 + a_{1,1}stu}{\Lambda^6} + \cdots\,. \tag{4.11}$$

Consistent factorization on the poles and assumptions on the particle spectrum of a theory can impose further restrictions on the $a_{i,j}$ coefficients. From this numerator seed, a single-copy scalar amplitude can be constructed following Eq. (3.14), which leads to

$$A_{\mathrm{R}} = \mathbf{m}_4 \cdot n_{0,\mathrm{R}} = -g_\phi^2 \Big(s\, g(s,t) + t\, g(t,s) + u\, g(u,s)\Big) \begin{pmatrix} 1/tu \\ 1/us \\ 1/st \end{pmatrix}, \tag{4.12}$$

where we note that the function $g(s,t)$ only appears through a permutation-invariant overall factor (so that it does not affect the traditional KK and BCJ relations).

4.2 Generalized numerator seeds

At zeroth order in Mandelstam invariants, one can verify that there is only one linear combination of single traces (c_{adj} defined in Eq. (4.2)) that satisfies the adjoint algebraic properties. Therefore, the only necessary numerator seed containing only color information is given by the c_0 vector defined in Eq. (4.3). At this order, any other seed is related to c_0 by the addition of a permutation invariant combination of traces. Such seeds also map to c_{adj} because permutation invariant combination of traces map to zero under multiplication by $\mathbf{J}$.

At higher orders in the kinematics, the most general functional form that is cyclically and reversal invariant is

$$\bar{c}_0^{\,\mathrm{h.d.}}(1,2,3,4) = g(t,s)\, c_0(1,2,3,4) + h(u,t)\, c_0(1,2,4,3) + h(s,t)\, c_0(1,3,2,4)\,, \tag{4.13}$$

from which one defines the generalized seed vector

$$\bar{c}_0^{\,\mathrm{h.d.}} = \big(\bar{c}_0^{\,\mathrm{h.d.}}(1,3,2,4),\ \bar{c}_0^{\,\mathrm{h.d.}}(1,2,3,4),\ \bar{c}^{\,\mathrm{h.d.}}(1,2,4,3)\big)^{\mathsf{T}}, \tag{4.14}$$

[10] To see this, first express $g(s,t)$ in term of $t+u$ and $t-u$. The requirement that $g(s,t) = g(s,-s-t) = g(s,u)$ imposes a symmetry under $t \leftrightarrow u$ exchange which requires that $t-u$ only arises in even powers. However, since $(t-u)^2 = s^2 - 4\,t\,u$, one concludes that $g(s,t)$ can be written as an expansion in powers of just s and $t\,u$, as claimed.

analogously to Eq. (4.3). Here $g(s,t) = g(s,-s-t) = g(s,u)$ which is the same constraint as before (Eq. (4.10)), and $h(s,t)$ is a priori a general function. Notice that it is not trivial that we can write the equation above in terms of the vector c_0 instead of the single traces separately. This is a feature of the 4-point kinematics, which is invariant under reversal of particle labels: $f'(1,2,3,4) \equiv f(s_{12}, s_{13}) = f(s_{43}, s_{42}) \equiv f'(4,3,2,1)$, for any function f of the Mandelstam invariants.[11]

While $\bar{c}_0^{\text{h.d.}}$ is the most general numerator seed, for the purpose of constructing independent adjoint numerators, the $g(t,s)\, c_0(1,2,3,4)$ term is redundant. It can be canceled by adding a permutation invariant function and redefining the arbitrary $h(s,t)$. This means that we can restrict to the numerator seed

$$\bar{c}_0^{\text{h.d.}} = \begin{pmatrix} 0 & h(t,s) & h(u,s) \\ h(s,t) & 0 & h(u,t) \\ h(s,u) & h(t,u) & 0 \end{pmatrix} \cdot c_0 \equiv \mathbf{H}^{\text{h.d.}} \cdot c_0 \,, \tag{4.15}$$

and still generate all possible adjoint numerators with color. We could have reached the same conclusion regarding the fact that the function g can be absorbed in h using the systematic algorithm which makes use of $\mathbf{J}^+$ which we discussed in Sec. 3.

4.3 Generalized gauge transformations

Up to this point, we have considered numerators independently from the amplitudes they generate. There does exist a freedom to shift a numerator without affecting the amplitude, if the other numerators they multiply satisfy Jacobi identities. For instance, at 4-point, the redefinition $n_s \to n_s + s\,\Delta,\ n_t \to n_t + t\,\Delta,\ n_u \to n_u + u\,\Delta$, for any function Δ, results in

$$\mathcal{A}_4 = \frac{c_s\, n_s}{s} + \frac{c_t\, n_t}{t} + \frac{c_u\, n_u}{u} \to \mathcal{A}_4 + (c_s + c_t + c_u)\Delta \,, \tag{4.16}$$

which is just $\mathcal{A}_4$ if the color vector $\vec{c}$ satisfies the Jacobi identity. In matrix notation, any shifts in the vector $\vec{n}$ proportional to $(s,t,u)^\mathsf{T}$ leave the amplitude $\mathcal{A} = c \cdot \mathbf{P}_4 \cdot n$ invariant. Here $(s,t,u)^\mathsf{T}$ is the null vector of $\mathbf{J}_4^\mathsf{T} \cdot \mathbf{P}_4$, where $\mathbf{J}_4^\mathsf{T}$ arises if $\vec{c}$ satisfies the Jacobi identity. Such shifts are called *generalized gauge transformations* because an actual gauge transformation, $\epsilon_i \to \epsilon_i + p_i$ for any particle label i, results in a similar vanishing shift of the amplitude. Nevertheless, generalized gauge transformations are also present in non-gauge theories.

At the level of the numerator seeds, the generalized gauge transformations allow for shifts proportional to the null-vectors of $\tilde{\mathbf{m}}_4 = \mathbf{J}_4^\mathsf{T} \cdot \mathbf{P}_4 \cdot \mathbf{J}_4$, which are $(u,0,-s)^\mathsf{T}$ and $(t,-s,0)^\mathsf{T}$. This includes the permutation invariant shift proportional to $(1,1,1)^\mathsf{T}$ (the null-vector of $\mathbf{J}_4$) that was used before and does not affect the constructed adjoint numerator. Other shifts are possible that change the permutation properties of the seed and, in turn, may correspond to a non-adjoint numerator. A particular generalized gauge

[11] At higher multiplicity n, this is not generally true. While generalized numerator seeds can always be organized in terms of an $(n-1)!/2$ dimensional vector, factorizing out the color from the kinematic dependence generally requires all $(n-1)!$ single traces separately.

transformation, given by

$$\bar{c}_0^{\text{h.d.}} \to c_0^{\text{h.d.}} = \bar{c}_0^{\text{h.d.}} + \begin{pmatrix} \frac{t\,h(s,t)+u\,h(s,u)}{s} & -h(t,s) & -h(u,s) \\ -h(s,t) & \frac{u\,h(t,u)+s\,h(t,s)}{t} & -h(u,t) \\ -h(s,u) & -h(t,u) & \frac{s\,h(u,s)+t\,h(u,t)}{u} \end{pmatrix} \cdot c_0\,, \tag{4.17}$$

has the property that the shift is itself a numerator seed and, therefore, maps into a valid seed $c_0^{\text{h.d.}}$. Hence, to capture all amplitudes in a CK-dual theory, $\bar{c}_0^{\text{h.d.}}$ in Eq. (4.15) can be replaced by

$$c_0^{\text{h.d.}} = \begin{pmatrix} g(s,t) & 0 & 0 \\ 0 & g(t,s) & 0 \\ 0 & 0 & g(u,s) \end{pmatrix} \cdot c_0 \equiv \mathbf{G}^{\text{h.d.}} \cdot c_0, \tag{4.18}$$

with the redefinition

$$g(s,t) = g(s,u) = \frac{t\,h(s,t) + u\,h(s,u)}{s}\,. \tag{4.19}$$

This defines the diagonal matrix $\mathbf{G}^{\text{h.d.}}$ for future reference. We thus find that one can restrict to the functional form

$$c_0^{\text{h.d.}}(1,2,3,4) = g(t,s)\,c_0(1,2,3,4) \tag{4.20}$$

(if $g(s,t)$ is allowed to have simple poles). We stress that this is not the most general numerator seed, nor does it construct the most general adjoint numerator, but it constructs the most general amplitude.

5 Seeds and the generalized KLT bootstrap at 4-point

The constructive approach to generalized adjoint numerators presented in the previous section enables a straightforward comparison with the generalized KLT approach by explicitly building BAS+h.d. amplitudes. Restricting to 4-point in this section, we show that the generalized CK and KLT formalisms are equivalent, and we study what this means for the possible double-copy amplitudes.

The matrix $\mathbf{m}_4^{\text{h.d.}}$ has the general form shown in Eq. (3.18) but generalized gauge transformations allow us to write

$$\begin{aligned} \mathbf{m}_4^{\text{h.d.}} &= \mathbf{G}_{\text{R}}^{\text{h.d.}} \cdot \mathbf{m}_4 \cdot \mathbf{G}_{\text{L}}^{\text{h.d.}} \\ &= g_\phi^2 \begin{pmatrix} -\frac{s\,g_{\text{R}}(s,t)\,g_{\text{L}}(s,t)}{t\,u} & -\frac{g_{\text{R}}(s,t)\,g_{\text{L}}(t,s)}{u} & -\frac{g_{\text{R}}(s,t)\,g_{\text{L}}(u,s)}{t} \\ -\frac{g_{\text{R}}(t,s)\,g_{\text{L}}(s,t)}{u} & -\frac{t\,g_{\text{R}}(t,s)\,g_{\text{L}}(t,s)}{s\,u} & -\frac{g_{\text{R}}(t,s)\,g_{\text{L}}(u,s)}{s} \\ -\frac{g_{\text{R}}(u,s)\,g_{\text{L}}(s,t)}{t} & -\frac{g_{\text{R}}(u,s)\,g_{\text{L}}(t,s)}{s} & -\frac{u\,g_{\text{R}}(u,s)\,g_{\text{L}}(u,s)}{s\,t} \end{pmatrix}, \end{aligned} \tag{5.1}$$

where $\mathbf{G}_{\text{L/R}}$ is defined as in Eq. (4.18) with $g_{\text{L/R}}(s,t)$ instead of $g(s,t)$. In fact, this corresponds to a generalized KLT matrix showed in Eq. (2.18), with the solution to the minimal-rank bootstrap Eqs. (2.16, 2.17) given by

$$f_2(s,t) = -\frac{g_\phi^2}{s}\,g_{\text{R}}(t,s)\,g_{\text{L}}(u,s)\,. \tag{5.2}$$

Another way to see that $\mathbf{m}_4^{\text{h.d.}}$ has minimal rank, is that the diagonal matrices $\mathbf{G}_{\text{L/R}}^{\text{h.d.}}$ have full rank and therefore preserve the rank of the matrix $\mathbf{m}_4$. Conversely, the relation above can be inverted to express $g_{\text{L/R}}$ in terms of f_2 (not uniquely), which means that they encompass any EFT solution f_2 of the bootstrap equations. One can for instance take

$$g_{\text{R}}(s,t) = \left(\frac{f(t,s)\,f(u,s)^2\,f(t,u)}{f(s,u)}\right)^{1/3}, \qquad g_{\text{L}}(s,t) = \left(\frac{f(t,u)\,f(u,t)}{f(s,t)\,f(s,u)}\right)^{1/3}, \tag{5.3}$$

with $f(s,t) \equiv -s\,f_2(s,t)/g_\phi^2$. Thanks to the bootstrap condition for f_2, these satisfy the constraint $g(s,t) = g(s,-s-t) = g(s,u)$, as required for numerator seeds. This shows that the generalized numerator seed of Eq. (4.20) generates any matrix of doubly ordered amplitudes that appears in the generalized KLT formalism. The choice above is not unique since $\mathbf{m}_4^{\text{h.d.}}$ only depends on the product of g_{L} and g_{R}.

To illustrate how Eq. (5.3) works in an EFT expansion, let us consider the lowest-order terms in the bootstrap solution for a pure scalar theory [36],

$$\frac{f_2(s,t)}{g_\phi^2} = -\frac{1}{s} + \frac{1}{\Lambda^4}(a_{1,0}\,t + a_{1,1}\,s) + \frac{a_{2,0}}{\Lambda^6}\,t\,(s+t) + \mathcal{O}\left(\frac{s_{ab}^3}{\Lambda^8}\right), \tag{5.4}$$

which determines the generalized matrix of doubly ordered amplitudes. Besides the bootstrap equation (Eq. (2.17)), further constraints have been imposed on this function, which ensure correct locality properties of the resulting BAS+h.d. theory. From Eq. (5.3), it follows that the matrix determined by f_2 is equivalently obtained through Eq. (5.1) with

$$\begin{aligned} g_{\text{R}}(s,t) &= 1 + \frac{4a_{1,1} - a_{1,0}}{3\Lambda^4}\,t\,(-t-s) + \frac{a_{1,0} - a_{1,1}}{3\Lambda^4}\,s^2 + \frac{a_{2,0}}{\Lambda^6}\,s\,t\,(-s-t) + \mathcal{O}\left(\frac{s_{ab}^4}{\Lambda^8}\right), \\ g_{\text{L}}(s,t) &= 1 - \frac{a_{1,0} - a_{1,1}}{3\Lambda^4}(s^2 + 2\,t\,(-s-t)) + \mathcal{O}\left(\frac{s_{ab}^4}{\Lambda^8}\right). \end{aligned} \tag{5.5}$$

These functions give rise to the numerator seeds and the associated adjoint numerators of the BAS+h.d. theory.

The minimal-rank bootstrap equations also allow for solutions that are not of the BAS+h.d. form. However, modifications to the lowest-order 4-point kernel were found to increase the rank at higher multiplicities and lead to unhealthy double-copy structures [36].

5.1 Generalized single and double copies

From Eq. (5.1), it is now straightforward to obtain the generalized KLT kernel $S_4^{\text{h.d.}}$ in terms of the traditional BAS one, $S_4[\alpha|\beta] = 1/m_4[\beta|\alpha]$:

$$S_4^{\text{h.d.}}[\alpha|\beta] = \frac{1}{\mathbf{G}_{\text{L}}^{\text{h.d.}}[\alpha|\alpha]}\,S_4[\alpha|\beta]\,\frac{1}{\mathbf{G}_{\text{R}}^{\text{h.d.}}[\beta|\beta]}. \tag{5.6}$$

Due to the diagonal structure of $\mathbf{G}_{\text{L/R}}^{\text{h.d.}}$, each entry of the generalized kernel only depends on that of the BAS kernel with the same color ordering. Since the rank of both $\mathbf{m}_4$ and $\mathbf{m}_4^{\text{h.d.}}$ is one, this relation holds for any choice of single color orderings α and β. One can explicitly check that this kernel obeys the KLT relation of Eq. (2.11),

$$m_4^{\text{h.d.}}[\alpha|\beta]\;S_4^{\text{h.d.}}[\beta|\gamma]\;m_4^{\text{h.d.}}[\gamma|\delta] = \mathbf{G}_{\text{R}}^{\text{h.d.}}[\alpha|\alpha]\;m_4[\alpha|\beta]\;S_4[\beta|\gamma]\;m_4[\gamma|\delta]\;\mathbf{G}_{\text{L}}^{\text{h.d.}}[\delta|\delta]$$

$$
\begin{aligned}
&= \mathbf{G}_{\mathrm{R}}^{\mathrm{h.d.}}[\alpha|\alpha]\; m_4[\alpha|\delta]\; \mathbf{G}_{\mathrm{L}}^{\mathrm{h.d.}}[\delta|\delta] \\
&= m_4^{\mathrm{h.d.}}[\alpha|\delta]\,, \qquad (5.7)
\end{aligned}
$$

and the generalized KKBCJ relations for color-ordered amplitudes are given in Eq. (3.21),

$$
\begin{aligned}
&\mathbf{G}_{\mathrm{R}}^{\mathrm{h.d.}}[\alpha|\alpha]\; m_4[\alpha|\beta]\, S_4[\beta|\gamma]\; \frac{1}{\mathbf{G}_{\mathrm{R}}^{\mathrm{h.d.}}[\gamma|\gamma]}\; A'_{\mathrm{R}}[\gamma] = A'_{\mathrm{R}}[\alpha]\,, \\
&A'_{\mathrm{L}}[\alpha]\; \frac{1}{\mathbf{G}_{\mathrm{L}}^{\mathrm{h.d.}}[\alpha|\alpha]}\; S_4[\alpha|\beta]\, m_4[\beta|\gamma]\; \mathbf{G}_{\mathrm{L}}^{\mathrm{h.d.}}[\gamma|\gamma] = A'_{\mathrm{L}}[\gamma]\,. \qquad (5.8)
\end{aligned}
$$

For example, Eq. (2.9) is generalized to

$$
A'_{\mathrm{R}}[1234] = \frac{t\, g_{\mathrm{R}}(t,s)}{u\, g_{\mathrm{R}}(u,s)}\, A'_{\mathrm{R}}[1243]\,, \qquad (5.9)
$$

which reduces to the traditional BCJ relation for $g_{\mathrm{R}}(s,t) = 1$. It is worth remarking that the KKBCJ relations for A'_{R} (A'_{L}) depend only on $\mathbf{G}_{\mathrm{R}}^{\mathrm{h.d.}}$ ($\mathbf{G}_{\mathrm{L}}^{\mathrm{h.d.}}$).

Interestingly, the explicit form of the generalized KKBCJ relations points to an object

$$
A_{\mathrm{R}}[\alpha] \equiv \frac{A'_{\mathrm{R}}[\alpha]}{\mathbf{G}_{\mathrm{R}}^{\mathrm{h.d.}}[\alpha|\alpha]}\,, \qquad (5.10)
$$

which obeys the traditional KK and BCJ relations, and similarly for A_{L}. Although the notation suggests otherwise, A_{R} may still contain higher-derivative corrections, but they are such that they do not affect the form of the traditional KK and BCJ relations. If A_{R} can be argued to be a valid amplitude, this could be an efficient method to construct generalized single copies. In addition, it would imply that the generalized KLT formalism does not lead to double copies with additional higher-derivative corrections besides the ones which can be obtained with the usual KLT kernel. Indeed,

$$
\mathcal{M}' = A'_{\mathrm{L}}[\alpha]\, S_4^{\mathrm{h.d.}}[\alpha|\beta]\, A'_{\mathrm{R}}[\beta] = A_{\mathrm{L}}[\alpha]\, S_4[\alpha|\beta]\, A_{\mathrm{R}}[\beta] = \mathcal{M}\,, \qquad (5.11)
$$

where we stress that both primed and unprimed amplitudes may contain higher-derivative corrections and S_4 stands for the traditional BAS KLT kernel.

However, as it stands, $A_{\mathrm{R}}[\alpha]$ cannot be interpreted as a physical amplitude, since it does not necessarily factorize properly on all channels, whose associated residues can be affected by $\mathbf{G}_{\mathrm{R}}^{\mathrm{h.d.}}$. We will study the functional form of $\mathbf{G}_{\mathrm{L/R}}^{\mathrm{h.d.}}$, under the assumption of a fixed BAS particle content in Sec. 5.2. This will lead to a slightly adapted but equivalent form for Eq. (5.10), such that $A_{\mathrm{R}}[\alpha]$ and $A'_{\mathrm{R}}[\alpha]$ have the same residues on all poles.

Leaving momentarily aside the question of physical residues, Eq. (5.10) also provides a means to construct generalized single copies. They can be obtained by multiplying amplitudes that satisfy the traditional KK and BCJ relations with $\mathbf{G}_{\mathrm{L/R}}^{\mathrm{h.d.}}$. The same conclusion is reached in the numerator formalism (see Eq. (3.21)). Explicitly, comparing with Eq. (4.12), we have that

$$
A'_{\mathrm{R}} = \mathbf{G}_{\mathrm{R}}^{\mathrm{h.d.}} \cdot \mathbf{m}_4 \cdot n_0 = -g_\phi^2 \Big(s\, g(s,t) + t\, g(t,s) + u\, g(u,s) \Big) \begin{pmatrix} g_{\mathrm{R}}(s,t)/tu \\ g_{\mathrm{R}}(t,s)/us \\ g_{\mathrm{R}}(u,s)/st \end{pmatrix}\,, \qquad (5.12)
$$

where $n_0 = \big(g(s,t), g(t,s), g(u,s)\big)^{\mathsf{T}}$. All generalized amplitudes can be constructed in this way, showing that one can associate a generalized color numerator to any amplitude obtained from the generalized KKBCJ relations. The first equality in Eq. (5.12) also applies to gauge theories in which case n_0 contains polarization vectors. Imposing particular locality properties on this amplitude restricts the coefficients inside $g(s,t)$ and $g_{\mathrm{R}}(s,t)$. We note that such constraints may be less restrictive than the constraints on $g_{\mathrm{L/R}}$ coming from imposing a fixed particle content on $\mathbf{m}_4^{\text{h.d.}}$. We will get back to this point in the following.

5.2 Factorization properties and particle spectrum

So far, we have not been concerned with the particle content (and factorization properties) of the BAS+h.d. amplitudes constructed through Eq. (5.1). Following [36], we now impose that $\mathbf{m}_4^{\text{h.d.}}$ reduces to the BAS matrix $\mathbf{m}_4$ at lowest order and that the particle content of the theory is fixed to one bi-adjoint scalar. This implies that $f(s,t) = -s\, f_2(s,t)/g_\phi^2$ in Eq. (5.3) is a polynomial of the form $1 + \mathcal{O}(s_{ab}/\Lambda^2)$. Interpreting Eq. (5.3) up to a fixed order in the $1/\Lambda$ expansion, the functions $g_{\mathrm{L/R}}(s,t)$ are then also of the same form.

While double or spurious poles are avoided in the construction of $\mathbf{m}_4^{\text{h.d.}}$ with polynomial $g_{\mathrm{L/R}}(s,t)$, the residues might be modified. Such modifications are either non-physical or can be interpreted as new particles appearing in the factorization channels. However, at 4-point, only contact-term higher-derivative corrections are allowed with a fixed single scalar particle content (since the 3-point amplitudes are not modified in the solution to the KLT bootstrap [36]). Imposing such conditions yields the following constraints:

$$\begin{aligned} m_4[1324|1324] &= \frac{g_{\mathrm{R}}(s,t)g_{\mathrm{L}}(s,t)}{t} + \frac{g_{\mathrm{R}}(s,t)g_{\mathrm{L}}(s,t)}{u} \underset{t\to 0}{\sim} \frac{1}{t} \implies g_{\mathrm{R}}(s,0)g_{\mathrm{L}}(s,0) = 1\,, \\ m_4[1324|1234] &= -\frac{g_{\mathrm{R}}(s,t)g_{\mathrm{L}}(t,s)}{u} \underset{u\to 0}{\sim} -\frac{1}{u} \implies g_{\mathrm{R}}(-t,t)g_{\mathrm{L}}(t,-t) = 1\,, \end{aligned} \tag{5.13}$$

which enforce in particular the consistency conditions

$$g_{\mathrm{L/R}}(s,0) = g_{\mathrm{L/R}}(-s,0)\,, \tag{5.14}$$

when using $g_{\mathrm{L/R}}(s,t) = g_{\mathrm{L/R}}(s,-s-t)$. This implies that, on the t and u poles, the first variable of $g(s,t)$ necessarily appears in even powers. The lowest order terms in the solutions are then

$$\begin{aligned} g_{\mathrm{L}}(s,t) &= 1 + \frac{a_{\mathrm{L}2,0}\, s^2 + a_{\mathrm{L}0,1}\, t\, u}{\Lambda^4} + \frac{a_{\mathrm{L}1,1}\, s\, t\, u}{\Lambda^6}\,, \\ g_{\mathrm{R}}(s,t) &= 1 + \frac{-a_{\mathrm{L}2,0}\, s^2 + a_{\mathrm{R}0,1}\, t\, u}{\Lambda^4} + \frac{a_{\mathrm{R}1,1}\, s\, t\, u}{\Lambda^6}\,, \end{aligned} \tag{5.15}$$

where we emphasize that the same coefficient $a_{\mathrm{L}2,0}$ appears in both expansions. These solutions are fully consistent with the expansions in Eq. (5.5), which were obtained from the bootstrap solution $f_2(s,t)$, assuming a fixed particle content [36].[12] With these solutions,

[12] The consistency conditions also imply that $f_2(s,t)$ does not contain even powers of just s. Since $g_{\mathrm{L/R}}(s,t) = g_{\mathrm{L/R}}(s,-s-t)$ takes $f_2(s,0) \propto g_{\mathrm{R}}(0,s)\, g_{\mathrm{L}}(-s,s)/s$ to $g_{\mathrm{R}}(0,-s)g_{\mathrm{L}}(-s,0)/s$, the conditions of Eq. (5.14) indeed constrain $f_2(s,0)$ to be an odd function of s.

one of the generalized KKBCJ relations, showed in Eq. (5.9), is given by

$$\left[1+\frac{(a_{\mathrm{L}2,0}-a_{\mathrm{R}0,1})\,s(t-u)}{\Lambda^4}+\mathcal{O}\left(\frac{s_{ab}^4}{\Lambda^8}\right)\right]\frac{t}{u}\,A'_{\mathrm{R}}[1243]=A'_{\mathrm{R}}[1234]\,. \tag{5.16}$$

Now let us turn our attention to the object $A_{\mathrm{R}}[\alpha]\equiv A'_{\mathrm{R}}[\alpha]/\mathbf{G}_{\mathrm{R}}^{\mathrm{h.d.}}[\alpha|\alpha]$ that appeared in Eq. (5.10). Given that the functions g_{R} in $\mathbf{G}_{\mathrm{R}}^{\mathrm{h.d.}}[\alpha|\alpha]$ are polynomials, the $A_{\mathrm{R}}[\alpha]$ and $A'_{\mathrm{R}}[\alpha]$ functions have the same poles. However, it is not guaranteed that the residues on these poles are consistent. In particular, on any of the poles of $A_{\mathrm{R}}[\alpha]$, the function $\mathbf{G}_{\mathrm{R}}^{\mathrm{h.d.}}[\alpha|\alpha]$ contributes a non-trivial inverse factor of $g_{\mathrm{R}}(s,0)$ (or $g_{\mathrm{R}}(t,0)$). On the poles, these factors can also be obtained from the function $g_{\mathrm{R}}(\sqrt{(s^2+t^2+u^2)/2},0)|_{t\to 0}=g_{\mathrm{R}}(s,0)$, where the square root always appears in even powers in the Taylor expansion over the first variable thanks to the consistency condition of Eq. (5.14). Since $g_{\mathrm{R}}(\sqrt{(s^2+t^2+u^2)/2},0)$ is permutation invariant, we can redefine the amplitude $A_{\mathrm{R}}[\alpha]$ as

$$A_{\mathrm{R}}[\alpha]\equiv\frac{g_{\mathrm{R}}\left(\sqrt{\frac{s^2+t^2+u^2}{2}},0\right)}{\mathbf{G}_{\mathrm{R}}^{\mathrm{h.d.}}[\alpha|\alpha]}A'_{\mathrm{R}}[\alpha] \tag{5.17}$$

which still satisfies the traditional KK and BCJ relations while also having the same poles and residues as $A'_{\mathrm{R}}[\alpha]$. Defining simultaneously

$$A_{\mathrm{L}}[\alpha]\equiv\frac{g_{\mathrm{L}}\left(\sqrt{\frac{s^2+t^2+u^2}{2}},0\right)}{\mathbf{G}_{\mathrm{L}}^{\mathrm{h.d.}}[\alpha|\alpha]}A'_{\mathrm{L}}[\alpha]\,, \tag{5.18}$$

it follows from Eq. (5.13) that $g_{\mathrm{L}}\left(\sqrt{(s^2+t^2+u^2)/2},0\right)=1/g_{\mathrm{R}}\left(\sqrt{(s^2+t^2+u^2)/2},0\right)$, which leads to the conclusion that the double-copy amplitude remains unchanged. In other words, there is an interplay between the left and right amplitudes and the generalized kernel that allows for the cancellation of any correction to the kernel, in a manner that does not affect the residues and poles of the amplitudes.

We therefore conclude that a double-copy amplitude obtained with a generalized kernel can equivalently be generated with the traditional BAS kernel (c.f. Eq. (5.11)). The single-copy amplitudes may then still include higher-derivative corrections, but only those that do not spoil the usual KK and BCJ relations. To prove this statement, we assumed that no extra particles are added to the BAS spectrum. At 4-point, the generalization of the KLT formalism does thus not enlarge the space of possible double copies, but it does enlarge the space of single copies that can be used as input.

In our derivation, it is also clear that Eqs (5.17, 5.18) can be used to obtain amplitudes that satisfy generalized KKBCJ relations ($A'_{\mathrm{L/R}}$) from amplitudes satisfying the usual KK and BCJ relations ($A_{\mathrm{L/R}}$), with the same particle content. This, indeed, has exactly the same form in the generalized numerators approach shown in Eq. (5.12).

6 Results at 5-point

Here, we use the numerator seeds described in Sec. 3 to construct adjoint numerators at 5-point. We also discuss the seed redundancies, previously described at 4-point in

Sec. 4. The statement that the generalized KLT formalism does not enlarge the space of double-copy amplitudes had two caveats, see Sec. 3.2. While these were fully addressed at 4-point in the previous sections, we only provide partial results at 5-point. We check explicitly that the generalized numerators generate all the leading-order KLT kernels of [36]. Furthermore, we study the factorization properties and the particle spectrum of the objects $A_{\mathrm{L/R}} = (\mathbf{H}^{\mathrm{h.d.}}_{\mathrm{L/R}})^{-1} \cdot A'_{\mathrm{L/R}}$ for a restricted set of higher-derivative corrections. For the corresponding generalized kernels, we achieve the same conclusion as at 4-point, namely that the double-copy amplitudes it produces can equally be obtained with the traditional BAS kernel and physical single-copy amplitudes.

6.1 Numerator seeds

At 5-point, the kinematic numerator seeds have the functional form

$$n_0(1,2,3,4,5) = \big(g(s_{12}, s_{23}, s_{34}, s_{45}, s_{51}) - g(s_{51}, s_{45}, s_{34}, s_{23}, s_{12}) \big) + \text{cyclic}\,, \tag{6.1}$$

where we have demanded invariance under cyclic permutations and antisymmetry under reversal. Imposing this antisymmetry is however optional, as the components symmetric under reversal are mapped to zero by the $\mathbf{J}$-matrix. We do not consider the parity-odd fully antisymmetric contraction of four independent momenta, which may lead to additional numerators. The numerator seed vector is then given by

$$n_0 = \big(n_0(1,2,3,4,5), \quad \text{all permutations of } \{2,3,...,n\}\big)^{\mathsf{T}}\,, \tag{6.2}$$

and all scalar adjoint numerators are easily built from $n_{\text{adj}} = \mathbf{J}_5 \cdot n_0$. The explicit form of $\mathbf{J}_5$ is given in Sec. C.1. We chose a cyclic basis of Mandelstam invariants, which is possible at any multiplicity (see e.g. [52]) and simplifies the particle permutation properties.

In addition, there are two independent adjoint numerators containing only color information. These are obtained from the following single-trace seeds,

$$\begin{aligned} c_{0,1} &= \big((12345), \quad \text{all permutations of } \{2,3,...,n\}\big)^{\mathsf{T}}\,, \\ c_{0,2} &= \big((13524), \quad \text{all permutations of } \{2,3,...,n\}\big)^{\mathsf{T}}\,. \end{aligned} \tag{6.3}$$

As before, it is not necessary to impose antisymmetry under reversal. Any other allowed single-trace color numerator maps to a linear combination of $\mathbf{J}_5 \cdot c_{0,1}$ and $\mathbf{J}_5 \cdot c_{0,2}$. For instance, the two independent adjoint numerators $c_{\mathfrak{a},1,2}$ of [35] are[13]

$$c_{\mathfrak{a},1} = \mathbf{J}_5 \cdot c_{0,1}, \qquad c_{\mathfrak{a},2} = \frac{1}{6}\,\mathbf{J}_5 \cdot (3c_{0,2} - c_{0,1})\,. \tag{6.4}$$

Besides numerator seeds featuring only kinematic and color information, we study generalized seeds that contain both, i.e.

$$c_0^{\mathrm{h.d.}}(1,2,3,4,5) = f(1,2,3,4,5) + \text{cyclic}, \tag{6.5}$$

[13] The original definitions of these two objects are $c_{\mathfrak{a},1}(1,2,3,4,5) \equiv f^{12x} f^{x3y} f^{y45}$ and $c_{\mathfrak{a},2}(1,2,3,4,5) \equiv 2\,d^{123x} f^{x45} + d^{124x} f^{x35} - d^{125x} f^{x34} + 2\,d^{234x} f^{x15} - 2\,d^{235x} f^{x14}$ [35].

where $f(...)$ is a function of Mandelstam invariants and color structures. Independent numerator seeds can give rise to redundant adjoint numerators, as discussed in the 4-point case. For example, if $g(...)$ is cyclic in its arguments (just like a single trace), then choosing $f(1,2,3,4,5) = g(1,2,3,5,4)$ generates the same adjoint numerator as a linear combination of $c_{0,1}^{\text{h.d.}}(1,2,3,4,5) = g(1,2,3,4,5)$ and $c_{0,2}^{\text{h.d.}}(1,2,3,4,5) = g(1,3,5,2,4)$. This can be established a priori, in analogy with the fact that $\bar{c}_0(1,2,3,4,5) = (12354) + \text{cyclic}$ is redundant with the two pure color seeds of Eq. (6.3). We leave the exploration of the independent basis of seeds for future work. In practice, it is straightforward to identify the independent numerators from the over-complete set built from seeds.

6.2 BAS+h.d. amplitudes

At any multiplicity, the BAS matrix of doubly color-ordered amplitudes that derives from the generalized numerators can be written as (see Eq. (3.18))

$$\mathbf{m}_n^{\text{h.d.}} = \mathbf{H}_{\text{R}}^{\text{h.d.}} \cdot \mathbf{m}_n \cdot \mathbf{H}_{\text{L}}^{\text{h.d.}} \,. \tag{6.6}$$

At 4-point, using generalized gauge transformations, we have previously shown that $\mathbf{H}_{\text{L/R}}^{\text{h.d.}}$ can be taken to be diagonal matrices, and that these capture all solutions to the KLT bootstrap. At 5-point, the generalized gauge transformations can reduce any single-trace numerator seed to the linear combination

$$c_0^{\text{h.d.}}(1,2,3,4,5) = g_1(1,2,3,4,5)\,(12345) \;+ g_2(1,2,3,4,5)\,(13524), \tag{6.7}$$

for independent g_1 and g_2 functions of the Mandelstam invariants that are cyclically symmetric. Even if an original seed is free of poles, capturing the same amplitude with $c_0^{\text{h.d.}}$ may introduce poles. For example, the amplitudes constructed from

$$\bar{c}_0^{\text{h.d.}}(1,2,3,4,5) = h(1,2,3,4,5)\,(12354) + \text{cyclic}\,, \tag{6.8}$$

for any function h depending only on the Mandelstam invariants, may equivalently be constructed from

$$g_1(1,2,3,4,5) = \frac{-h(3,4,5,2,1) - h(5,1,2,4,3)}{s_{12}s_{34}\; m_5[12345|12345]} + \text{cyclic}\,, \tag{6.9}$$

and

$$g_2(1,2,3,4,5) = \frac{h(3,5,2,1,4)}{s_{12}s_{34}\; m_5[12345|12345]} + \text{cyclic}\,, \tag{6.10}$$

where $m_5[12345|12345] = 1/s_{12}s_{34} + \text{cyclic}$, showing that the functions g_1 and g_2 are in general likely to have poles. This means that an amplitude built with $c_0^{\text{h.d.}}$ for those special values of $g_{1,2}$ is equal to an amplitude built with $\bar{c}_0^{\text{h.d.}}$, because both seeds are related by a generalized gauge transformation (nevertheless, the adjoint numerators obtained via these seeds are linearly independent). Therefore any 5-point BAS matrix of doubly color-ordered amplitudes can be written as

$$\mathbf{m}_5^{\text{h.d.}} = \left(\mathbf{G}_{1,\text{R}}^{\text{h.d.}} + \mathbf{G}_{2,\text{R}}^{\text{h.d.}}\right) \cdot \mathbf{m}_5 \cdot \left(\mathbf{G}_{1,\text{L}}^{\text{h.d.}} + \mathbf{G}_{2,\text{L}}^{\text{h.d.}}\right)\,, \tag{6.11}$$

where $\mathbf{G}^{\text{h.d.}}_{1,2}$ follow from the numerator seed defined by Eq. (6.7) by stripping off the single traces. The $\mathbf{G}^{\text{h.d.}}_{1}$ is *diagonal* and $\mathbf{G}^{\text{h.d.}}_{2}$ is *non-diagonal*, with only one non-zero entry on each row/column, as can be seen in Sec. C.1. The non-diagonal matrix is necessary because the 5-point BAS matrix $\mathbf{m}_5$ contains zero entries, which may become non-zero when higher-derivative corrections are included.

It is non-trivial to verify whether Eq. (6.11) covers all solutions to the KLT bootstrap of [36]. As argued in general in Sec. 3.2, $\mathbf{m}^{\text{h.d.}}_5$ has minimal rank just as $\mathbf{m}_5$ since $\mathbf{G}^{\text{h.d.}}_{1,\text{R/L}} + \mathbf{G}^{\text{h.d.}}_{2,\text{R/L}}$ has full rank. We have reproduced the solution of the KLT bootstrap for all orders explicitly provided in [36] and the forms of the necessary functions are listed in Sec. C.2. Going beyond this, we also checked that the numerator seeds reproduce the lowest-order 5-point contact terms, which are cubic in the Mandelstam invariants. These are captured by Eq. (6.11) with

$$\begin{aligned} g_{1,\text{L}}(1,2,3,4,5) &= 1 + \frac{c_1}{\Lambda^{10}}\, s^5_{\text{cyclic}}\,, \quad g_{2,\text{L}}(1,2,3,4,5) = \frac{c_2}{\Lambda^{10}}\, s^5_{\text{cyclic}}\,, \\ g_{1,\text{R}}(1,2,3,4,5) &= 1 + \frac{c_3}{\Lambda^{10}}\, s^5_{\text{cyclic}}\,, \quad g_{2,\text{R}}(1,2,3,4,5) = \frac{c_4}{\Lambda^{10}}\, s^5_{\text{cyclic}}\,, \end{aligned} \tag{6.12}$$

where $s^5_{\text{cyclic}} \equiv s_{12}s_{23}s_{34}s_{45}s_{51}$ and c_i are free parameters. Recall that $\mathbf{m}_5$ is $\mathcal{O}(1/s^2_{ab})$ so that the resulting $\mathbf{m}^{\text{h.d.}}_5$ is of third order in the Mandelstam invariants. This exposes a simple structure of higher-order corrections.

6.3 Factorization properties and particle spectrum

As discussed in Sec. 3.2, given a single-copy amplitude A' that satisfies the generalized KKBCJ relation, the object $A = (\mathbf{H}^{\text{h.d.}})^{-1} \cdot A'$ satisfies the usual KK and BCJ relations (where we momentarily omitted the subscript L/R, for simplicity). Moreover, the same double copy can be constructed using either of these two single copies. However, it is not immediately clear that A and A' share the same analytic properties. At 4-point, we showed that this is indeed the case. At 5-point, achieving a fully general proof seems far more challenging. Therefore, we start the exploration of this question with simplifying assumptions.

Since one can write $\mathbf{H}^{\text{h.d.}} = (\mathbf{G}_1 + \mathbf{G}_2)$ at 5-point, it follows that $A' = (\mathbf{G}_1 + \mathbf{G}_2) \cdot A$. Restricting to the two independent 12345 and 13524 color orderings, we have that

$$\begin{aligned} A'[12345] &= g_1(1,2,3,4,5)A[12345] + g_2(1,2,3,4,5)A[13524]\,, \\ A'[13524] &= g_1(1,3,5,2,4)A[13524] + g_2(1,3,5,2,4)A[15432]\,. \end{aligned} \tag{6.13}$$

This is a simple choice of BCJ basis since the amplitudes do not share any poles (see Sec. C.3 for more details). The full set of ordered amplitudes can be reconstructed using the KKBCJ relations. The fact that $A[15432]$ satisfies the traditional KK reflection relation $A[15432] = -A[12345]$ allows to write

$$\begin{pmatrix} A[12345] \\ A[13524] \end{pmatrix} = \begin{pmatrix} g_1(1,2,3,4,5) & g_2(1,2,3,4,5) \\ -g_2(1,3,5,2,4) & g_1(1,3,5,2,4) \end{pmatrix}^{-1} \cdot \begin{pmatrix} A'[12345] \\ A'[13524] \end{pmatrix}\,. \tag{6.14}$$

Therefore, the unprimed amplitude can be written as

$$A[12345] = \frac{A'[12345]g_1(1,3,5,2,4) + A'[13524]g_2(1,3,5,2,4)}{g_1(1,2,3,4,5)g_1(1,3,5,2,4) + g_2(1,2,3,4,5)g_2(1,3,5,2,4)} \,. \tag{6.15}$$

Studying the analytic properties of the amplitude A is challenging for two main reasons. First, as discussed in Sec. 6.2 and in contrast to the 4-point case, the functions g_1 and g_2 may contain poles. So a general parametrization has a complicated form and studying whether the poles of A and A' agree is non-trivial. Second, even if we take analytic g_1, g_2 so that A and A' poles are identical, a suitable redefinition of A may be necessary to guarantee that its residues match those of A' (similarly to the 4-point case presented in Eq. (5.17)).

Working out such a redefinition at 5-point, or even proving that one always exists, is beyond the scope of this paper. However, in the simplest setup, namely that of a kernel whose first EFT correction is a 5-point contact term constructed from a vanishing g_2 and a polynomial g_1, we can identify a suitable redefinition of $A_{\mathrm{L/R}}$ such that their residues agree with those of $A'_{\mathrm{L/R}}$ while leaving the double copy unchanged (we checked this property up to order $\mathcal{O}(1/\Lambda^{10})$). Explicitly, we found that the locality properties of the kernel at 5-point impose that (for seeds up to order $\mathcal{O}(1/\Lambda^{10})$)

$$g_{1,\mathrm{L/R}}(1,2,3,4,5) = p_{\mathrm{L/R}}(1,2,3,4,5) + \frac{c_{\mathrm{L/R}}}{\Lambda^{10}} s_{12}s_{23}s_{34}s_{45}s_{51} \,, \tag{6.16}$$

where $c_{\mathrm{L/R}}$ are free constants and $p_{\mathrm{L/R}}(1,2,3,4,5) = p_{\mathrm{L/R}}$ are permutation-invariant functions such that $p_{\mathrm{L}}p_{\mathrm{R}} = 1$ whenever a Mandelstam invariant vanishes. Consequently,

$$p_{\mathrm{L}}A_{\mathrm{L}}[\alpha] \quad \text{and} \quad \frac{A_{\mathrm{R}}[\alpha]}{p_{\mathrm{L}}} \,, \tag{6.17}$$

have the same residues on the poles as A'_{L} and A'_{R}, respectively. Therefore, the double-copy amplitudes associated to the kernel obtained from the seeds of Eq. (6.16) can also be obtained with the traditional KLT kernel and single-copy amplitudes which verify the usual KKBCJ relations. Whether such manipulations can be performed in full generality at 5-point is a question left for future investigation.

7 Conclusions and outlook

In this paper, we revisited recent proposals for the systematic double copy of effective field theories. Inspired by the decomposition of adjoint color factors into single traces of Lie algebra generators, we proposed a method to construct generalized adjoint numerators. These satisfy Jacobi-like and antisymmetry relations, while depending on both color and kinematics. Starting from *numerator seeds* satisfying the permutation properties of single traces, we proved that all adjoint numerators can be obtained through the linear map $\mathbf{J}$ between single-trace and adjoint color factors. While generalized numerators have previously been constructed up to 5-point in [34, 35], the construction from numerator seeds is advantageous because the algebraic properties of single traces are simpler than the adjoint algebraic relations. We showed that this method works for any multiplicity

and it is convenient to explore the higher-derivative corrections that allow for a color-dual representation.

The matrix $\mathbf{J}$ is also instrumental in relating the amplitudes represented with trivalent graphs involving (generalized) adjoint numerators to color-ordered ones. The construction of generalized adjoint numerators therefore facilitates the comparison between the generalized numerators construction of [34, 35] and the generalized KLT formalism of [36]. At 4-point, we showed that the generalized adjoint numerators encode all the KLT bootstrap solutions, to any order in the EFT expansion. The two approaches therefore allow for exactly the same higher-derivative corrections to the bi-adjoint scalar amplitudes. The single-copy amplitudes are also the same in the two formalisms. This insight consequently exposed the structure of double-copy amplitudes. While the generalized KLT formalism does expand the range of operators in the single-copy amplitudes, we find (at 4-point) that any resulting double copy can also be obtained with the traditional KLT kernel. We provide partial 5-point results suggesting that these conclusions may extend to higher multiplicity. However, due to the more complicated KLT bootstrap and structures involved, further investigations on higher multiplicities are left to future work.

There are several directions that deserve further attention. For example, we have focused on the construction of scalar numerators, as opposed to gauge-theory numerators involving polarization vectors. The relevant purely kinematic numerator seeds have been considered for Yang–Mills theory in [40–44], but not in the EFT context or for generalized numerators. Gauge invariance must hold at the level of amplitudes (i.e. $\mathcal{A}|_{\epsilon_i \to p_i} = 0$), and individual entries in a numerator are typically not gauge invariant. The necessary extra constraint on numerator seeds may therefore take a complicated form, especially beyond 4-point. Methods proposed to identify the possible gauge-invariant structures in Yang–Mills theories [52, 53] could be useful. In particular, a basis of cyclically invariant structures which can be used as numerator seeds was provided in [53].

Furthermore, we have only considered color factors consisting of single traces of Lie algebra generators. However, the construction of adjoint numerators from numerator seeds does apply more generally. For instance, at 4-point there exist theories with double traces, which can be combined into the seed $c_0(1,3,2,4) = \mathrm{Tr}(T^{a_1}T^{a_2})\mathrm{Tr}(T^{a_3}T^{a_4}) = \delta^{12}\delta^{34}$. The resulting generalized adjoint numerator was previously identified in [31]. Constructing generalized color factors involving products of traces may lead to interesting new single or double copies, which would also be worth studying in the generalized KLT formalism.

Besides the assumption of single traces, we have not included the possibility of extra particles beyond the bi-adjoint scalar discussed in Sec. 5. For the double copy of a gauge theory with matter in the fundamental representation, the KLT kernel is for instance constructed from a *bicolor* theory containing two scalars [10, 12]. Including new particles in factorization channels, or even externally, in the generalized KLT formalism would allow for a larger space of single-copy amplitudes. It would then be worthwhile to extend the numerator seeds to amplitudes with a more complicated particle spectrum.

Altogether, the simple construction of adjoint numerators from numerator seeds has been useful to explore the structures in the generalized CK and KLT double-copy formalisms. Still, the double copy of effective field theories retains various unexplored aspects

which promise exciting new findings for the years to come.

Acknowledgments

We thank Jan Plefka and Johannes Broedel for helpful discussions. This work is supported by the Deutsche Forschungsgemeinschaft under Germany's Excellence Strategy EXC 2121 "Quantum Universe" - 390833306. J.R.N. is supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Projektnummer 417533893/GRK2575 "Rethinking Quantum Field Theory".

A Proof: numerator seeds construct all adjoint numerators

In this appendix, we prove that any adjoint numerator can be constructed from a numerator seed, as in Eq. (3.9). For any (not necessarily square) matrix $\mathbf{J}$, there exists a unique *pseudoinverse* $\mathbf{J}^+$ (see for instance [54]) satisfying

$$\mathbf{J} \cdot \mathbf{J}^+ \cdot \mathbf{J} = \mathbf{J} \tag{A.1}$$

(together with $\mathbf{J}^+ \cdot \mathbf{J} \cdot \mathbf{J}^+ = \mathbf{J}^+$, $(\mathbf{J} \cdot \mathbf{J}^+)^\mathsf{T} = \mathbf{J} \cdot \mathbf{J}^+$, and $(\mathbf{J}^+ \cdot \mathbf{J})^\mathsf{T} = \mathbf{J}^+ \cdot \mathbf{J}$ for real $\mathbf{J}$, which will be of less importance in the following) That is, $\mathbf{J} \cdot \mathbf{J}^+$ maps all columns of $\mathbf{J}$ to themselves. Another word for pseudoinverse is the Moore–Penrose inverse. It can be obtained, for example, by the rank decomposition as in Eq. (3.5). Here $(2n-5)!!$ is the number of trivalent graphs. We then have

$$\mathbf{J}^+ = \mathbf{B}^\mathsf{T} \cdot (\mathbf{B} \cdot \mathbf{B}^\mathsf{T})^{-1} \cdot (\mathbf{A}^\mathsf{T} \cdot \mathbf{A})^{-1} \cdot \mathbf{A}^\mathsf{T} \,. \tag{A.2}$$

At 4- and 5-point, the pseudoinverse takes a particularly simple form: $\mathbf{J}_4^+ = \mathbf{J}_4^\mathsf{T}/6$ and $\mathbf{J}_5^+ = \mathbf{J}_5^\mathsf{T}/20$. At 6-point, the pseudoinverse is not proportional to $\mathbf{J}_6^\mathsf{T}$, but can still be obtained algorithmically.

Since the adjoint color factor c_{adj} is constructed by $\mathbf{J}$, Eq. (A.1) allows to express the Jacobi identities in matrix form,

$$(\mathbb{1} - \mathbf{J} \cdot \mathbf{J}^+) \cdot \vec{c}_{\text{adj}} = 0 \,. \tag{A.3}$$

By definition, (generalized) color-kinematics duality states that adjoint numerators should obey the same algebraic relations as c_{adj}. In other words, as pointed out in e.g. [18, 44], it implies that all adjoint numerators should live in the null-space of $\mathbb{1} - \mathbf{J} \cdot \mathbf{J}^+$. Therefore, *any* adjoint numerator $\vec{n}$ (with only kinematics or both color and kinematics) that satisfies the Jacobi relations obeys

$$\mathbf{J} \cdot \mathbf{J}^+ \cdot \vec{n} = \vec{n} \,. \tag{A.4}$$

This crucially implies that any adjoint numerator at any multiplicity can be written as the $\mathbf{J}$-matrix multiplying a vector,

$$\vec{n}_0 = \mathbf{J}^+ \cdot \vec{n} \,. \tag{A.5}$$

Below, we show that $\vec{n}_0$ constructed in this way transforms in the same way as c_0 under permutations of the particle labels. It is thus a numerator seed. Therefore, all adjoint numerators can be constructed from numerator seeds by multiplication with $\mathbf{J}$.

$\vec{n}_0 = \mathbf{J}^+ \cdot \vec{n}$ is a numerator seed. In this subsection, we prove that the numerator seeds constructed through Eq. (A.5), transform according to the same rule as the color factor c_0,

$$c_0 \underset{\sigma}{\to} M_{c_0,\sigma} \cdot c_0 \,, \tag{A.6}$$

under a permutation of the particle labels, σ.

First, note that adjoint numerators transform analogously under σ,

$$n_{\text{adj}} \underset{\sigma}{\to} M_{c_{\text{adj}},\sigma} \cdot n_{\text{adj}} \,. \tag{A.7}$$

with the same $M_{c_{\text{adj}},\sigma}\,(\neq M_{c_0,\sigma})$ for any adjoint numerator. Since the entries of c_0 are linearly independent,

$$\begin{aligned} c_{\text{adj}} = \mathbf{J} \cdot c_0 \;\underset{\sigma}{\to}\; M_{c_{\text{adj}},\sigma} \cdot \mathbf{J} \cdot c_0 = \mathbf{J} \cdot M_{c_0,\sigma} \cdot c_0 \\ \implies M_{c_{\text{adj}},\sigma} \cdot \mathbf{J} = \mathbf{J} \cdot M_{c_0,\sigma} \,. \end{aligned} \tag{A.8}$$

The numerator seeds transform as

$$n_0 = \mathbf{J}^+ \cdot n_{\text{adj}} \;\underset{\sigma}{\to}\; \mathbf{J}^+ \cdot M_{c_{\text{adj}},\sigma} \cdot n_{\text{adj}} \,, \tag{A.9}$$

which we want to show is the same as $\mathbf{J}^+ \cdot n_{\text{adj}} \overset{?}{\underset{\sigma}{\to}} M_{c_0,\sigma} \cdot \mathbf{J}^+ \cdot n_{\text{adj}}$. We will prove this by showing that $[\mathbf{J}^+\mathbf{J}, M_{c_0,\sigma}] = 0$.

Each row and each column of the matrix $M_{c_0,\sigma}$ contains exactly one 1 and the rest 0's. Such matrices are orthogonal, $M^{\mathsf{T}}_{c_0,\sigma} = M^{-1}_{c_0,\sigma}$. In addition, any permutation can be written as a product of transpositions (the interchange of two labels, $\sigma = a \leftrightarrow b$). Therefore, it will be enough to prove that $[\mathbf{J}^+\mathbf{J}, M_{c_0,a\leftrightarrow b}] = 0$. Performing the same transposition twice is the same as doing nothing. That is, transpositions are their own inverse,

$$M_{c_0,a\leftrightarrow b} = M^{-1}_{c_0,a\leftrightarrow b} = M^{\mathsf{T}}_{c_0,a\leftrightarrow b} \,. \tag{A.10}$$

As in Eq. (3.5), $\mathbf{J}$ can be decomposed as $\mathbf{J} = \mathbf{A} \cdot \mathbf{B}$, where $\mathbf{B}$ is the matrix that relates c_0 to some choice of DDM basis, $c_{\text{DDM}} = \mathbf{B} \cdot c_0$. This implies

$$\mathbf{J}^+ \cdot \mathbf{J} = \mathbf{B}^{\mathsf{T}} \cdot (\mathbf{B} \cdot \mathbf{B}^{\mathsf{T}})^{-1} \cdot \mathbf{B} \,. \tag{A.11}$$

The entries in the DDM basis correspond to all half-ladder graphs with fixed endpoints. If a, b are not those endpoints, then $\sigma = a \leftrightarrow b$ maps entries of the DDM basis to other entries of the same DDM basis, so $c_{\text{DDM}} \underset{\sigma}{\to} M_{c_{\text{DDM}},a\leftrightarrow b} \cdot c_{\text{DDM}}$. Therefore for each transposition there exists a choice of $\mathbf{B}$ such that

$$\begin{aligned} \mathbf{B} \cdot c_0 \underset{\sigma}{\to} M_{c_{\text{DDM}},a\leftrightarrow b} \cdot \mathbf{B} \cdot c_0 = \mathbf{B} \cdot M_{c_0,a\leftrightarrow b} \cdot c_0 \\ \implies M_{c_{\text{DDM}},a\leftrightarrow b} \cdot \mathbf{B} = \mathbf{B} \cdot M_{c_0,a\leftrightarrow b} \,. \end{aligned} \tag{A.12}$$

Now, using Eq. (A.10) and Eq. (A.12), it is straightforward to show that

$$M_{c_0,a\leftrightarrow b} \cdot \mathbf{B}^{\mathsf{T}} \cdot (\mathbf{B} \cdot \mathbf{B}^{\mathsf{T}})^{-1} \cdot \mathbf{B} = \mathbf{B}^{\mathsf{T}} \cdot (\mathbf{B} \cdot \mathbf{B}^{\mathsf{T}})^{-1} \cdot \mathbf{B} \cdot M_{c_0,a\leftrightarrow b} \,. \tag{A.13}$$

k	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
4-pt	1	2	2	3	3	4	4	5	5	6	6	7	7	8	8
5-pt	0	0	2	4	10	16	28	42	64	90	126	168	224	288	370
6-pt	2	8	22	58	133	298	600	1166	2132	3754	6324	10351	16368	25266	38004

Table 2. Counting of scalar kinematic numerator seeds up to 6-point and at $\mathcal{O}(1/\Lambda^{2k})$ for $k \leq 15$ in the EFT expansion. The Gram determinant constraints relevant in 4 spacetime dimensions have been accounted for.

Therefore

$$\begin{aligned} M_{c_0,\sigma} \cdot \mathbf{J}^+ \cdot \mathbf{J} = \mathbf{J}^+ \cdot \mathbf{J} \cdot M_{c_0,\sigma} &= \mathbf{J}^+ \cdot M_{c_{\text{adj}},\sigma} \cdot \mathbf{J} \\ \implies \mathbf{J}^+ \cdot M_{c_{\text{adj}},\sigma} \cdot n_{\text{adj}} &= \mathbf{J}^+ \cdot M_{c_{\text{adj}},\sigma} \cdot \mathbf{J} \cdot \mathbf{J}^+ \cdot n_{\text{adj}} \\ &= M_{c_0,\sigma} \cdot \mathbf{J}^+ \cdot \mathbf{J} \cdot \mathbf{J}^+ \cdot n_{\text{adj}} \\ &= M_{c_0,\sigma} \cdot \mathbf{J}^+ \cdot n_{\text{adj}} \,. \end{aligned} \tag{A.14}$$

Thus, as we wanted to show, the numerator seed $n_0 = \mathbf{J}^+ \cdot n_{\text{adj}}$ (which exists for any adjoint numerator) transforms as

$$n_0 \underset{\sigma}{\rightarrow} M_{c_0,\sigma} \cdot n_0 \,. \tag{A.15}$$

B Examples of seeds and adjoint numerators

In this appendix, we present the construction of kinematic adjoint numerators from numerator seeds. In general, a numerator seed can be obtained from any function $f(1, 2, ..., n)$ through

$$n_0(1, 2, ..., n) = \big(f(1, 2, ..., n) + (-1)^n f(n, ..., 2, 1)\big) + \text{cyclic} \,. \tag{B.1}$$

The number of independent numerator seeds built using this equation, up to 6-point and dimension 30, is provided in Table 2. After multiplying them by $\mathbf{J}$, one needs to explicitly verify the linear independence of the resulting adjoint numerators. To determine the counting provided in Table 1, we first construct vectors by evaluating the numerators numerically for different values of the momenta. The rank of the matrix formed with these vectors as columns is the number of independent adjoint numerators.

A systematic correspondence can be established between the entries of Table 1 and Table 2. At 5-point, the number of independent numerators is, for instance, exactly half of that of independent seeds. This can be understood from the algebraic properties of color factors decomposed into structure constants f^{abc} and symmetric $d^{abc\cdots}$, as discussed in [35] up to 5-point. The 4- and 5-point cases are detailed below.

4-point. At 4-point, the numerator seeds are given by Eq. (4.11). At lowest orders, we have

$$\begin{aligned} n_0(1, 3, 2, 4) = s &\implies n_{\text{adj},s} = t - u \,, \\ n_0(1, 3, 2, 4) = s^2 &\implies n_{\text{adj},s} = t^2 - u^2 = s(u - t) \,. \end{aligned} \tag{B.2}$$

The other valid numerator seed at second order, $n_0(1,2,3,4) = t\,u$ maps to the same adjoint numerator, because the permutation invariant $s^2 + t^2 + u^2 = 2(s^2 - t\,u)$ maps to zero. At third order (and any higher order), the adjoint numerators are permutation invariant functions multiplying the lowest two orders [34]:

$$n_0(1,3,2,4) = s^3 \implies n_{\text{adj},s} = t^3 - u^3 = (t-u)(s^2+t^2+u^2)/2\,. \tag{B.3}$$

Upon comparing Table 1 and Table 2, one can identify the difference between the numbers of independent seeds and numerators at 4-point as being the number of permutation-invariant functions of the Mandelstam invariants. The latter are combinations of terms of the form $(s^2+t^2+u^2)^m(stu)^n$ for integer m, n. Indeed, the seeds at 4-point are symmetric under reversal symmetry, hence they capture all expressions which are combinations of reversal-symmetric terms of the form

$$f(\sigma(a),\sigma(b),\sigma(c),\sigma(d)) + f(\sigma(d),\sigma(c),\sigma(b),\sigma(a)) \tag{B.4}$$

for some permutation σ and some function f, consistently with Eq. (B.1). They do however not capture expressions that are antisymmetric under the reversal of particle labels. To understand what algebraic structures appear in the symmetric combinations, one can consider the permutation properties of color structures generated by single traces of group generators, which are identical to those of $f(1,2,...,n)+$cyclic in Eq. (B.1). At 4-point, the relevant color structures are $d^{abx}f^{xcd}$, $f^{abx}f^{xcd}$ and d^{abcd} [35]. The first one is expressed in terms of antisymmetric combinations of traces only, whereas the other two are expressed in terms of symmetric combinations. Therefore, the seeds generate expressions with the same algebraic properties as the two last color structures, which are respectively adjoint and permutation-invariant.

5-point. At the lowest two orders in the Mandelstam invariants, there are no numerator seeds. For example, the candidate function $f(1,2,3,4,5) = s_{12} + s_{23} + s_{34} + s_{45} + s_{51}$ is invariant under cyclic permutations, but symmetric instead of antisymmetric under reversing the order of its arguments, therefore it is mapped to zero under the action of $\mathbf{J}$. At third order in the Mandelstams, there exists

$$n_0(1,2,3,4,5) = \frac{1}{2}(s_{12}^2\, s_{34} - s_{45}^2\, s_{23}) + \text{cyclic}\,, \tag{B.5}$$

which is mapped to the adjoint numerator

$$\begin{aligned} n_{\text{adj},s_{12}s_{45}} = \;& \big(n_0(1,2,3,4,5) - n_0(1,5,4,3,2)\big) + \big(n_0(1,4,5,3,2) - n_0(1,2,3,5,4)\big) \\ &+\big(n_0(1,2,5,4,3) - n_0(1,3,4,5,2)\big) + \big(n_0(1,3,5,4,2) - n_0(1,2,4,5,3)\big) \\ =\;& (s_{24}s_{35}^2 - s_{24}s_{13}^2) + \big(s_{24}^2 s_{13} - s_{24}^2 s_{35}\big) + \big(s_{25}s_{13}^2 - s_{14}s_{35}^2\big) \\ &+\big(s_{14}^2 s_{35} - s_{25}^2 s_{13}\big) + \big(s_{14}s_{23}^2 - s_{25}s_{34}^2\big) + \big(s_{25}^2 s_{34} - s_{14}^2 s_{23}\big) \\ &+\big(s_{15}s_{34}^2 - s_{15}s_{23}^2\big) + \big(s_{15}^2 s_{23} - s_{15}^2 s_{34}\big)\,. \end{aligned} \tag{B.6}$$

Here, the subscript $s_{12}s_{45}$ indicates the pole structure of the trivalent graph to which this entry belongs, i.e. $c_{\text{adj},s_{12}s_{45}} = f^{12x}f^{x3y}f^{y45}$. The first line in Eq. (B.6) is analogous to

$$\begin{aligned} f^{12x}f^{x3y}f^{y45} = \ & \big((12345) - (15432)\big) + \big((14532) - (12354)\big) \\ & + \big((12543) - (13452)\big) + \big((13542) - (12453)\big)\,. \end{aligned} \tag{B.7}$$

(recall $(12...5) \equiv \text{Tr}(T^{a_1}T^{a_2}...T^{a_5})$, etc.).

The comparison between Table 1 and Table 2 shows that the number of independent adjoint numerators is exactly half that of independent seeds. As in the 4-point case, this can be understood from the algebraic properties of the color factors which are generated by combinations of single traces of group generators with the same behavior under reversal symmetry as the seeds. At 5-point, the seeds are antisymmetric under reversal, while the color factors which can be decomposed onto antisymmetric combinations of single traces can be identified from the classification in [35]. In the language of this reference, they correspond to adjoint and hybrid structures, which correspond to combinations of color factors of the form $f^{abx}f^{xcy}f^{yde}$ and $d^{abcx}f^{xde}$. Therefore, antisymmetric seeds generate all expressions having the algebraic properties of these two kinds. In addition, there exists a bijection between adjoint and hybrid structures [35]. Namely, for each expression with adjoint properties, there exists one with hybrid properties, and reciprocally. It then follows that the number of independent seeds is exactly twice that of independent adjoint numerators.

6-point. At lowest order, there exist two functional forms that are cyclically invariant and symmetric under argument reversal,

$$\begin{aligned} n_0(1,2,3,4,5,6) &= s_{12} + s_{23} + s_{34} + s_{45} + s_{56} + s_{16}\,, \\ n_0(1,2,3,4,5,6) &= s_{123} + s_{234} + s_{345}\,, \end{aligned} \tag{B.8}$$

where $s_{abc} = (p_a + p_b + p_c)^2$. These map to one independent adjoint numerator after multiplication by $\mathbf{J}$. At second order, there are three independent adjoint numerators. Out of the eight independent seeds, these can for instance be constructed from

$$\begin{aligned} n_0(1,2,3,4,5,6) &= s_{123}\,(s_{234} + s_{345}) + s_{234}\,s_{345}\,, \\ n_0(1,2,3,4,5,6) &= s_{16}s_{34} + s_{12}s_{45} + s_{23}s_{56}\,, \\ n_0(1,2,3,4,5,6) &= s_{123}(s_{16} + s_{34}) + s_{345}(s_{23} + s_{56}) + s_{234}(s_{12} + s_{45})\,. \end{aligned} \tag{B.9}$$

The adjoint numerator then has the form

$$\begin{aligned} n_{\text{adj},s_{12}s_{34}s_{56}} = \big[& n_0(1,2,3,4,5,6) - n_0(1,2,3,4,6,5) - n_0(1,2,4,3,5,6) + n_0(1,2,4,3,6,5) \\ & - n_0(1,2,5,6,3,4) + n_0(1,2,5,6,4,3) + n_0(1,2,6,5,3,4) - n_0(1,2,6,5,4,3)\big] \\ & + \text{reversed orderings}\,, \end{aligned} \tag{B.10}$$

for the adjoint entry corresponding to $c_{\text{adj},s_{12}s_{34}s_{56}} = f^{12x}f^{34y}f^{56z}f^{xyz}$, and

$$n_{\text{adj},s_{12}s_{123}s_{56}} = \big[n_0(1,2,3,4,5,6) - n_0(1,2,3,4,6,5) - n_0(1,2,3,5,6,4) + n_0(1,2,3,6,5,4)$$

$$
\begin{aligned}
&- n_0(1,2,4,5,6,3) + n_0(1,2,4,6,5,3) + n_0(1,2,5,6,4,3) - n_0(1,2,6,5,4,3)\big] \\
&+ \text{reversed orderings}\,, \qquad \text{(B.11)}
\end{aligned}
$$

for the adjoint entry corresponding to $c_{\text{adj},s_{12}s_{123}s_{56}} = f^{12x} f^{x3y} f^{y4z} f^{z56}$. At third order, there are nine independent adjoint numerators. They can be constructed from, for example,

$$
\begin{aligned}
n_0(1,2,3,4,5,6) &= s_{12}s_{34}s_{56} + s_{23}s_{45}s_{16}\,, \qquad \text{(B.12)} \\
n_0(1,2,3,4,5,6) &= s_{123}s_{234}s_{345}\,, \\
n_0(1,2,3,4,5,6) &= s_{12}s_{23}s_{34} + s_{12}s_{23}s_{16} + s_{23}s_{34}s_{45} + s_{12}s_{16}s_{56} + s_{16}s_{45}s_{56} + s_{34}s_{45}s_{56}\,, \\
n_0(1,2,3,4,5,6) &= (s_{56}s_{12}^2 + s_{34}^2 s_{12} + s_{16}s_{23}^2 + s_{23}s_{45}^2 + s_{34}s_{56}^2 + s_{16}^2 s_{45}) \\
&\qquad + (1 \leftrightarrow 6,\ 2 \leftrightarrow 5,\ 3 \leftrightarrow 4)\,, \\
n_0(1,2,3,4,5,6) &= s_{234}(s_{12}^2 + s_{45}^2) + s_{123}(s_{16}^2 + s_{34}^2) + s_{345}(s_{23}^2 + s_{56}^2)\,, \\
n_0(1,2,3,4,5,6) &= s_{123}s_{345}(s_{12} + s_{45}) + s_{123}s_{234}(s_{23} + s_{56}) + s_{234}s_{345}(s_{16} + s_{34})\,, \\
n_0(1,2,3,4,5,6) &= s_{123}^2\,(s_{234} + s_{345}) + s_{234}^2\,(s_{123} + s_{345}) + s_{345}^2\,(s_{123} + s_{234})\,, \\
n_0(1,2,3,4,5,6) &= (s_{123} + s_{234} + s_{345})^{\,3}\,, \\
n_0(1,2,3,4,5,6) &= (s_{12} + s_{23} + s_{34} + s_{45} + s_{56} + s_{16})^{\,3}\,.
\end{aligned}
$$

We leave the detailed comparison between the entries of Table 1 and Table 2 at 6-point (and higher) to future work.

C Details at 5-point

C.1 Matrices

At 5-point, for the trace ordering given by

$$
\begin{aligned}
c_0 = \big(&(12345),(12354),(12435),(12453),(12534),(12543),(13245),(13254), \\
&(13425),(13524),(14235),(14325),(15432),(14532),(15342),(13542), \\
&(14352),(13452),(15423),(14523),(15243),(14253),(15324),(15234)\big)^{\mathsf{T}}\,,
\end{aligned} \qquad \text{(C.1)}
$$

and the adjoint color numerators by

$$
\begin{aligned}
c_{\text{adj}} = \Big(&\, f^{12x} f^{34y} f^{5xy}, f^{12x} f^{35y} f^{4xy}, f^{12x} f^{45y} f^{3xy}, f^{13x} f^{24y} f^{5xy}, \\
&f^{13x} f^{25y} f^{4xy}, f^{13x} f^{45y} f^{2xy}, f^{14x} f^{23y} f^{5xy}, f^{14x} f^{25y} f^{3xy}, \\
&f^{14x} f^{35y} f^{2xy}, f^{15x} f^{23y} f^{4xy}, f^{15x} f^{24y} f^{3xy}, f^{15x} f^{34y} f^{2xy}, \\
&f^{23x} f^{45y} f^{1xy}, f^{24x} f^{35y} f^{1xy}, f^{25x} f^{34y} f^{1xy}\,\Big)^{\mathsf{T}},
\end{aligned} \qquad \text{(C.2)}
$$

the matrix $\mathbf{J}_5 = \left(\mathbf{J}_{15\times 12}, -\mathbf{J}_{15\times 12}\right)^{\mathsf{T}}$ with

$$\mathbf{J}_{15\times 12} = \begin{pmatrix}
1 & 0 & -1 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & -1 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-1 & 1 & 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & -1 & -1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -1 & 0 & 1 & -1 & -1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & -1 & -1 & 1 & 0 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & -1 \\
0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & -1 & 0 & -1 \\
0 & 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & -1 & -1 & 0 \\
-1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & -1 \\
0 & 0 & -1 & 0 & 0 & 0 & 1 & 0 & -1 & 0 & 1 & 0 \\
1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 \\
1 & -1 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & 0 \\
0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & -1 & 0 & 0 & 1
\end{pmatrix}. \tag{C.3}$$

A block structure in $\mathbf{J}_5$ has been made manifest by ordering c_0 into the schematic form: $c_0 = (c_{012}, \text{reversed}(c_{012}))^{\mathsf{T}}$, where c_{012} contains 12 entries of c_0 that are not related by reversing the order of the particle labels.

The matrix $\mathbf{G}_2$, which can be obtained from the numerator seed Eq. (6.7) by stripping of the single traces, for $\tilde{g}_{12345} \equiv g_2(1,2,3,4,5)$, is

$$\mathbf{G}_2 = \begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \tilde{g}_{12345} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \tilde{g}_{12354} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \tilde{g}_{12435} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \tilde{g}_{12453} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \tilde{g}_{12534} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \tilde{g}_{12543} & 0 \\
0 & 0 & 0 & 0 & \tilde{g}_{13245} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \tilde{g}_{13254} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \tilde{g}_{13425} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \tilde{g}_{13524} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \tilde{g}_{14235} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \tilde{g}_{14325} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \tilde{g}_{15432} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \tilde{g}_{14532} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \tilde{g}_{15342} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \tilde{g}_{13542} \\
0 & 0 & 0 & 0 & 0 & 0 & \tilde{g}_{14352} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \tilde{g}_{13452} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \tilde{g}_{15423} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \tilde{g}_{14523} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \tilde{g}_{15243} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\tilde{g}_{14253} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \tilde{g}_{15324} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \tilde{g}_{15234} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}. \tag{C.4}$$

C.2 Solution of the KLT bootstrap

We confirmed explicitly that the solution to the KLT bootstrap at 5-point is reproduced by Eq. (6.11), with

$$\begin{aligned}
g_{1,\mathrm{L}}(1,2,3,4,5)\times m[12345|12345] &= 3\,\frac{e_1}{\Lambda^4} - \frac{e_1}{\Lambda^4}\left(\frac{s_{12}}{s_{34}}+\frac{s_{12}}{s_{45}}+\text{cyclic}\right)\\
&\quad + \frac{e_2}{\Lambda^6}\left(-2\,s_{12}+\frac{s_{12}^2}{s_{34}}+\frac{s_{12}^2}{s_{45}}+2\,\frac{s_{12}s_{23}}{s_{45}}+\text{cyclic}\right)\\
g_{2,\mathrm{L}}(1,2,3,4,5)\times m[12345|12345] &= \frac{e_1}{\Lambda^4} - \frac{e_2}{\Lambda^6}\,(s_{12}+\text{cyclic})\\
g_{1,\mathrm{R}}(1,2,3,4,5)\times m[12345|12345] &= 3\,\frac{e_3}{\Lambda^4} - \frac{e_3}{\Lambda^4}\left(\frac{s_{12}}{s_{34}}+\frac{s_{12}}{s_{45}}+\text{cyclic}\right)\\
g_{2,\mathrm{R}}(1,2,3,4,5)\times m[12345|12345] &= \frac{e_3}{\Lambda^4}
\end{aligned} \tag{C.5}$$

where

$$m[12345|12345] = \frac{1}{s_{12}s_{34}} + \text{cyclic}\,. \tag{C.6}$$

and e_i are free parameters, which reproduce the results of [36] when $e_1 = a_{1,0} - a_{1,1}$, $e_2 = a_{2,0}$, and $e_3 = a_{1,1}$. Note that the e_2 parameter could have equivalently been part of $g_{1,\mathrm{R}}$ and $g_{2,\mathrm{R}}$ without changing the amplitudes.

C.3 Generalized KLT kernel from seeds

At 4-point, the generalized kernel for any choice of BCJ bases was obtained by multiplying the left and right side of the traditional KLT kernel by the diagonal matrices $\mathbf{G}^{\text{h.d.}}_{\mathrm{L/R}}$ for the *same* BCJ bases, see Eq. (5.6). This simple structure extends to 5-point only for a particular BCJ bases, as the structure gets more involved due the presence of the non-diagonal matrix $\mathbf{G}^{\text{h.d.}}_2$ in Eq. (6.11). This is not necessarily a problem as one is always allowed to choose a particular basis to compute double-copy amplitudes.

For example, for the usual biadjoint scalar theory, the sub-matrix of ordered amplitudes for $\alpha,\beta\in\{12345, 13524\}$ reads

$$m_5[\alpha|\beta] = \begin{pmatrix} m_5[12345|12345] & 0 \\ 0 & m_5[13524|13524] \end{pmatrix}, \tag{C.7}$$

where $m_5[12345|13524] = m_5[13524|12345] = 0$ because the color orderings 12345 and 13524 do not share any poles. The associated kernel is then simply

$$S_5[\alpha|\beta] = \begin{pmatrix} 1/m_5[12345|12345] & 0 \\ 0 & 1/m_5[13524|13524] \end{pmatrix}. \tag{C.8}$$

It turns out that $\mathbf{m}^{\text{h.d.}}_5$ is also simplified when restricted to the BCJ basis $\{12345, 13524\}$. For any color ordering a, and δ restricted to $\{12345, 13524\}$, Eq. (6.11) becomes

$$m^{\text{h.d.}}_5[a|\delta] = \sum_{b,c}^{24} (G_{1,\mathrm{R}}+G_{2,\mathrm{R}})[a|b]\;\; m_5[b|c]\;\; (G_{1,\mathrm{L}}+G_{2,\mathrm{L}})[c|\delta]\,. \tag{C.9}$$

This can be simplified by noting that $(G_{1,\mathrm{L}} + G_{2,\mathrm{L}})[c|12345]$ is nonzero only for $c = \{12345, 14253\}$, while $(G_{1,\mathrm{L}} + G_{2,\mathrm{L}})[c|13524]$ is nonzero only for $c = \{12345, 13524\}$. Moreover, $m_5[\alpha|14253] = -m_5[\alpha|13524]$ for any α, which means that the second index of m_5 can be limited to the BCJ basis $\{12345, 13524\}$. This means that we can restrict to the sub-matrix (appearing in Eq. (6.14))

$$\mathcal{G}_{\mathrm{L}} \equiv \begin{pmatrix} (G_{1,\mathrm{L}} + G_{2,\mathrm{L}})[12345|12345] & (G_{1,\mathrm{L}} + G_{2,\mathrm{L}})[12345|13524] \\ -(G_{1,\mathrm{L}} + G_{2,\mathrm{L}})[14253|12345] & (G_{1,\mathrm{L}} + G_{2,\mathrm{L}})[13524|13524] \end{pmatrix} , \tag{C.10}$$

and similarly for the right side with

$$\mathcal{G}_{\mathrm{R}} \equiv \begin{pmatrix} (G_{1,\mathrm{R}} + G_{2,\mathrm{R}})[12345|12345] & -(G_{1,\mathrm{R}} + G_{2,\mathrm{R}})[12345|14253] \\ (G_{1,\mathrm{R}} + G_{2,\mathrm{R}})[14253|12345] & (G_{1,\mathrm{L}} + G_{2,\mathrm{R}})[13524|13524] \end{pmatrix} , \tag{C.11}$$

We can then write the BAS amplitudes for this specific BCJ basis as

$$m_5^{\mathrm{h.d.}}[\alpha|\delta] = \sum_{\beta,\gamma\in\{12345,13524\}} \mathcal{G}_{\mathrm{R}}[\alpha|\beta] \cdot m_5[\beta|\gamma] \cdot \mathcal{G}_{\mathrm{L}}[\gamma|\delta] , \tag{C.12}$$

Therefore, in this case, it is simple to invert the generalized BAS matrix to obtain the generalized KLT kernel:

$$S_5^{\mathrm{h.d.}}[\alpha|\delta] = \sum_{\beta,\gamma\in\{12345,13524\}} \mathcal{G}_{\mathrm{L}}^{-1}[\alpha|\beta]\, S[\beta|\gamma]\, \mathcal{G}_{\mathrm{R}}^{-1}[\gamma|\delta] . \tag{C.13}$$

We let the investigation of similar formula for a generic BCJ basis for a future work.

References

[1] H. Kawai, D.C. Lewellen and S.H.H. Tye, *A Relation Between Tree Amplitudes of Closed and Open Strings*, *Nucl. Phys. B* **269** (1986) 1.

[2] Z. Bern, J.J.M. Carrasco and H. Johansson, *New Relations for Gauge-Theory Amplitudes*, *Phys. Rev. D* **78** (2008) 085011 [0805.3993].

[3] Z. Bern, J.J.M. Carrasco and H. Johansson, *Perturbative Quantum Gravity as a Double Copy of Gauge Theory*, *Phys. Rev. Lett.* **105** (2010) 061602 [1004.0476].

[4] F. Cachazo, S. He and E.Y. Yuan, *Scattering equations and Kawai-Lewellen-Tye orthogonality*, *Phys. Rev. D* **90** (2014) 065001 [1306.6575].

[5] F. Cachazo, S. He and E.Y. Yuan, *Scattering of Massless Particles in Arbitrary Dimensions*, *Phys. Rev. Lett.* **113** (2014) 171601 [1307.2199].

[6] F. Cachazo, S. He and E.Y. Yuan, *Scattering of Massless Particles: Scalars, Gluons and Gravitons*, *JHEP* **07** (2014) 033 [1309.0885].

[7] J.J.M. Carrasco and L. Rodina, *UV considerations on scattering amplitudes in a web of theories*, *Phys. Rev. D* **100** (2019) 125007 [1908.08033].

[8] H. Johansson and A. Ochirov, *Pure Gravities via Color-Kinematics Duality for Fundamental Matter*, *JHEP* **11** (2015) 046 [1407.4772].

[9] H. Johansson and A. Ochirov, *Color-Kinematics Duality for QCD Amplitudes*, *JHEP* **01** (2016) 170 [1507.00332].

[10] L. de la Cruz, A. Kniss and S. Weinzierl, *Double Copies of Fermions as Matter that Interacts Only Gravitationally*, *Phys. Rev. Lett.* **116** (2016) 201601 [1601.04523].

[11] R.W. Brown and S.G. Naculich, *Color-factor symmetry and BCJ relations for QCD amplitudes*, *JHEP* **11** (2016) 060 [1608.05291].

[12] R.W. Brown and S.G. Naculich, *KLT-type relations for QCD and bicolor amplitudes from color-factor symmetry*, *JHEP* **03** (2018) 057 [1802.01620].

[13] H. Johansson and A. Ochirov, *Double copy for massive quantum particles with spin*, *JHEP* **09** (2019) 040 [1906.12292].

[14] J. Plefka, C. Shi and T. Wang, *Double copy of massive scalar QCD*, *Phys. Rev. D* **101** (2020) 066004 [1911.06785].

[15] M. Chiodaroli, M. Gunaydin, H. Johansson and R. Roiban, *Spontaneously Broken Yang-Mills-Einstein Supergravities as Double Copies*, *JHEP* **06** (2017) 064 [1511.01740].

[16] L.A. Johnson, C.R.T. Jones and S. Paranjape, *Constraints on a Massive Double-Copy and Applications to Massive Gravity*, *JHEP* **02** (2021) 148 [2004.12948].

[17] A. Momeni, J. Rumbutis and A.J. Tolley, *Massive Gravity from Double Copy*, *JHEP* **12** (2020) 030 [2004.07853].

[18] A. Momeni, J. Rumbutis and A.J. Tolley, *Kaluza-Klein from colour-kinematics duality for massive fields*, *JHEP* **08** (2021) 081 [2012.09711].

[19] Y.-F. Hang and H.-J. He, *Structure of Kaluza-Klein graviton scattering amplitudes from the gravitational equivalence theorem and double copy*, *Phys. Rev. D* **105** (2022) 084005 [2106.04568].

[20] Y. Li, Y.-F. Hang, H.-J. He and S. He, *Scattering amplitudes of Kaluza-Klein strings and extended massive double-copy*, *JHEP* **02** (2022) 120 [2111.12042].

[21] G. Chen and Y.-J. Du, *Amplitude Relations in Non-linear Sigma Model*, *JHEP* **01** (2014) 061 [1311.1133].

[22] F. Cachazo, S. He and E.Y. Yuan, *Scattering Equations and Matrices: From Einstein To Yang-Mills, DBI and NLSM*, *JHEP* **07** (2015) 149 [1412.3479].

[23] Z. Bern, J.J. Carrasco, M. Chiodaroli, H. Johansson and R. Roiban, *The Duality Between Color and Kinematics and its Applications*, 1909.01358.

[24] J. Broedel and L.J. Dixon, *Color-kinematics duality and double-copy construction for amplitudes from higher-dimension operators*, *JHEP* **10** (2012) 091 [1208.0876].

[25] G. Menezes, *Color-kinematics duality, double copy and the unitarity method for higher-derivative QCD and quadratic gravity*, *JHEP* **03** (2022) 074 [2112.00978].

[26] H. Elvang, M. Hadjiantonis, C.R.T. Jones and S. Paranjape, *Soft Bootstrap and Supersymmetry*, *JHEP* **01** (2019) 195 [1806.06079].

[27] H. Elvang, M. Hadjiantonis, C.R.T. Jones and S. Paranjape, *Electromagnetic Duality and D3-Brane Scattering Amplitudes Beyond Leading Order*, *JHEP* **04** (2021) 173 [2006.08928].

[28] M. Carrillo González, R. Penco and M. Trodden, *Shift symmetries, soft limits, and the double copy beyond leading order*, *Phys. Rev. D* **102** (2020) 105011 [1908.07531].

[29] K. Kampf, *The ChPT: top-down and bottom-up*, *JHEP* **12** (2021) 140 [2109.11574].

[30] I. Low and Z. Yin, *New Flavor-Kinematics Dualities and Extensions of Nonlinear Sigma Models*, *Phys. Lett. B* **807** (2020) 135544 [1911.08490].

[31] I. Low, L. Rodina and Z. Yin, *Double Copy in Higher Derivative Operators of Nambu-Goldstone Bosons*, *Phys. Rev. D* **103** (2021) 025004 [2009.00008].

[32] A. Brandhuber, G. Chen, G. Travaglini and C. Wen, *A new gauge-invariant double copy for heavy-mass effective theory*, *JHEP* **07** (2021) 047 [2104.11206].

[33] K. Haddad and A. Helset, *The double copy for heavy particles*, *Phys. Rev. Lett.* **125** (2020) 181603 [2005.13897].

[34] J.J.M. Carrasco, L. Rodina, Z. Yin and S. Zekioglu, *Simple encoding of higher derivative gauge and gravity counterterms*, *Phys. Rev. Lett.* **125** (2020) 251602 [1910.12850].

[35] J.J.M. Carrasco, L. Rodina and S. Zekioglu, *Composing effective prediction at five points*, *JHEP* **06** (2021) 169 [2104.08370].

[36] H.-H. Chi, H. Elvang, A. Herderschee, C.R.T. Jones and S. Paranjape, *Generalizations of the double-copy: the KLT bootstrap*, *JHEP* **03** (2022) 077 [2106.12600].

[37] S. Mizera, *Inverse of the String Theory KLT Kernel*, *JHEP* **06** (2017) 084 [1610.04230].

[38] N.E.J. Bjerrum-Bohr, P.H. Damgaard, T. Sondergaard and P. Vanhove, *The Momentum Kernel of Gauge and Gravity Theories*, *JHEP* **01** (2011) 001 [1010.3933].

[39] M. Kiermaier, *Gravity as the square of gauge theory*, presentation at *"Amplitudes 2010"* (7 May 2010).

[40] Z. Bern and T. Dennen, *A Color Dual Form for Gauge-Theory Amplitudes*, *Phys. Rev. Lett.* **107** (2011) 081601 [1103.0312].

[41] N.E.J. Bjerrum-Bohr, P.H. Damgaard, R. Monteiro and D. O'Connell, *Algebras for Amplitudes*, *JHEP* **06** (2012) 061 [1203.0944].

[42] Y.-J. Du, B. Feng and C.-H. Fu, *The Construction of Dual-trace Factor in Yang-Mills Theory*, *JHEP* **07** (2013) 057 [1304.2978].

[43] C.-H. Fu, Y.-J. Du and B. Feng, *Note on Construction of Dual-trace Factor in Yang-Mills Theory*, *JHEP* **10** (2013) 069 [1305.2996].

[44] S.G. Naculich, *Scattering equations and virtuous kinematic numerators and dual-trace functions*, *JHEP* **07** (2014) 143 [1404.7141].

[45] A. Brandhuber, G. Chen, H. Johansson, G. Travaglini and C. Wen, *Kinematic Hopf Algebra for Bern-Carrasco-Johansson Numerators in Heavy-Mass Effective Field Theory and Yang-Mills Theory*, *Phys. Rev. Lett.* **128** (2022) 121601 [2111.15649].

[46] R. Kleiss and H. Kuijf, *Multigluon cross sections and 5-jet production at hadron colliders*, *Nuclear Physics B* **312** (1989) 616.

[47] H. Elvang, *Effective field theories and the double-copy*, presentation at *"QCD meets Gravity 2021"* (15 December 2021).

[48] H. Elvang and Y.-t. Huang, *Scattering Amplitudes*, 1308.1697.

[49] C. Cheung, *TASI Lectures on Scattering Amplitudes*, in *Proceedings, Theoretical Advanced Study Institute in Elementary Particle Physics : Anticipating the Next Discoveries in*

Particle Physics (TASI 2016): Boulder, CO, USA, June 6-July 1, 2016, R. Essig and I. Low, eds., pp. 571–623 (2018), DOI [1708.03872].

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

[51] D. de Neeling, D. Roest and S. Veldmeijer, *Flavour-kinematic duality for Goldstone modes*, 2204.11629.

[52] R.H. Boels and R. Medina, *Graviton and gluon scattering from first principles*, *Phys. Rev. Lett.* **118** (2017) 061602 [1607.08246].

[53] Z. Bern, A. Edison, D. Kosower and J. Parra-Martinez, *Curvature-squared multiplets, evanescent effects, and the U(1) anomaly in $N = 4$ supergravity*, *Phys. Rev. D* **96** (2017) 066004 [1706.01486].

[54] A. Ben-Israel and T.N. Greville, *Generalized Inverses: Theory and Applications*, Wiley, New York (1980).