

The double copy effective action: a quantum (chromodynamics) approach to space-time

John Joseph M. Carrasco^{a,b} Suna Zekioglu^a

^a*The Amplitudes and Insights Group, Department of Physics & Astronomy, Northwestern University, Evanston, Illinois 60208, USA*

^b*Center for Interdisciplinary Exploration and Research in Astrophysics (CIERA), Northwestern University, 1800 Sherman Ave, Evanston, IL 60201, USA*

ABSTRACT: Conventional Lagrangian formulations of gauge and gravity theories emphasize compactness and off-shell symmetry. This often obscures the structure of on-shell physical observables. In this work, we present a constructive framework that elevates gauge-invariant scattering amplitudes to the defining data for quantum field theory actions, including effective field theories. Focusing on double-copy theories, we promote color-dual amplitude numerators to quantum operators. This enables the systematic identification of novel local operator content at each multiplicity and the construction of double-copy-compatible actions. By applying this framework to the well-established double-copy relationship between Einstein gravity and Yang-Mills theory, which holds for all-multiplicity tree-level amplitudes, we demonstrate a systematic path to constructing the operator expansion of $\sqrt{-g}R$ from factorized gauge-theory components. This clarifies how gravitational interactions can be understood as emerging from simpler gauge-theoretic structures at the action level. This formalism extends color-kinematics duality from amplitude data to operator constructions, naturally realizing the double copy at the level of actions and asymptotic quantum states. We illustrate the method with Yang-Mills theory, Einstein gravity, and its application to generating higher-derivative operators inspired by Z-theory and open superstring amplitudes. This work provides a concrete bridge between structured amplitudes and effective actions, offering a physically grounded alternative to traditional EFT basis-building. It reveals at the operator level deep structural connections between gauge theory and gravity (connections long recognized in scattering amplitudes) from fundamental interactions to their quantum state descriptions and higher-derivative extensions.

Contents

1	Introduction	3
2	Review	6
2.1	Color-dual representations	6
2.2	Higher-derivative compositional bootstrap	9
3	An Invitation: The Simple Scalar	10
3.1	Defining the simple scalar	10
3.1.1	Bootstrapping the theory from physical principles	10
3.1.2	Traditional form of the action	15
3.2	Constructing operators	16
3.2.1	Three field operator	18
3.2.2	Four field operator	20
3.3	Writing the actual Lagrangian	23
3.3.1	Removing redundancy at four points	23
3.4	Summary	29
4	Extracting Novel Contact Terms From Scattering Amplitudes	30
4.1	General Procedure	30
4.2	Five point example	33
4.3	Six point example	37
4.4	More on gravitational contact isolation	39
4.5	Spurious non-locality	41
5	The Operator Promotion	41
5.1	Double-Copy Structure	42
5.2	Operators can look like Amplitudes	44
5.3	Generalized gauge accounting in $\hat{\mathcal{O}}(\mathcal{A}_m)$	48
6	Applications and Examples	49
6.1	Pedagogical Case: Yang-Mills and Einstein-Hilbert Gravity	49
6.2	Higher Derivative Operators from Pointlike Fields to String Theory	55
7	Path to Quantum Gravity via Double Copy	61
7.1	From Yang-Mills amplitudes to the Einstein-Hilbert action	61
7.2	State level description	63

1 Introduction

Recent advances in explicit scattering amplitude constructions [1–3] have shown that it is straightforward to bootstrap color-dual higher-derivative corrections to all orders in mass dimension—at least at low multiplicity. But what local operators should be added to the action to generate these bootstrapped amplitudes?

The conventional route is to postulate field-level ansätze and fix their coefficients by matching to known amplitude data. While this approach can succeed at low multiplicity or when guided by symmetry constraints, it becomes increasingly inefficient and opaque at higher orders. Previous efforts to incorporate color-kinematics duality at the Lagrangian level have often focused on constructing effective Yang-Mills Lagrangians whose Feynman rules are designed to directly yield BCJ-satisfying kinematic numerators for all graph contributions (e.g., [4–6]). Such approaches typically involve adding either auxiliary fields or higher-valency interaction terms to the standard Lagrangian, which may be non-local and are often defined recursively such that the full Lagrangian remains on-shell equivalent to Yang-Mills but is expressed in a form that makes the duality manifest at the level of Feynman rules.

Additionally significant insights into the origins and off-shell validity of color-kinematics duality also come from analyzing specific action principles in various frameworks. Examples include the pure spinor formalism for supersymmetric Yang-Mills theories [7], where dual numerators emerge from the action’s structure, and studies of Chern-Simons theory [8], where kinematic algebras satisfying Jacobi identities can be identified directly from the action, leading to off-shell CK duality for currents and correlators. These formalism-centric approaches further our understanding of how such dualities can be inherent, even off-shell, properties of particular fundamental actions.

Our approach differs significantly in its philosophy and goals. We take well-structured, color-dual (or double-copy) amplitudes as the primary input. Our framework then provides a systematic method to derive the minimal set of local operators in an effective action that are necessary to reproduce these physical observables. We achieve this by sharply distinguishing novel m -point contact interactions from contributions reconstructible via unitarity from lower-multiplicity amplitudes. While color-kinematics duality is a crucial guiding principle that ensures the consistency and structural integrity of our input amplitudes (and is essential for applications like the gravitational double copy), our method does not aim to generate specific numerator forms from Feynman rules. Instead, it translates the physical information encoded in these already-structured numerators directly into local operator terms in the action. This allows us to maintain locality in the fundamental operator additions at each order and provides a direct bridge from on-shell S-matrix data to a standard effective action.

Specifically, we describe a promotion procedure, fig. 1, that maps each graph contribution in an m -point amplitude to a corresponding local m -field operator. This approach preserves color-kinematics duality and, when applicable, maintains manifest double-copy structure. While the method is particularly well-suited to constructing higher-derivative color-dual operators, it also provides a path to rewriting entire double-copy theories in terms of their

color-dual graph basis. This S-matrix-driven approach to constructing actions has important historical precedents. Of related prior work, we find the closest in spirit and approach to be the landmark work of Bern and Grant [9] which, preceding the convenience of color-dual double-copy, still constructed higher-multiplicity gravity amplitudes from Yang-Mills amplitudes via the $\alpha' \rightarrow 0$ KLT relations [10, 11] and subtracted all cuts manually to achieve incredibly compact representations of the Lagrangian level graviton contact terms through five points.

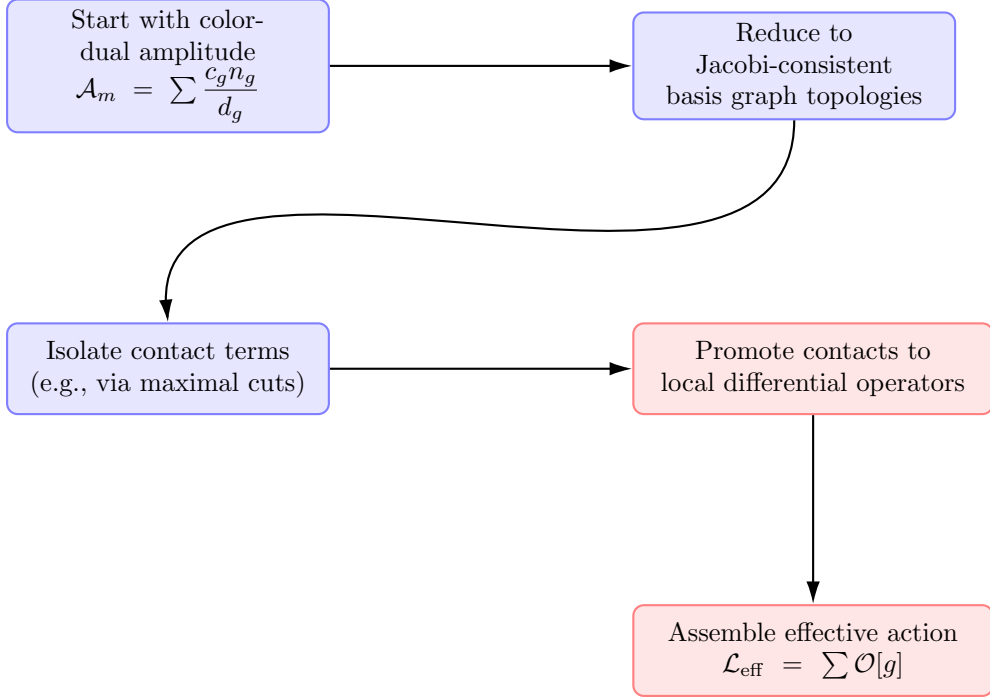


Figure 1: From amplitudes to operators. The operator promotion procedure begins with a color-dual amplitude, reduces to a basis of cubic graphs using Jacobi identities, isolates local contact terms (e.g., via maximal cut 2), promotes each numerator to a field-space operator, and assembles the resulting operators into the effective action. The structure is preserved under double copy.

The operator promotion procedure we present utilizes a systematic method, akin to a generalization of maximal cuts, to isolate novel contact terms from amplitude data at each multiplicity, thereby enforcing locality and removing redundancy from lower-order physics. This not only streamlines operator construction but also makes the double-copy structure of the resulting actions manifest. Operators derived in this way inherit the symmetries of the input amplitudes.

This framework finds a particularly compelling application in relating Yang-Mills theory to Einstein gravity. Given that tree-level gravitational scattering amplitudes are understood to be constructible as a double copy of Yang-Mills amplitudes to all multiplicities,

and that color-dual representations for Yang-Mills numerators are systematically available, our method provides a direct bridge to the corresponding operator structure. Specifically, it furnishes a constructive algorithm for deriving the operator expansion of the Einstein-Hilbert action, $\sqrt{-g}R$, from factorized Yang-Mills building blocks. This process reveals how the complete tower of gravitational interactions can be systematically assembled from gauge-theoretic components at the level of the action. This finds a particular resonance in the class of twofold-symmetry representations of the Einstein-Hilbert action of Cheung and Remmen [12]. Reference [12] presents a well-structured set of gauge transformations and gauge fixing so that ‘left’-indices only contract with ‘left’-indices, and ‘right’-indices only contract with ‘right’-indices. While motivated by double-copy they did not establish ‘left’ and ‘right’ with distinct Yang-Mills-type operators. Here such a structure is clear – although we identify gravitational states in a more traditional gauge involving a symmetrization and projection out of any dilatonic trace, and explicitly use propagators that enforce such a projection to physical graviton states.

Beyond this result for fundamental forces, our approach readily extends to constructing higher-derivative operators, and we illustrate its application to structures found in Z-theory and open superstring amplitudes. Furthermore, we explore how this operator-level double copy informs a consistent state-level encoding for quantum gravity.

This work finds itself at the intersection of a number of fields. For amplitudes practitioners, it offers a concrete and structurally faithful map from bootstrapped amplitude data to local operators. For effective field theorists, it provides an efficient alternative to traditional ansatz- and basis-building methods, yielding manifestly gauge-invariant, double-copy-compatible operators imported from their natural habitat. For researchers interested in the foundational structure of QFT and gravity, the method opens a path toward reconstructing semi-classical gravitational actions from gauge-theoretic constituents—an open invitation to flat-space holography. Finally, we hope to present the method with sufficient procedural clarity and pedagogical grounding to serve as a practical and easily automatable tool.

This paper is organized as follows. In section 2, we review the duality between color and kinematics and the graph-based double-copy structure shared by many theories. In section 3, we introduce a pedagogical covariantized free-scalar bootstrap to illustrate our method in a transparent setting. Section 4 describes how contact terms are isolated from amplitude data using a generalized maximal-cut strategy. Section 5 presents the core operator promotion procedure, mapping Jacobi-consistent graph numerators to local field-space operators. In section 6, we apply the method across a range of theories—gauge, string-inspired, and gravitational—highlighting its efficiency and generality. In section 7 we lay out the path to quantum gravity via double copy, first highlighting how every YM tree-level amplitude encodes the necessary Yang-Mills gravity contact required by $\sqrt{-g}R$ at that multiplicity, then laying out a dictionary for gravity states as double-copied Yang-Mills states. We conclude in section 8 with discussion of broader applications and structural implications.

2 Review

2.1 Color-dual representations

The duality between color-and-kinematics originally identified at tree-level [13] and soon thereafter generalized to the multiloop integrand level [14] as well as perturbative and complete classical solutions to the equation of motion [15–18] now bridges many aspects of physics from particle physics, to string theory, to mathematical physics, and from gravitational wave astrophysics, to inflationary cosmology. We would not be able to do justice to the field here, and instead defer to a number of tutorials and reviews [19–25].

In this paper we will focus on the most familiar type of color-dual representations – adjoint antisymmetric where kinematics obey the same structural relations as color-weights dressed with antisymmetric adjoint color factors (f^{abc}). Namely this will mean satisfying Jacobi and antisymmetry. This is sufficient to bootstrap Yang-Mills and the Nonlinear Sigma Model to all multiplicity, and when allowing additional color-structures to combine with kinematics allows for many higher derivative interactions including the open and closed bosonic and superstring theories at tree-level. The methods presented here generalize trivially to double-copies that require symmetric structure constants (d^{abc}) [26], but that will not be the main focus of this paper.

All scattering amplitudes can be expressed in terms of cubic (trivalent) graphs. Higher point contact contributions can be absorbed into cubic graph dressings by including relevant inverse propagators. At tree level for m particles scattering, we need consider a maximum of $(2m - 5)!!$ distinct cubic graphs, Γ_3^m . For the Yang-Mills theory we map these graphs to color-weights (dressing all vertices with f^{abc} structure constants) c_g , kinematic-weights (functions of momenta and polarizations) n_g , and propagators $1/d_g$. As such the full amplitude is given by:

$$\mathcal{A}_m = \sum_{g \in \Gamma_3^m} \frac{c_g n_g}{d_g} \quad (2.1)$$

where we have suppressed the coupling whose power goes as $m - 2$ with multiplicity m . This set grows factorially with m because of all the different labels each topology can have. While it is possible to choose individual mappings for each distinctly labeled graph, this is in some sense artificial for gluons whose polarizations have not yet been specified.

Fortunately there are only an exponential number¹ of distinct cubic topologies, $\mathcal{T}_3^m$, at each multiplicity. Hence we will consider a single functional dressings for each graph's topology. We can therefore dress each topology and sum over permutations of labels to capture all $(2m - 5)!!$ channels,

$$\mathcal{A}_m = \sum_{g \in \mathcal{T}_3^m} S_g \frac{c_g n_g}{d_g} + \text{permutations} . \quad (2.2)$$

¹A000672.

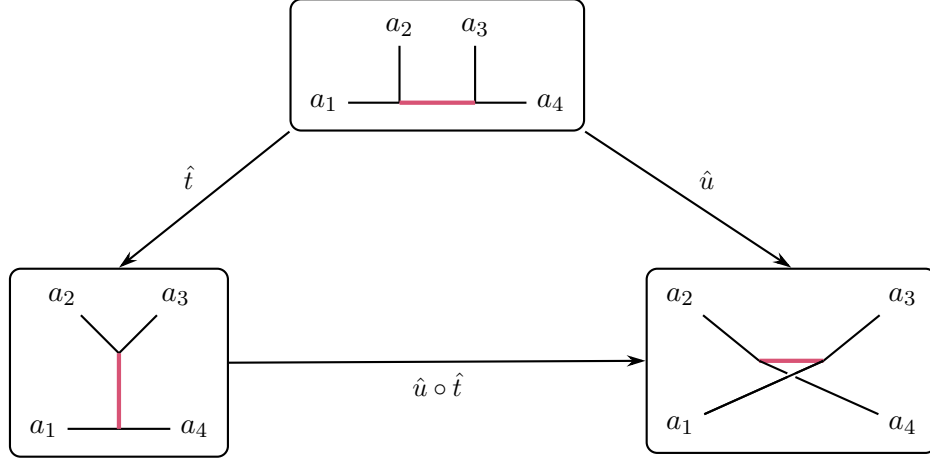


Figure 2: Jacobi moves at four points. Each trivalent graph at four-point multiplicity is related to the others by a local graph transformation, shown as an edge in this graph-of-graphs. These moves correspond to the algebraic Jacobi identities in both color and kinematics. The cyclic ordering of the arrows corresponds to choosing an orientation convention for the antisymmetric vertex structure.

We have introduced symmetry factors S_n to account for the overcounting of relabeling due to each topology's symmetry.

Each one of the distinct topologies in $\mathcal{T}_3^m$ can be expressed as a finite number of Jacobi moves² (see fig. 2) from a single basis graph topology, the so called half-ladder m -point graph,

$$b_m = \begin{array}{c} 2 \qquad \qquad m-1 \\ | \qquad | \qquad \dots \qquad | \\ \hline | \qquad \qquad \qquad | \\ 1 \qquad \qquad \qquad m \end{array} \quad (2.3)$$

This is a familiar concept for color-weights and was used by Dixon, Del Duca, and Maltoni to prove the Kleiss-Kuijff field theory relations. By way of example, consider the first occasion for a non-half-ladder graph, the trimers topology that contributes at 6-points:

$$g_{\text{tri}} = \begin{array}{c} 2 \quad 3 \quad 4 \quad 5 \\ | \quad \diagdown \quad \diagup \quad | \\ | \quad \quad \quad | \\ \hline | \quad \qquad \qquad | \\ 1 \qquad \qquad \qquad 6 \end{array} \quad (2.4)$$

²Sometimes called Whitehead moves [20, 21, 27].

Antisymmetric adjoint structure constant (f^{abc}) based color weights $c(g_{\text{tri}})$ satisfy a Jacobi relation with the color weights of the two differently labeled half-ladder graphs:

$$c \circ \begin{array}{c} b \quad c \quad d \quad e \\ | \quad \diagdown \quad / \quad | \\ \text{---} \\ | \quad | \quad | \quad | \\ a \quad \quad \quad f \end{array} = c \circ \begin{array}{c} b \quad c \quad d \quad e \\ | \quad | \quad | \quad | \\ \text{---} \\ | \quad | \quad | \quad | \\ a \quad \quad \quad f \end{array} - c \circ \begin{array}{c} b \quad d \quad c \quad e \\ | \quad | \quad | \quad | \\ \text{---} \\ | \quad | \quad | \quad | \\ a \quad \quad \quad f \end{array}. \quad (2.5)$$

Color-dual representations of Yang-Mills amplitudes simply have the kinematic weights satisfy exactly the self-same Jacobi and antisymmetry rules, so the kinematic numerators $n(g_{\text{tri}})$ will be given as the difference between the relabeled basis graphs.

$$n \circ \begin{array}{c} b \quad c \quad d \quad e \\ | \quad \diagdown \quad / \quad | \\ \text{---} \\ | \quad | \quad | \quad | \\ a \quad \quad \quad f \end{array} = n \circ \begin{array}{c} b \quad c \quad d \quad e \\ | \quad | \quad | \quad | \\ \text{---} \\ | \quad | \quad | \quad | \\ a \quad \quad \quad f \end{array} - n \circ \begin{array}{c} b \quad d \quad c \quad e \\ | \quad | \quad | \quad | \\ \text{---} \\ | \quad | \quad | \quad | \\ a \quad \quad \quad f \end{array}. \quad (2.6)$$

The fact that adjoint color-dual kinematic numerators also satisfy Jacobi imposes additional relations on ordered amplitudes known as BCJ relations to a basis of $(n-3)!$ orderings.

For theories color-dual to f^{abc} color, we write every tree-level amplitude as:

$$\mathcal{A}_m = S(b_m) \frac{c(b_m)n(b_m)}{d(b_m)} + \text{Jacobi moves} + \text{permutations}. \quad (2.7)$$

We will give various explicit kinematic weights $n(b_m)$ for a variety of theories in later sections. At this time perhaps it can help ground the discussion to generically offer the propagator and adjoint color weights associated with b_m .

For theories like Yang-Mills the color-weights are simply given by dressing every vertex with a gauge theory structure constant labelled according to the graph,

$$c(b_m) = f^{e_1 e_2 i_1} f^{i_1 e_3 i_2} \dots f^{i_{m-3} e_{m-1} e_m}. \quad (2.8)$$

For convenience we distinguish between external color labels e_j and internal color-labels i_k .

For massless fields of any spin, with an all outgoing momentum convention, we can dress the propagators of the half-ladder, b_m ,

$$d(b_m) = (k_1 + k_2)^2 (k_1 + k_2 + k_3)^2 \dots (k_{m-2} + k_{m-1} + k_m)^2 (k_{m-1} + k_m)^2. \quad (2.9)$$

Famously, local individual kinematic weights n_g are not generically gauge invariant. There must be cancelations between distinct channels. Algebraic relations between color-

weights c_g associated with distinct channels ensure the gauge invariance of the full amplitude. Therefore we can preserve this gauge invariance by replacing the c_g with any other graph weights $\tilde{n}_g$ that obey the same algebraic relations – this is known as taking the double-copy between two theories. Critically that means removing a weight in the numerator from each (usually a color-weight) and then combining the remaining kinematic weights,

$$\mathcal{A}_m = S(b_m) \frac{\tilde{n}(b_m)n(b_m)}{d(b_m)} + \text{Jacobi moves} + \text{permutations} . \quad (2.10)$$

Taking both copies to be vector (spin-1) kinematic weights we arrive at amplitudes for gravitational (spin-2) theories. Linearized diffeomorphism invariance emerges from the double-copy of linearized gauge invariance [21].

Generically we will write tree-level amplitudes in theories that participate in the *double-copy* web of theories as follows:

$$\mathcal{A}_m^{A \otimes B} = S(b_m) \frac{n_A(b_m)n_B(b_m)}{d(b_m)} + \text{Jacobi moves} + \text{permutations} \quad (2.11)$$

In each copy n_A , every external field contributes a particular little-group weight corresponding to its representation of the Lorentz group. This is manifested in arbitrary D -dimensions via formal polarization vectors and formal spinors. Here we consider double-copy amplitudes for theories of maximum spin-2, so each copy every external leg can contribute maximum spin-1 little-group weight.

2.2 Higher-derivative compositional bootstrap

The compositional bootstrap of [1, 2, 26] provides a systematic framework for constructing and classifying the predictions of higher derivative operators in gauge and gravity theories. This approach leverages color-kinematics duality to efficiently encode higher derivative corrections to all orders while maintaining manifest double-copy structure. The key insight is that the functional forms associated with various graphical algebraic structures can be composed to generate additional functional forms of higher mass dimension. The algebraic properties of color and kinematic numerators, such as Jacobi identities and vertex symmetries, are crucial for ensuring the overall consistency and physical properties of the resulting amplitudes, including Bose symmetry where applicable.

With a scalar graph weight of linear order one can create a ladder to generate all higher order scalar interactions all the way to the ultra-violet. Antisymmetric f^{abc} structure constants generate color factors that participate in the algebraic relations that characterize the original duality between color-and kinematics.

We summarize some of the results that will refer to directly, but refer the interested reader to the above references for more details and context. We require three antisymmetric and Jacobi-satisfying building blocks to exhaust all such higher-derivative modifications to single-trace color-weights at four-points – up to products of scalar permutation invariants.

As such any arbitrary higher derivative generalization of $c_s \equiv f^{a_1 a_2 e} f^{e a_3 a_4}$, is spanned by,

$$\mathbf{c}_s = \sum_i \alpha'^i (a_{XY} \mathbf{c}_s^{(XY)} + a_{XY}^{ss} \mathbf{c}_s^{(XY,ss)} + a_{XY}^d \mathbf{c}_s^{(XY,d)}) . \quad (2.12)$$

The three building blocks are as follows

$$\mathbf{c}_s^{(X,Y)} \equiv c_s (stu)^X (s^2 + t^2 + u^2)^Y \alpha'^{3X+2Y} , \quad (2.13)$$

$$\mathbf{c}_s^{(X,Y,ss)} \equiv (c_t(u-s) + c_u(s-t)) (stu)^X (s^2 + t^2 + u^2)^Y \alpha'^{1+3X+2Y} . \quad (2.14)$$

$$\mathbf{c}_s^{(XY,d)} \equiv d^{abcd} s(u-t) (stu)^X (s^2 + t^2 + u^2)^Y \alpha'^{(2+3X+2Y)} . \quad (2.15)$$

We introduce the mass-dimension carrying α' , the other antisymmetric channel adjoint weights via relabeling, $c_t = f^{a_4 a_1 e} f^{a_2 a_3 e}$ and $c_u = f^{a_3 a_1 e} f^{a_4 a_2 e}$, and admit the fully permutation invariant color $d^{abcd} = \frac{1}{3!} \sum_{\sigma \in S_3(b,c,d)} \text{Tr}(\mathbf{T}^a \mathbf{T}^{\sigma_1} \mathbf{T}^{\sigma_2} \mathbf{T}^{\sigma_3})$. The other higher derivative channels follow $\mathbf{c}_s$ via relabeling, and all higher-derivative color-numerators satisfy antisymmetry about each vertex, and Jacobi as written,

$$\mathbf{c}_s = \mathbf{c}_t + \mathbf{c}_u . \quad (2.16)$$

This allows us to lift the eight adjoint vector color-weights to all orders in Mandelstams spanning every higher-derivative (parity preserving) vector weight that can be written as an antisymmetric adjoint double-copy. One can complete all higher derivative vector weights by admitting kinematic and color weights that obey symmetric-adjoint double-copy. One can of course carry out a similar program for parity odd higher derivative operators by fixing in dimensions and admitting dual vector dressings.

Having established how tree-level scattering amplitudes in a wide web of theories can be expressed in a color-dual or double-copy form using a basis of graph numerators, including all-orders higher-derivative corrections, we now turn to the central challenge: constructing the corresponding local operators in an effective action that generate these structured amplitudes.

3 An Invitation: The Simple Scalar

We will illustrate our method for writing down actions using the example of the covariantized free scalar, a theory of adjoint scalars and gluons that we will define in two ways: an amplitudes-based approach that defines the theory by its physical properties and a traditional approach of writing down gauge-invariant operators. This will serve as a pedagogic invitation to our method of operator promotion, introducing the ideas one by one within the example, later to be generalized.

3.1 Defining the simple scalar

3.1.1 Bootstrapping the theory from physical principles

We define the simple scalar to be a *gauge-invariant, color-dual theory of a real adjoint scalar minimally coupled to Yang-Mills*, at the lowest mass dimension possible. These basic physical

principles immediately allow us to write down three-point and four-point amplitudes in this theory, which tell us what operators need to be present in the theory.

We can write down the three-point amplitude of two scalars and one gluon by constraining an ansatz to satisfy the Ward identity on the gluon leg and Bose invariance:

$$\mathcal{A}_{\varphi\varphi g} \equiv \mathcal{A}\left(\varphi^a(k_1), \varphi^b(k_2), g^c(k_3)\right) = \alpha f^{abc} (k_2 \cdot \varepsilon_3) \quad (3.1)$$

where α is a free parameter that will later be fixed from considerations of gauge invariance and consistent factorization at higher multiplicity. An ansatz at lowest mass dimension for an amplitude of one scalar interacting with two gluons admits no gauge invariant solutions, so we do not consider such a three-point interaction as part of the simple scalar theory.

From the existence of three gluon interactions (from pure Yang-Mills) and two scalar, one gluon interactions, we can consider two distinct four-point amplitudes: two external scalars and two external gluons, or four external scalars. We write the four-scalar case as a sum over the three cubic graph channels (as defined in Figure 4), dressing each graph with adjoint color c , a kinematic numerator n , and the associated massless cubic propagator d :

$$\mathcal{A}_{\varphi\varphi\varphi\varphi} = \sum_{g \in \{s, t, u\}} \frac{c_g n_g}{d_g} \quad (3.2)$$

The color factor associated with a graph g with leg labels $(abcd)$ is given simply by the adjoint factor built from structure constant contractions, $c_g = f^{abe} f^{ecd}$. Each graph is dressed with a functional numerator n , which we constrain to obey adjoint color-kinematics duality (antisymmetry around each vertex and a Jacobi identity on the internal leg):

$$\begin{aligned} n(abcd) &= -n(bacd) \\ n(abcd) &= -n(abdc) \\ n(abcd) &= +n(dcba) \end{aligned} \quad (3.3)$$

$$n(abcd) = n(dabc) + n(dbca) \quad (3.4)$$

These conditions fix the form of the numerator to be:

$$n(abcd) = \beta (s_{bc} - s_{ac}) \quad (3.5)$$

where the free parameter β can be fixed on factorization of the full four scalar amplitude to

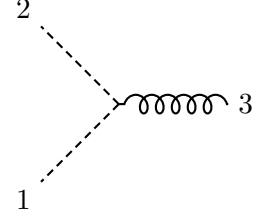


Figure 3: $\varphi\varphi g$ vertex

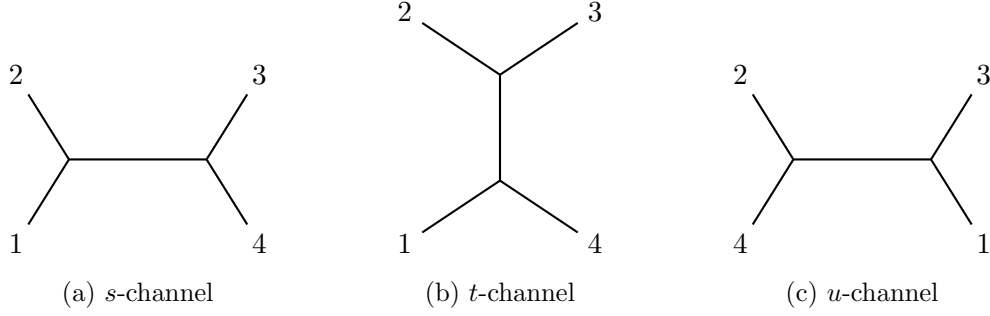


Figure 4: Standard definitions of *s*, *t*, and *u* channel cubic four-point graph labelings.

be $\beta = \alpha^2/2$, according to:

$$\lim_{s \rightarrow 0} s \mathcal{A} \left(\varphi^a(k_1), \varphi^b(k_2), \varphi^c(k_3), \varphi^d(k_4) \right) = \sum_{states} \mathcal{A} \left(\varphi^a(k_1), \varphi^b(k_2), g^{e,(h)}(k_3 + k_4) \right) \times \mathcal{A} \left(\varphi^c(k_3), \varphi^d(k_4), g^{e,(\bar{h})}(k_3 + k_4) \right) \quad (3.6)$$

Computation of the full amplitude indicates that no new information is present in the four scalar amplitude that was not already contained within the three point vertex (i.e., there is no four scalar contact term necessary to ensure the theory satisfies color-kinematics duality at four points):

$$\mathcal{A}_{\varphi\varphi\varphi\varphi} = \frac{\alpha^2}{2} \left(\frac{c_s(u-t)}{s} + \frac{c_t(u-s)}{t} + \frac{c_u(s-t)}{u} \right). \quad (3.7)$$

We can write down the two scalar, two gluon amplitude $\mathcal{A}(\varphi(k_1), \varphi(k_2), g(k_3), g(k_4))$ from similar considerations. There are now two cubic graph topologies, one with a scalar propagator and one with a vector propagator, defined as follows:

$n_\varphi(abcd)$
(3.8)

$n_v(abcd)$
(3.9)

We will proceed in a standard manner: assign each topology an ansatz; constrain each ansatz on its graph symmetries, factorization, and gauge invariance of the constructed full amplitude;

and finally relate the topologies using the kinematic Jacobi relation. As is typical in adjoint color-kinematics considerations, this construction does not exclude the possibility of a four point contact term of the form:

$$\begin{array}{ccc}
 & b & c \\
 & \diagdown & \diagup \\
 & \times & \\
 & \diagup & \diagdown \\
 a & & d
 \end{array} \quad (3.10)$$

Rather, if such a contact term is demanded by gauge invariance, it will be generated by this ansatz procedure as assigned to the cubic graphs in a manner consistent with color-dual functional representations, just as the Yang-Mills four point contact can be absorbed onto the cubic graphs by multiplying by appropriate factors of unity (s/s , t/t , u/u).

The full amplitude can then be written as a sum over the dressings of the relevant cubic graphs — represented schematically,

$$\mathcal{A}_{\varphi\varphi gg} \sim \begin{array}{c} \text{Diagram 1: } k_1 \text{ (dashed), } k_4 \text{ (wavy), } k_2 \text{ (dashed), } k_3 \text{ (wavy)} \\ \text{Diagram 2: } k_1 \text{ (wavy), } k_3 \text{ (dashed), } k_2 \text{ (dashed), } k_4 \text{ (wavy)} \\ \text{Diagram 3: } k_2 \text{ (dashed), } k_1 \text{ (dashed), } k_2 \text{ (dashed), } k_3 \text{ (wavy), } k_4 \text{ (wavy)} \end{array} \cdot (3.11)$$

In terms of our functional dressings and adjoint color factors, the amplitude is given by:

$$\mathcal{A}_{\varphi\varphi gg} = \frac{c_t n_\varphi(1234)}{t} - \frac{c_u n_\varphi(1243)}{u} + \frac{c_s n_v(1234)}{s}. \quad (3.12)$$

Graph symmetries provide the following constraints on the n_φ and n_v functional numerators (these can be understood from drawing the automorphisms of each topology and considering vertex antisymmetry, in line with color-kinematics duality):

$$n_\varphi(abcd) = +n_\varphi(badc) \quad (3.13)$$

$$n_v(abcd) = -n_v(bacd) \quad (3.14)$$

$$n_v(abcd) = -n_v(abdc)$$

Such partially-fixed numerator functions are then further constrained to consistently factorize down to appropriate products of three point amplitudes; for example, the s -channel cut of our amplitude, described by the appropriate limit of the n_v numerator dressing, can be fixed

as follows:

$$\begin{aligned} \lim_{s \rightarrow 0} s \mathcal{A} \left(\varphi^a(k_1), \varphi^b(k_2), g^c(k_3), g^d(k_4) \right) &= \lim_{s \rightarrow 0} c_s n_v(1234) \\ &= \sum_{states} \mathcal{A} \left(\varphi^a(k_1), \varphi^b(k_2), g^{e,(h)}(k_3 + k_4) \right) \\ &\quad \times \mathcal{A} \left(g^c(k_3), g^d(k_4), g^{e,(\bar{h})}(k_3 + k_4) \right) \end{aligned} \quad (3.15)$$

where the first three point amplitude arises from our simple scalar theory, and the second from pure Yang-Mills theory,

$$\mathcal{A}_3^{\text{YM}} = g f^{abc} [(k_1 - k_2) \cdot \varepsilon_3 (\varepsilon_1 \cdot \varepsilon_2) + \text{cyclic}] \quad (3.16)$$

$$= -2g f^{abc} [(k_2 \cdot \varepsilon_3)(\varepsilon_1 \cdot \varepsilon_2) + (k_3 \cdot \varepsilon_1)(\varepsilon_2 \cdot \varepsilon_3) + (k_1 \cdot \varepsilon_2)(\varepsilon_3 \cdot \varepsilon_1)] \quad (3.17)$$

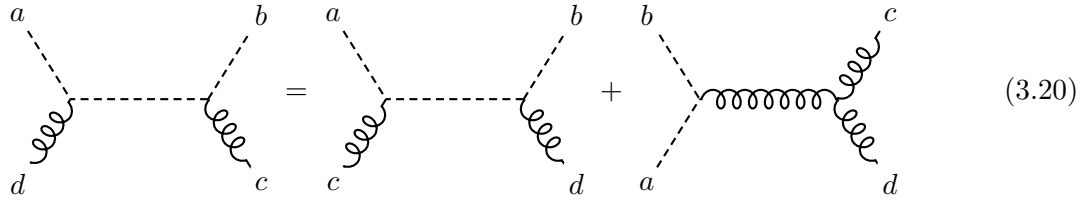
Upon fixing on all distinct cuts, the two scalar, two gluon amplitude $\mathcal{A}_{\varphi\varphi gg}$ contains just two parameters: g , the Yang-Mills coupling, and α , the unfixed overall coefficient of the three point simple scalar amplitude $\mathcal{A}_{\varphi\varphi g}$. By demanding this amplitude be gauge invariant, α is fixed to $-2g$, consistent with the knowledge that the couplings of particles to non-abelian gauge bosons are determined entirely by the pure gauge theory coupling and the relevant group representation. The fully fixed numerators are as follows:

$$n_\varphi(1234) = -4g^2(k_2 \cdot \varepsilon_3)(k_1 \cdot \varepsilon_4) - 2g^2(k_2 \cdot k_3)(\varepsilon_3 \cdot \varepsilon_4) \quad (3.18)$$

$$\begin{aligned} n_v(1234) &= 4g^2(k_1 \cdot \varepsilon_3)(k_2 \cdot \varepsilon_4) - 4g^2(k_1 \cdot \varepsilon_4)(k_2 \cdot \varepsilon_3) \\ &\quad - 4g^2(k_2 \cdot k_3)(\varepsilon_3 \cdot \varepsilon_4) - 2g^2(k_1 \cdot k_2)(\varepsilon_3 \cdot \varepsilon_4) \end{aligned} \quad (3.19)$$

One can note that each dressing contains a term proportional to its associated cubic propagator, confirming that this theory does indeed require a two scalar, two gluon contact term to preserve gauge invariance that we have assigned to the cubic graphs in a notion consistent with color-kinematics duality.

These fully fixed numerator dressings automatically satisfy the kinematic Jacobi identity which relates the two distinct graph topologies:



$$(3.20)$$

$$n_\varphi(abcd) = n_\varphi(abdc) + n_v(abcd) \quad (3.21)$$

This confirms our ability to define this simple scalar theory as a consistent, color-dual, gauge-invariant theory of real adjoint scalars coupled to Yang-Mills.

3.2 Constructing operators

From what we learned building the amplitudes for this theory, we expect the following schematic terms to appear in the Lagrangian for the simple scalar:

$$\mathcal{L} = \mathcal{L}_{\text{Yang-Mills}} + \mathcal{L}_{\text{free-scalar}} + \mathcal{L}_{\varphi\varphi g} + \mathcal{L}_{\varphi\varphi gg}, \quad (3.27)$$

consistent with the types of terms seen when the traditional action formalism for this theory is expanded. We will take a different approach to constructing our action, by instead using the content of the three-point amplitude $\mathcal{A}_{\varphi\varphi g}$ to write down $\mathcal{L}_{\varphi\varphi g}$, and, likewise, the content of the four-point amplitude $\mathcal{A}_{\varphi\varphi gg}$ to write down $\mathcal{L}_{\varphi\varphi gg}$.

We will here introduce a method for promoting an amplitude to a corresponding operator in this simplified case, and provide the more general prescription in section 5. This will work by promoting kinematic numerators n and propagators d to operators, denoted symbolically as $\hat{n}$, $\hat{d}$, given in terms of fields and derivatives. Recall that the amplitude for particle content $\mathcal{P}$ is taken generically to be written a sum over all relevant cubic graphs, each dressed with a kinematic numerator n_g , a color factor c_g , and the graph's associated cubic propagators d_g ,

$$\mathcal{A}_{\mathcal{P}} = \sum_{g \in \Gamma_{\mathcal{P}}^{(3)}} \frac{c_g n_g}{d_g}. \quad (3.28)$$

Promoting will allow us to write down the relevant contribution to the Lagrangian density by constructing a sum over *distinct four-point trivalent-graph topologies* $\tau \in \mathcal{T}_3^4$, each dressed with a corresponding numerator operator, color factor, propagator operator, and symmetry factor (which will be explained shortly):

$$\mathcal{L}_{\mathcal{P}} = \int \mathcal{D}_{|\mathcal{P}|} \sum_{\tau \in \mathcal{T}_3^4 \mathcal{P}} s_{\tau} \frac{c_{\tau} \hat{n}_{\tau}}{\hat{d}_{\tau}} \mathcal{P} \quad (3.29)$$

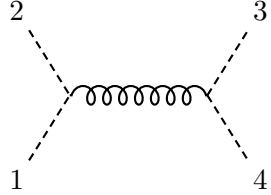
where the sum runs over all distinct cubic graph topologies $\mathcal{T}_3^4 \mathcal{P}$ available given the relevant field content. We have the promoted operators acting on field operators represented here by $\mathcal{P}$, and we introduced an integration as a convenience to localize any labeling coordinates we may have given our fields to make it easy for derivatives to land in the right spot. Color-kinematics duality can then be exploited to write these topology dressings in terms of those corresponding to the minimal basis topologies that arise from solving the Jacobi relations.

This operator $\mathcal{L}_{\mathcal{P}}$ encodes the relevant information in the theory concerning particle content $\mathcal{P}$ in the following sense: dressing *only the contact graph* for this particle content with the vertex rule derived from $\mathcal{L}_{\mathcal{P}}$ reproduces the entire amplitude $\mathcal{A}_{\mathcal{P}}$, even if the theory contains no such contact interaction! This works because the denominators $\hat{d}_{\tau}$ in our operator construction encode the propagator structure associated with all graph contributions to the amplitude (cubic, contact, or otherwise). Simply put, the Feynman vertex rule already contains all the information about graph connectivity and causal structure in the theory, allowing

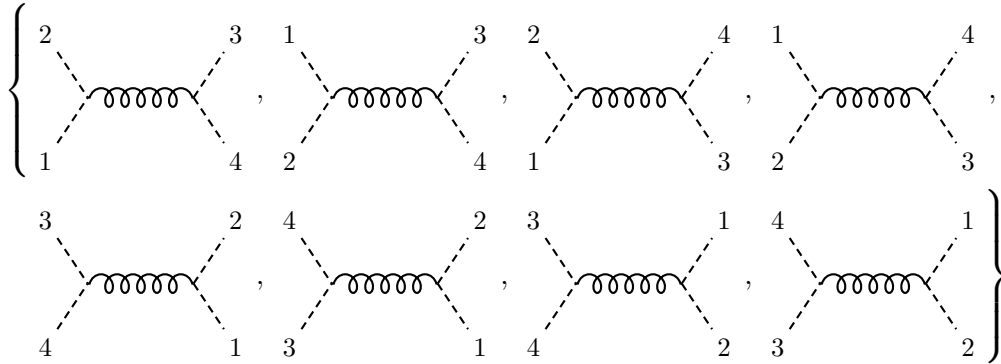
us to assign everything to the contact diagram.

Of course, this is not how Feynman rule calculations are meant to proceed; rather, all possible graph topologies are drawn and dressed with all the rules obtained from the full Lagrangian of the theory. Following such a prescription naively using our operators will result in some redundancy (in which our contact diagrams are dressed to include cubic graph information, and then, separately, cubic graphs are dressed with the lower-point vertices). We will remove such redundancy systematically by constructing additional operators corresponding to the cut contributions to the amplitudes of the theory.

The generation of Feynman rules considers all the possible ways that fields in the action could line up with the fields in initial and final states, effectively constituting a sum over all relevant permutations. This will lead to each topology's contribution being overcounted by precisely its number of automorphisms, necessitating a symmetry factor to compensate. Each symmetry factor s_τ is simply the *reciprocal of the number of automorphisms* of the graph topology τ . For example, for four external scalars in our theory, there is only one topology,


(3.30)

and thus only one term in our candidate operator. In calculating the corresponding Feynman rule, all permutations of (1234) will be generated, but these 24 permutations contain only three distinct variations (the s , t , and u channels, as visualized in Figure 4). Each physically distinguishable channel appears 8 total times, meaning we must divide the result by 8 to correct for this overcounting. We see this factor of 8 precisely because it is the number of automorphisms of this graph topology – the total number of ways we could have labeled it that are, by Bose symmetry, indistinguishable from the original:


(3.31)

So, for this topology, the appropriate symmetry factor would be $s = 1/8$. It is also precisely for this reason that Feynman rules generate all permutations that we do not specify particular

We can achieve this by assigning each field in the operator its own dummy position x_i , which is set by a Dirac delta function to be evaluated only at the spacetime point x that is universal to all fields in the operator, so as not to corrupt the locality of the interaction. Essentially, our procedure will always be to write an arbitrary field ϕ as follows,

$$\phi(x) = \int d^d x_i \delta^{(d)}(x - x_i) \phi(x_i) \quad (3.37)$$

with a unique label i for each field in the operator. This allows us to rewrite the operator as follows,

$$\hat{n}_{\varphi\varphi g}^\mu = 2g \partial_2^\mu, \quad (3.38)$$

so that the full contribution to the Lagrangian $\mathcal{L}_{\varphi\varphi g} = \frac{1}{2} \int \mathcal{D}_3 (f^{abc} \hat{n}_{\varphi\varphi g}^\mu) \phi^a \phi^b A_\mu^c$ will hence take the form:

$$\mathcal{L}_{\varphi\varphi g} = g f^{abc} \int \left(\prod_i^3 d^d x_i \delta^{(d)}(x - x_i) \right) \frac{\partial}{\partial x_{2\mu}} \left[\varphi^a(x_1) \varphi^b(x_2) A_\mu^c(x_3) \right]. \quad (3.39)$$

To write this more succinctly, we will introduce two pieces of notation: first, the shorthand $\partial_2^\mu \equiv \partial/\partial x_{2\mu}$ denotes derivatives with respect to the different spacetime position labels. Second, we define a convenient measure to encode the dummy spacetime variables,

$$\mathcal{D}_n^d x \equiv \left(\prod_i^n d^d x_i \delta^{(d)}(x - x_i) \right) \quad (3.40)$$

so that our $\varphi\varphi g$ contribution to the Lagrangian can be expressed in the compact form:

$$\mathcal{L}_{\varphi\varphi g} = g f^{abc} \int \mathcal{D}_3^d x \partial_2^\mu \left[\varphi^a(x_1) \varphi^b(x_2) A_\mu^c(x_3) \right] \quad (3.41)$$

$$= g f^{abc} \varphi^a(\partial^\mu \varphi^b) A_\mu^c \quad (3.42)$$

We see we got there by basically taking $k_i^\mu \rightarrow \partial_i^\mu$, and $\eta_{\mu\nu} \epsilon^\nu \rightarrow \eta_{\mu\nu}$, a procedure to be elaborated on in more detail in section 5.

Now that we have our operator, we will confirm it gives rise to the correct amplitude by explicitly computing its corresponding vertex rule. The operator can be rewritten in momentum-space via a Fourier transform (all particles are taken to be outgoing to match the conventions used in our amplitudes-based approach):

$$\mathcal{L}_{\varphi\varphi g} = g f^{abc} \int \left(\prod_i^3 d^d k_i e^{ik_i \cdot x} \right) (i k_2^\mu) \left[\tilde{\varphi}^a(k_1) \tilde{\varphi}^b(k_2) \tilde{A}_\mu^c(k_3) \right] \quad (3.43)$$

We now proceed to calculate its contribution to the action $\mathcal{S} = \int d^d x \mathcal{L}$ and relabel dummy

variables and indices:

$$\mathcal{S}_{\varphi\varphi g} = g f^{xyz} \int d^d x \left(\prod_i^3 d^d p_i \right) e^{i(p_1+p_2+p_3) \cdot x} (i p_2^\nu) \left[\tilde{\varphi}^x(p_1) \tilde{\varphi}^y(p_2) \tilde{A}_\nu^z(p_3) \right] \quad (3.44)$$

$$\mathcal{S}_{\varphi\varphi g} = g f^{xyz} \int \left(\prod_i^3 d^d p_i \right) (2\pi)^d \delta^{(d)}(p_1 + p_2 + p_3) (i p_2^\nu) \left[\tilde{\varphi}^x(p_1) \tilde{\varphi}^y(p_2) \tilde{A}_\nu^z(p_3) \right] \quad (3.45)$$

$$(3.46)$$

We can write down the associated Feynman rule for the $\varphi\varphi g$ vertex by taking the functional derivative of the action:

$$\mathcal{V}_\mu^{abc} \equiv \frac{\delta \mathcal{S}}{\delta \tilde{\varphi}^a(k_1) \delta \tilde{\varphi}^b(k_2) \delta \tilde{A}^{\mu,c}(k_3)} \quad (3.47)$$

$$\begin{aligned} \mathcal{V}_\mu^{abc} = g f^{xyz} \int \left(\prod_i^3 d^d p_i \right) (2\pi)^d \delta^{(d)}(p_1 + p_2 + p_3) (i p_{2,\nu}) \delta^{cz} \delta_\mu^\nu \delta^{(d)}(k_3 - p_3) \\ \left[\delta^{ax} \delta^{by} \delta^{(d)}(k_1 - p_1) \delta^{(d)}(k_2 - p_2) + \delta^{bx} \delta^{ay} \delta^{(d)}(k_2 - p_1) \delta^{(d)}(k_1 - p_2) \right] \end{aligned} \quad (3.48)$$

$$\mathcal{V}_\mu^{abc} = i g (2\pi)^d \delta^{(d)}(k_1 + k_2 + k_3) f^{abc} (k_2 - k_1)_\mu \quad (3.49)$$

The full on-shell all-outgoing three-point S -matrix element $\langle \varphi\varphi g | S | 0 \rangle$ can be computed directly from this vertex rule by contracting it with the external gluon's polarization vector:

$$\langle \varphi\varphi g | S | 0 \rangle = \mathcal{V}_\mu^{abc} \varepsilon_3^\mu = 2 i g (2\pi)^d \delta^{(d)}(k_1 + k_2 + k_3) f^{abc} (k_2 \cdot \varepsilon_3) \quad (3.50)$$

Finally, from the standard definition $\langle f | S | i \rangle = (2\pi)^d \delta^{(d)}(k_f - k_i) i \mathcal{A}$, we can read off the scattering amplitude,

$$\mathcal{A} = 2 g f^{abc} (k_2 \cdot \varepsilon_3) . \quad (3.51)$$

This result is in perfect agreement with the result found using amplitudes-based considerations, as desired.

3.2.2 Four field operator

Now we would like to encode the information from the two scalar, two gluon amplitude $\mathcal{A}_{\varphi\varphi gg}$ in a (admittedly nonlocal) four-field operator $\mathcal{L}_{\varphi\varphi gg}$. Of course this will introduce redundancies as we discuss and remove below, but for now we want a four-field operator that reproduces the four-field amplitude in its entirety.

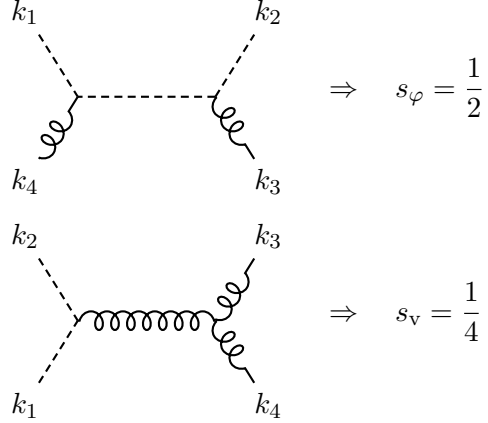
This amplitude is expressed as the following sum over graphs, depicted in equation 3.11 and written functionally as:

$$\mathcal{A}_{\varphi\varphi gg} = \frac{c_t n_\varphi(1234)}{t} - \frac{c_u n_\varphi(1243)}{u} + \frac{c_s n_v(1234)}{s} , \quad (3.52)$$

In this case, we have two topologies, labeled φ and v (according to the particle-type of their propagators). We must then construct our Lagrangian in terms of the associated dressings of both topologies:

$$\mathcal{L}_{\varphi\varphi gg} = \int \mathcal{D}_4 \left(s_\varphi \frac{c_\varphi^{abcd} \hat{n}_\varphi^{\mu\nu}}{\hat{d}_\varphi} + s_v \frac{c_v^{abcd} \hat{n}_v^{\mu\nu}}{\hat{d}_v} \right) \varphi^a \varphi^b A_\mu^c A_\nu^d \equiv \mathcal{O}_\varphi + \mathcal{O}_v \quad (3.53)$$

Again, we arrive at the symmetry factors by counting automorphisms for each topology:



$$\Rightarrow s_\varphi = \frac{1}{2} \quad (3.54)$$

$$\Rightarrow s_v = \frac{1}{4}$$

The numerator operators can be written in the same manner as at three points. The scalar propagator graph dressing n_φ is translated into the operator $\hat{n}_\varphi$ as follows:

$$n_\varphi(1234) = -2g^2 [2(k_2 \cdot \varepsilon_3)(k_1 \cdot \varepsilon_4) + (k_2 \cdot k_3)(\varepsilon_3 \cdot \varepsilon_4)] \quad (3.55)$$

$$\hat{n}_\varphi^{\mu\nu} = -2g^2 (2\partial_2^\mu \partial_1^\nu + \eta^{\mu\nu} \partial_2^\rho \partial_{3,\rho}) \quad (3.56)$$

Finally, we encode the propagator structure of the graphs using derivatives in the denominator, so the $d_\varphi = s_{23} = (k_2 + k_3)^2 = 2k_2 \cdot k_3$ propagator is promoted to $\hat{d}_\varphi = 2(\partial_2 \cdot \partial_3) = 2(\partial_{2,\mu} \partial_3^\mu)$. We can then arrive at this topology's contribution to the Lagrangian as:

$$\mathcal{O}_\varphi = -g^2 f^{dae} f^{ebc} \int \mathcal{D}_4 \frac{(2\partial_2^\mu \partial_1^\nu + \eta^{\mu\nu} \partial_2^\rho \partial_{3,\rho})}{2(\partial_2 \cdot \partial_3)} \varphi^a(x_1) \varphi^b(x_2) A_\mu^c(x_3) A_\nu^d(x_4) \quad (3.57)$$

While the non-local appearance of this operator, with derivatives in the denominator, can ring alarm bells signaling action at a distance, computation of the associated vertex rule will show that this non-locality is actually precisely encoding the causal structure of our theory: the presence of a denominator in this expression, which has been written as a four-field contact term, gives rise to the cubic propagators appropriate for this operator's contribution to the consistently-factorizing amplitude.

The vector propagator graph dressing n_v follows from the kinematic Jacobi relation:

$$n_v(abcd) = n_\varphi(abcd) - n_\varphi(abdc) \quad (3.58)$$

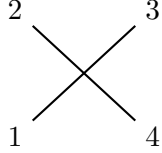
We opt to simply promote this Jacobi identity between functional numerators n to one between operators $\hat{n}$, and then construct $\mathcal{O}_v$ in the same way as the scalar-propagator case, with $c_v = f^{abe} f^{ecd}$, $\hat{d}_v = 2(\partial_1 \cdot \partial_2)$, and $s_v = 1/4$.

$$\mathcal{O}_v = -\frac{1}{2} g^2 f^{abe} f^{ecd} \int \mathcal{D}_4 \frac{(2 \partial_2^{[\mu} \partial_1^{\nu]} + \eta^{\mu\nu} \partial_2^\rho \partial_{3,\rho} - \eta^{\mu\nu} \partial_2^\rho \partial_{4,\rho})}{2(\partial_1 \cdot \partial_2)} \varphi^a(x_1) \varphi^b(x_2) A_\mu^c(x_3) A_\nu^d(x_4) \quad (3.59)$$

Thus, we are able to systematically write down a Lagrangian contribution $\mathcal{L}_{\varphi\varphi gg} = \mathcal{O}_\varphi + \mathcal{O}_v$. A complete calculation of its associated four-point vertex rule (as it is a four-field operator, albeit non-local) $\mathcal{V}_{\mu\nu}^{abcd}$ which yields

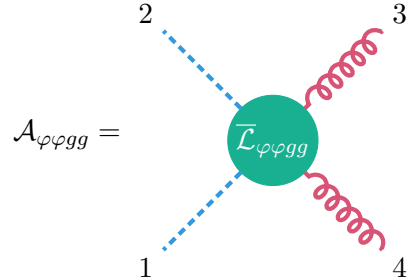
$$\mathcal{V}_{\mu\nu}^{abcd} \epsilon^\mu \epsilon^\nu = \mathcal{A}_{\varphi\varphi gg}. \quad (3.60)$$

In other words we can calculate the corresponding on-shell amplitude for two scalars and two gluons by dressing just the contact diagram:



$$(3.61)$$

with the Feynman rule $\mathcal{V}_{\mu\nu}^{abcd}$; this successfully reproduces $\mathcal{A}_{\varphi\varphi gg}$ as desired, with color-dual structure manifest and the non-localities in the operators $\mathcal{O}$ converted into the appropriate cubic propagators demanded by causality. We will denote this amplitude (and further amplitudes) schematically by drawing the graph contributions, with vertices labeled with the specific operator from which its dressing arises,



$$\mathcal{A}_{\varphi\varphi gg} = \quad (3.62)$$

We emphasize here that we are not done! This is quite a distinct prescription from the traditional method of calculating an amplitude via Feynman diagrams. One must dress *all*

possible graphs at the specified multiplicity and order in quantum correction, not just the contact diagram. If one does this with the operators we have introduced so far we will have a problem. Remedying this problem involves removing redundancies by considering cuts as we now discuss.

3.3 Writing the actual Lagrangian

We have thus far given a prescription for generating a three-field operator, which, when turned into a Feynman vertex rule, reproduces the three-point amplitude; and a four-field operator, which, when used to dress a four-point contact diagram *in isolation*, reproduces the four-point amplitude. By this, we mean that $\mathcal{A}_{\varphi\varphi gg}$ is computed solely by dressing a four-point contact diagram with the vertex rule arising from $\mathcal{L}_{\varphi\varphi gg}$. But of course, this is not how Feynman rule calculations transpire: one must also write down the cubic graphs at four points, and dress them with the rules arising from lower-multiplicity operators $\mathcal{L}_{\varphi\varphi g}$ and $\mathcal{L}_{\text{Yang-Mills}}$.

With our operators as currently written, carrying out the standard procedure of writing both the cubic and quartic graphs and dressing with appropriate vertices specified by the candidate Lagrangian $\mathcal{L}_{\text{Yang-Mills}} + \mathcal{L}_{\text{free-scalar}} + \mathcal{L}_{\varphi\varphi g} + \mathcal{L}_{\varphi\varphi gg}$ will not yield the correct amplitude, but rather, a gauge-dependent result containing redundant contributions. To construct a formally correct Lagrangian, then, we must remove this redundancy, which we will achieve by subtracting operators corresponding to the amplitude's cuts, so that the resulting four-field operator only contains the true local contact term.

3.3.1 Removing redundancy at four points

For the two scalar, two gluon case, using the traditional form of the Lagrangian in equation 3.24, such a calculation is depicted schematically as follows, dressing the three cubic graphs in addition to the contact diagram:

$$\begin{aligned}
 \mathcal{A}_{\varphi\varphi gg} = & \quad \begin{array}{c} 1 \quad 2 \\ \diagdown \quad \diagup \\ \text{---} \mathcal{L}_{\varphi\varphi g} \text{---} \mathcal{L}_{\varphi\varphi g} \text{---} \\ \diagup \quad \diagdown \\ 4 \quad 3 \end{array} + \begin{array}{c} 1 \quad 2 \\ \diagdown \quad \diagup \\ \text{---} \mathcal{L}_{\varphi\varphi g} \text{---} \mathcal{L}_{\varphi\varphi g} \text{---} \\ \diagup \quad \diagdown \\ 3 \quad 4 \end{array} \\
 & + \begin{array}{c} 2 \quad 3 \\ \diagdown \quad \diagup \\ \text{---} \mathcal{L}_{\varphi\varphi g} \text{---} \mathcal{L}_{\text{YM}} \text{---} \\ \diagup \quad \diagdown \\ 1 \quad 4 \end{array} + \begin{array}{c} 2 \quad 3 \\ \diagdown \quad \diagup \\ \text{---} \mathcal{L}_{\varphi\varphi gg} \text{---} \\ \diagup \quad \diagdown \\ 1 \quad 4 \end{array} \quad (3.63)
 \end{aligned}$$

If we attempt to naively replicate this calculation but using the Feynman rules generated from our operators as constructed, we will arrive at an unphysical answer,

$$\mathcal{A}_{\varphi\varphi gg} \neq \begin{array}{c} \begin{array}{ccc} 1 & & 2 \\ & \searrow & \nearrow \\ & \bar{\mathcal{L}}_{\varphi\varphi g} & \bar{\mathcal{L}}_{\varphi\varphi\varphi g} \\ & \nearrow & \searrow \\ 4 & & 3 \end{array} + \begin{array}{ccc} 1 & & 2 \\ & \searrow & \nearrow \\ & \bar{\mathcal{L}}_{\varphi\varphi g} & \bar{\mathcal{L}}_{\varphi\varphi\varphi g} \\ & \nearrow & \searrow \\ 3 & & 4 \end{array} \\ + \begin{array}{ccc} 2 & & 3 \\ & \searrow & \nearrow \\ & \bar{\mathcal{L}}_{\varphi\varphi g} & \bar{\mathcal{L}}_{YM} \\ & \nearrow & \searrow \\ 1 & & 4 \end{array} + \begin{array}{ccc} 2 & & 3 \\ & \searrow & \nearrow \\ & \bar{\mathcal{L}}_{\varphi\varphi gg} & \\ & \nearrow & \searrow \\ 1 & & 4 \end{array} \end{array}, \quad (3.64)$$

precisely because we have already encoded the content of the three cubic graph contributions within the four-point contact operator $\bar{\mathcal{L}}_{\varphi\varphi gg}$, so this calculation is redundant. While the resolution is simple – only dress the contact diagram – we now will aim to write down a full Lagrangian that yields the correct amplitude even using the traditional method of dressing all possible Feynman diagrams. This will necessitate the introduction of additional operators to compensate for this overcounting of the cubic graph contributions. Such information can be extracted from our bootstrap approach by careful consideration of the cuts of the desired amplitude.

Schematically, we will write the Lagrangian in the following form:

$$\mathcal{L} = \bar{\mathcal{L}}_{\text{Yang-Mills}} + \bar{\mathcal{L}}_{\text{free-scalar}} + \bar{\mathcal{L}}_{\varphi\varphi g} + \left(\bar{\mathcal{L}}_{\varphi\varphi gg} - \bar{\not{\mathcal{L}}}_{\varphi\varphi gg} \right), \quad (3.65)$$

where we introduce the slashed operator $\bar{\not{\mathcal{L}}}_{\varphi\varphi gg}$ to remove this redundancy corresponding to the overcounting of cubic graph contributions. This will be a (non-local) four field contact operator, just like $\bar{\mathcal{L}}_{\varphi\varphi gg}$; hence, we will derive from it an additional Feynman rule for dressing the four particle vertex. The traditional application of this Lagrangian's Feynman rules to

all possible graphs will then take the form:

$$\begin{aligned}
\mathcal{A}_{\varphi\varphi gg} = & \text{Diagram 1} + \text{Diagram 2} \\
& + \text{Diagram 3} + \text{Diagram 4} - \text{Diagram 5} .
\end{aligned}
\tag{3.66}$$

For this equality to hold, we must construct the new operator $\bar{\mathcal{Z}}_{\varphi\varphi gg}$ to encode precisely the same information as the three cubic graph dressings, so that we can achieve a cancellation yielding the correct (and original form of) the amplitude:

$$\mathcal{A}_{\varphi\varphi gg} = \text{Diagram 6} .
\tag{3.67}$$

We will achieve precisely this result by considering the cuts of the desired amplitude, and encoding said information in $\bar{\mathcal{Z}}_{\varphi\varphi gg}$.

To understand the relationship between the cubic graph overcounting and the cuts of the amplitude, let's consider in closer detail precisely what goes wrong without the inclusion of this slashed operator $\bar{\mathcal{Z}}_{\varphi\varphi gg}$. The schematic form of the calculation, as depicted in equation 3.64, does not initially raise any red flags: all diagrams are dressed with the rules from their corresponding operators. The issue lies in the fact that the four point contact diagram is dressed with non-local rule, rather than a truly local contact contribution, as would be the case when using the traditional Lagrangian, when instead this graph is dressed with the rule

arising from the familiar genuinely local four-field operator

$$\mathcal{L}_{\varphi\varphi gg} = \frac{g^2}{2} f^{abe} f^{ecd} \varphi^a A_\mu^b \varphi^c A_\nu^d \eta^{\mu\nu} \quad (3.68)$$

To remedy this issue, we insist that the *difference* of these four-field operators, $(\bar{\mathcal{L}}_{\varphi\varphi gg} - \bar{\mathcal{F}}_{\varphi\varphi gg})$, be a truly local contact operator. Since the original operator $\bar{\mathcal{L}}_{\varphi\varphi gg}$ was constructed from the full amplitude, it stands to reason that, to obtain just the local contact contribution, we must subtract the information living on the amplitude's cuts.

To make this a bit more precise, let's write down the local contact piece of the amplitude in the following way,

$$\mathcal{C}_{\varphi\varphi gg} = \sum_g \left(\frac{c_g n_g}{d_g} - \frac{c_g \not{n}_g}{d_g} \right) \quad (3.69)$$

where the slashed numerators $\not{n}_g$ are written down to encode that particular graph's unique cut contributions (this will be defined generically in the following subsection, but will be sufficiently intuitive at four points to construct using this explicit example). We then promote these slashed numerators to operators in precisely the same manner as standard numerators, so that the desired difference of operators can be written as:

$$\bar{\mathcal{L}}_{\varphi\varphi gg} - \bar{\mathcal{F}}_{\varphi\varphi gg} = \int \mathcal{D}_4 \left(\sum_{\tau \in \mathcal{T}_3^4 \varphi\varphi gg} s_\tau \frac{c_\tau (\hat{n}_\tau - \hat{\not{n}}_\tau)}{\hat{d}_\tau} \right) \varphi^a \varphi^b A_\mu^c A_\nu^d. \quad (3.70)$$

The distinct cubic topologies are our familiar scalar propagator and vector propagator graphs, so the sum expands as follows:

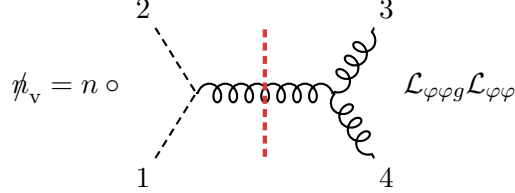
$$\sum_{\tau \in \mathcal{T}_3^4 \varphi\varphi gg} \left(s_\tau \frac{c_\tau (\hat{n}_\tau - \hat{\not{n}}_\tau)}{\hat{d}_\tau} \right) = \left(s_\varphi \frac{c_\varphi (\hat{n}_\varphi - \hat{\not{n}}_\varphi)}{\hat{d}_\varphi} + s_v \frac{c_v (\hat{n}_v - \hat{\not{n}}_v)}{\hat{d}_v} \right). \quad (3.71)$$

Schematically, these slashed numerators correspond to how the graph topology can contribute to any possible cuts of the amplitude. We note that in this form the color-dual nature of the amplitudes themselves is manifest. The scalar propagator topology contains information that survives the s_{14} cut of the amplitude, depicted schematically as:

This can be calculated by imposing the cut conditions to arrive at:

$$\not{n}_\varphi(1234) = -4g^2(k_2 \cdot \varepsilon_3)(k_1 \cdot \varepsilon_4) \quad (3.73)$$

Similarly, the vector propagator topology dressing has a contribution that survives the s_{12} channel cut,



$$\not{n}_v = n \circ \quad \mathcal{L}_{\varphi\varphi g} \mathcal{L}_{\varphi\varphi g} \quad (3.74)$$

$$\begin{aligned} \not{n}_v(1234) &= 4g^2(k_1 \cdot \varepsilon_3)(k_2 \cdot \varepsilon_4) - 4g^2(k_1 \cdot \varepsilon_4)(k_2 \cdot \varepsilon_3) \\ &\quad - 4g^2(k_2 \cdot k_3)(\varepsilon_3 \cdot \varepsilon_4) \end{aligned} \quad (3.75)$$

One should not be concerned by the absence of s_{13} cut information, as our procedure, in summing over all possible permutations of particle labels, will encode this content in the Feynman rule.

These cut numerators $\not{n}$ are promoted to operators just like the standard numerators,

$$\Phi_\varphi = -g^2 f^{dae} f^{ebc} \int \mathcal{D}_4 \frac{(2\partial_2^\mu \partial_1^\nu)}{2(\partial_2 \cdot \partial_3)} \varphi^a(x_1) \varphi^b(x_2) A_\mu^c(x_3) A_\mu^d(x_4), \quad (3.76)$$

$$\Phi_v = -\frac{1}{2}g^2 f^{abe} f^{ecd} \int \mathcal{D}_4 \frac{(2\partial_2^{[\mu} \partial_1^{\nu]} + 2\eta^{\mu\nu} \partial_2^\rho \partial_{3,\rho})}{2(\partial_1 \cdot \partial_2)} \varphi^a(x_1) \varphi^b(x_2) A_\mu^c(x_3) A_\mu^d(x_4). \quad (3.77)$$

Using conservation of momentum and the results in equations 3.57 and 3.59 we therefore find,

$$\begin{aligned} \mathcal{L}_{\varphi\varphi gg} - \not{\mathcal{L}}_{\varphi\varphi gg} &= -g^2 f^{dae} f^{ebc} \int \mathcal{D}_4 \frac{(\eta^{\mu\nu} \partial_2 \cdot \partial_3)}{2(\partial_2 \cdot \partial_3)} \varphi^a(x_1) \varphi^b(x_2) A_\mu^c(x_3) A_\mu^d(x_4) \\ &\quad - \frac{1}{2}g^2 f^{abe} f^{ecd} \int \mathcal{D}_4 \frac{(\eta^{\mu\nu} \partial_1 \cdot \partial_2)}{2(\partial_1 \cdot \partial_2)} \varphi^a(x_1) \varphi^b(x_2) A_\mu^c(x_3) A_\mu^d(x_4) \end{aligned} \quad (3.78)$$

Noting the desired explicit cancellation of the propagators, we can simplify this to recover the familiar manifestly local four-field contact term of eq. (3.68),

$$\mathcal{L}_{\varphi\varphi gg} - \not{\mathcal{L}}_{\varphi\varphi gg} = -\frac{1}{4}g^2 \int \mathcal{D}_4 \left(2f^{dae} f^{ebc} + f^{abe} f^{ecd} \right) \varphi^a(x_1) \varphi^b(x_2) A^{c\mu}(x_3) A_\mu^d(x_4) \quad (3.79)$$

$$= -\frac{1}{4}g^2 \left(2f^{dae} f^{ebc} + f^{abe} f^{ecd} \right) \varphi^a(x) A_\mu^d(x) \varphi^b(x) A^{c\mu}(x) \quad (3.80)$$

$$= -\frac{1}{4}g^2 \left(2f^{dae} f^{ebc} \right) \varphi^a(x) A_\mu^d(x) \varphi^b(x) A^{c\mu}(x) \quad (3.81)$$

$$= -\frac{1}{2}g^2 \left(f^{bae} f^{ecd} \right) \varphi^a(x) A_\mu^b(x) \varphi^c(x) A^{d\mu}(x) \quad (3.82)$$

$$= \frac{1}{2} g^2 \left(f^{abe} f^{ecd} \right) \varphi^a(x) A_\mu^b(x) \varphi^c(x) A^{d\mu}(x) \quad (3.83)$$

$$= \mathcal{C}_{\varphi\varphi gg} . \quad (3.84)$$

In going from the second to third line we note that the contraction of the antisymmetric structure constants with symmetric field labels (both in scalars and gluons) vanishes after the fields are all brought locally, and in going from the third to fourth line simply relabel indices to line up field labeling with the usual form of the contact written in the action eq. (3.68).

As a contrasting example, we will carry out the same procedure for the four scalar amplitude (calculated in equation 3.7) and its associated operators. We again note that there is no apparent contact contribution:

$$\mathcal{A}_{\varphi\varphi\varphi\varphi} = 2g^2 \left(\frac{c_s(u-t)}{s} + \frac{c_t(u-s)}{t} + \frac{c_u(s-t)}{u} \right) . \quad (3.85)$$

The operator in can be written down in the same manner as elaborated in the previous subsection; we write down $\mathcal{L}_{\varphi\varphi\varphi\varphi} = \int \mathcal{D}_4 s_4 \left(c_4 \hat{n}_4 / \hat{d}_4 \right) \varphi^4$ from the constructed numerator n_4 :

$$n_4(1234) = 2g^2 (s_{13} - s_{23}) \quad (3.86)$$

$$\mathcal{L}_{\varphi\varphi\varphi\varphi} = \frac{1}{8} (2g^2) f^{abe} f^{ecd} \int \mathcal{D}_4 \frac{(\partial_1 \cdot \partial_3) - (\partial_2 \cdot \partial_3)}{(\partial_1 \cdot \partial_2)} \varphi^a(x_1) \varphi^b(x_2) \varphi^c(x_3) \varphi^d(x_4) \quad (3.87)$$

This operator, when used to dress the four-scalar contact graph, yields the desired four-scalar amplitude. To remove redundancy associated with dressing both the contact and the relevant cubic graphs, we must write:

$$\mathcal{L}_{\varphi\varphi\varphi\varphi} - \not{\mathcal{L}}_{\varphi\varphi\varphi\varphi} = \int \mathcal{D}_4 \left(s_4 \frac{c_4 \hat{n}_4}{\hat{d}_4} - s_4 \frac{c_4 \not{n}_4}{\hat{d}_4} \right) \varphi^4 \quad (3.88)$$

The cut contribution $\not{n}_4$ is simply the operator promotion of the numerator's unique cut contributions, $\not{n}_4$. In this case, as no contact is required for the duality between color and kinematics, this leaves the numerator unchanged,

$$\not{n}_4(1234) = 2g^2 (s_{13} - s_{23}) = n_4(1234) , \quad (3.89)$$

so the slashed operator will give the same contribution as the original operator. Hence, the overall contribution to the Lagrangian vanishes,

$$\mathcal{L}_{\varphi\varphi\varphi\varphi} - \not{\mathcal{L}}_{\varphi\varphi\varphi\varphi} = 0 , \quad (3.90)$$

consistent with our finding that there is no contact contribution to the full four-scalar amplitude $\mathcal{A}_{\varphi\varphi\varphi\varphi}$.

3.4 Summary

Let us take the opportunity to summarize what we accomplished this invitation. We bootstrapped the amplitudes associated with the theory of the minimally coupled real adjoint scalar by imposing the duality between color and kinematics, and consistent factorization. We chose this theory for the invitation because the presence of external scalars allows for color-dual scattering amplitudes with non-trivial kinematic factors, but without necessarily drowning in polarization vectors. We then took that perturbative interacting data and constructively wrote down the corresponding interaction operators required to generate that data. In this section we restricted ourselves to operators involving scalar fields, and will demonstrate the promotion of gluon operators after we discuss our procedure in full

The familiar action for the covariantized free scalar theory is given by the Lagrangian,

$$\mathcal{L} = -\frac{1}{4} \text{Tr} (F^2) + \frac{1}{2} (D_\mu \varphi)^a (D^\mu \varphi)^a \quad (3.91)$$

$$= -\frac{1}{4} \text{Tr} (F^2) + \frac{1}{2} (\partial \varphi)^2 + g f^{abc} (\partial^\mu \varphi^a) A_\mu^b \varphi^c + \frac{g^2}{2} f^{abe} f^{ecd} \varphi^a A_\mu^b \varphi^c A_\nu^d \eta^{\mu\nu} \quad (3.92)$$

We reproduced the cubic scalar interaction $g f^{abc} \phi^a A_\mu^b \partial^\mu \phi^c$ by considering the three-point bootstrapped amplitude eq. (3.32), $\mathcal{A}_3 = 2g f^{abc} (k_2 \cdot \varepsilon_3)$. Applying eq. (3.29) to the 3-point amplitudes with graph symmetry factor 1/2 associated with the automorphic exchange of scalar legs $1 \leftrightarrow 2$, we simply found, eq. (3.93),

$$\mathcal{L}_{\varphi\varphi g} = \int \mathcal{D}_3(s_3 c_3 \hat{n}_{\varphi\varphi g}) \varphi^2 A \quad (3.93)$$

$$= g f^{abc} \int \mathcal{D}_3 \partial_2^\mu \left[\varphi^a(x_1) \varphi^b(x_2) A_\mu^c(x_3) \right] \quad (3.94)$$

$$= g f^{abc} \varphi^a (\partial^\mu \varphi^b) A_\mu^c \quad (3.95)$$

We reproduced the four point contact interaction $\frac{g^2}{2} f^{bce} f^{eda} \varphi^b \varphi^d \eta^{\mu\nu} A_\mu^c A_\nu^a$ by isolating the contact in the four-point amplitude and promoting it to an operator. How did we isolate the contact? We exploited the fact that Feynman rules encode permutation symmetry. We bootstrapped the necessary topologies to construct the full 4-point amplitude and promoted them to operators– which of course would result in overcounting, so then we subtracted the operators associated with the cut of the amplitude, yielding (eq. (3.84)) precisely

$$\mathcal{L}_{\varphi\varphi gg} - \not\mathcal{L}_{\varphi\varphi gg} = \int \mathcal{D}_4 \left(s_\varphi \frac{c_\varphi (\hat{n}_\varphi - \not{n}_\varphi)}{\hat{d}_\varphi} + s_v \frac{c_v (\hat{n}_v - \not{n}_v)}{\hat{d}_v} \right) \varphi^a \varphi^b A_\mu^c A_\nu^d \quad (3.96)$$

$$= -\frac{1}{4} g^2 \int \mathcal{D}_4 \left(2 f^{dae} f^{ebc} + f^{abe} f^{ecd} \right) \varphi^a(x_1) \varphi^b(x_2) A^{c\mu}(x_3) A_\mu^d(x_4) \quad (3.97)$$

$$= -\frac{1}{4} g^2 \left(2 f^{dae} f^{ebc} + f^{abe} f^{ecd} \right) \varphi^a(x) \varphi^b(x) A^{c\mu}(x) A_\mu^d(x) \quad (3.98)$$

$$= \frac{g^2}{2} f^{bce} f^{eda} \varphi^b \varphi^d \eta^{\mu\nu} A_\mu^c A_\nu^a. \quad (3.99)$$

We also demonstrated that the approach of building operators and subtracting cuts suggests no four-scalar contact associated with this theory even though there is a non-trivial color-dual scattering amplitude.

4 Extracting Novel Contact Terms From Scattering Amplitudes

The core challenge in constructing higher-derivative operators from scattering amplitudes lies in identifying what information is genuinely new at each multiplicity. That is, which terms in an m -point amplitude require the introduction of new m -field operators — and which simply arise from lower-multiplicity physics via unitarity. This section presents a sharp and general method for isolating these novel contributions using graph-local data in color-dual or double-copy representations inspired by the systematic application of generalized unitarity [28–33] known as the method of maximal cuts [20, 28, 34–36].

4.1 General Procedure

An m -point amplitude contains both *old* and *new* information. The old content arises from lower-point interactions — it propagates through factorization channels and corresponds to residues on physical poles. The new content, by contrast, originates from local m -field contact operators that contribute only starting at m -points.

To determine which new operators are needed at multiplicity m , we must isolate these novel contact contributions. This is possible because all cut-constructible terms — those with physical poles — are already determined by lower-multiplicity data. The remaining contact terms represent genuine new operator content.

Schematically we can simply say each amplitude is the sum over any new contacts, $\mathcal{C}$, plus residues of cuts over their uncut propagators, $\mathcal{A}$,

$$\mathcal{A}_m = \mathcal{C}_m + \mathcal{A}_m. \quad (4.1)$$

This can be used as a functional definition:

$$\boxed{\mathcal{C}_m \equiv \mathcal{A}_m - \mathcal{A}_m.} \quad (4.2)$$

This functional identity defines the m -point contact term as the difference between the full amplitude and the sum of all contributions reconstructible from lower multiplicity via cuts.

In Section 5 we will define a prescription, $\hat{\mathcal{O}}$, for promoting m -point amplitude expressions, X_m , to m -field operators, $\mathcal{O}_m$. Specifically the promotion $\hat{\mathcal{O}}$ will be defined such that the Feynman rule associated with the operator, $\mathcal{O}_m \equiv \hat{\mathcal{O}} \circ (X_m)$, dresses the m -point contact graph to contribute precisely the expression X_m . As such, we can write the unique m -field

interaction, $\mathcal{L}_m$, from the novel contact information within the m -point amplitude,

$$\mathcal{L}_m = \hat{\mathcal{O}} \circ \mathcal{C}_m \quad (4.3)$$

$$= \hat{\mathcal{O}} \circ (\mathcal{A}_m - \mathcal{A}_m) \quad (4.4)$$

$$= \hat{\mathcal{O}}(\mathcal{A}_m) - \hat{\mathcal{O}}(\mathcal{A}_m) \quad (4.5)$$

This is morally how we will promote the color-dual representations of our amplitudes to operator form in our Lagrangian density without overcounting the redundant information that lives on the cuts,

$$\mathcal{S} = \int d^d x \left[\mathcal{L}_{\text{free}} + \sum_m \mathcal{L}_m \right] \quad (4.6)$$

$$= \int d^d x \left[\mathcal{L}_{\text{free}} + \sum_m \hat{\mathcal{O}}(\mathcal{A}_m) - \sum_m \hat{\mathcal{O}}(\mathcal{A}_m) \right]. \quad (4.7)$$

Of course the alignment of cut construction with Feynman rules in a particular gauge only has to hold on-shell after sum over all diagrams contributing to cuts, and generalized gauge choice consistency between contact terms is a matter of book-keeping given particular gauge choices. This means that practically one should ensure that one's generalized gauge choices are consistent with the desired amplitude form of the representation of cut data, and adjust the presentation of $\hat{\mathcal{O}}(\mathcal{A}_m)$ consistent with the gauge choices being made in the action. Such representation accounting will be discussed in the operator promotion section 5.3. Being careful about the compatibility between operators at different multiplicity will not be new for EFT. For amplitudes practitioners this is nothing more than the requirement that one's representation of the k -collapsed propagator contact-graph contributions depend crucially on decisions made in assigning cut data to fewer collapsed propagator contributions.

Now we will introduce a simple method to entirely isolate contact terms, multiplicity by multiplicity, via a systematic generalization of the method of maximal cuts. The method of maximal cuts can always be applied to isolate novel contact contributions directly to the full amplitude. In general this involves a factorial number of operations due to the $(2n-5)!!$ graphs contributing. If an amplitude can be decomposed into smaller gauge-invariant blocks such as color-ordered (or stripped) amplitudes all the better. Each ordered amplitude of multiplicity n has only $n(n-3)/2$ channels contributing to it, which means a Catalan number of cuts are required to isolate any novel contact. If however we are in the privileged position of having a color-dual representation we can simply look at the half-ladder graph that has $(n-3)$ propagators, so only requires at most $2^{(n-3)} - 1$ cuts to entirely isolate its contact. In more general local representations, the same recursive subtraction can be applied graph-by-graph to each cubic topology—though the cut isolation can be performed on each graph independently.

To do this, we make manifest the contribution of each cubic graph $g \in \mathcal{T}_3^m$ to the m -point contact term in the full amplitude. This is achieved by dressing the numerator $n(g)$ in

terms of a minimal independent basis of momentum invariants chosen such that all inverse propagators of the graph g appear as basis elements; this will allow us to identify the contact contribution as arising from terms proportional to the product of all inverse propagators of the graph. Take the set of $(m - 3)$ inverse-propagators $\Delta(g)$ of a cubic m -point graph, g , such that $d(g) = \prod_{i=1}^{|\Delta|} \Delta_i$. We decompose the numerator such that each term is classified by the specific subset $P \subset \Delta(g)$ of inverse propagators to which it is proportional,

$$\begin{aligned} n(g) &= \sum_{P \subset \Delta(g)} (P_{\times} \equiv P_1 \times \cdots \times P_{|P|}) n_P^{(|P|)}(g) \\ &= n^{(0)}(g) + \sum_{p \in \Delta(g)} p n_p^{(1)}(g) + \sum_{\{p,q\} \subset \Delta(g)} p \times q \times n_{\{p,q\}}^{(2)}(g) + \cdots + d(g) n_{\Delta}^{(|\Delta|)}(g). \end{aligned} \quad (4.8)$$

For generic kinematics in D dimensions (i.e., avoiding specific kinematic configurations leading to Gram determinant constraints among external momenta), the $m - 3$ inverse propagators $\Delta(g) = \{p_1, \dots, p_{m-3}\}$ associated with a specific m -point tree-level cubic graph topology g form a set of linearly independent kinematic variables. The presence of non-zero external masses m_i^2 does not diminish, and generally enhances, the space of available independent kinematic invariants, further ensuring the independence of these $m - 3$ internal propagator variables for generic external momenta. This linear independence guarantees that any polynomial numerator $n(g)$ can be uniquely decomposed into terms classified by their explicit dependence on products of these inverse propagators, as in Eq. (4.8). The coefficients $n_P^{(|P|)}(g)$ in this expansion are, by construction, polynomial functions of the kinematic invariants but are free of any further explicit factors from $\Delta(g) \setminus P$.

We therefore uniquely identify $n_P^{(|P|)}(g)$ such that it contains no factors of any inverse propagators present in the complement of P in Δ ; it can, however, contain additional powers of $P_i \in P$. This means that $n^{(0)}(g)$ has no dependence on any inverse propagators Δ_i relevant to the graph g , and in general, for higher-derivative contributions, $n_{\Delta}^{(|\Delta|)}(g)$ could contain any number of additional powers of inverse propagators. For an m -point amplitude organized in terms of cubic graphs, it should be clear that the m -point contact contribution to the full amplitude is given simply by summing over the $n_{\Delta(g)}^{(|\Delta(g)|)}(g)$ contribution from each graph.

$$\mathcal{C}_m = \sum_{g \in \Gamma_m^3} \frac{\left(d(g) n_{\Delta(g)}^{(|\Delta(g)|)}(g) \right)}{d(g)} \quad (4.9)$$

$$= \sum_{g \in \Gamma_m^3} n_{\Delta(g)}^{(|\Delta(g)|)}(g). \quad (4.10)$$

This decomposition of the kinematic numerators into contributions $n_P^{(|P|)}(g)$ identified by their inverse-propagator dependence has a recursive definition in terms of applying cut conditions (taking inverse-propagators to vanish for massless particles). The following subtraction recursively removes all subleading cut contributions from $n(g)$, isolating the term

proportional to the full set of inverse propagators in P .

Recursive extraction of contact numerators:

$$(P_\times) n_P^{(|P|)}(g) = n(g)|_{(\Delta \setminus P) \rightarrow 0} - \sum_{Q \subsetneq P} (Q_\times) n_Q^{(|Q|)}(g) \quad (4.11)$$

The first term corresponds to the graph numerator under cut conditions corresponding to all inverse propagators in the complement of P in Δ , written as $(\Delta \setminus P)$, taken on-shell. The second term is a sum over numerator contributions from $n_Q^{(|Q|)}(g)$ with inverse-propagator dependence Q that is a subset of P , but excluding P itself. We can then define the unique cut contributions to an amplitude, graph by graph, as:

$$\not{n}(g) = n(g) - d(g) n_\Delta^{(|\Delta|)}(g). \quad (4.12)$$

Now it is important to realize that what we mean here is the full numerator dressing over the propagators, so in gravitational theories, or double-copy theories where both copies contain kinematics, the situation is more subtle. Contact extraction must act on the full double-copy numerator $n\tilde{n}$, since cut conditions act on kinematics but not on color. While in Yang–Mills $\not{n}_{\text{YM}} = c \cdot \not{n}$, in gravity one must evaluate

$$\not{n}_{\text{GR}} = \mathbf{cut}(n\tilde{n}).$$

This ensures that $\mathcal{C}_m = \mathcal{A}_m - \not{\mathcal{A}}_m$ continues to isolate purely contact contributions even in the double-copy theories with two kinematic numerator weights.

We summarize the algorithm as follows:

Contact Extraction Algorithm

1. Choose a graph g in the amplitude.
2. Identify the set of inverse propagators $\Delta(g)$.
3. Write $n(g)$ as a sum over terms proportional to subsets $P \subset \Delta(g)$.
4. Recursively extract $n_P^{(|P|)}(g)$ using cut conditions on $\Delta \setminus P$.
5. The contact numerator is $n_\Delta^{(|\Delta|)}(g)$.
6. Sum over all graphs to get $\mathcal{C}_m$.

4.2 Five point example

As a schematic example, we consider a generic theory of a single particle type at five point tree level. There is a single cubic graph topology (the half ladder); we parameterize its numerator

dressing in the following form:

$$n(12345) = n \circ \begin{array}{c} 2 \quad 3 \quad 4 \\ | \quad | \quad | \\ \hline | \quad | \quad | \\ 1 \quad \quad 5 \end{array} = n^{(0)} + s_{12} n_{12}^{(1)} + s_{45} n_{45}^{(1)} + s_{12}s_{45} n_{12,45}^{(2)} \quad (4.13)$$

The generic kinematic functions $n^{(0)}$, $n_{12}^{(1)}$, $n_{45}^{(1)}$, and $n_{12,45}^{(2)}$ are defined such that dependence on the graph's inverse-propagators is explicit. Specifically, $n^{(0)}$ does not depend on s_{12} or s_{45} ; $n_{12}^{(1)}$ does not depend on s_{45} ; and $n_{45}^{(1)}$ does not depend on s_{12} .

This form is highly suggestive of the existence of contact operators in the theory. As the denominator for this graph is simply $d_g = s_{12}s_{45}$, the contributions arising from five-point contact operators are encoded entirely and exclusively by the function $n_{12,45}^{(2)}$. The other functions ($n^{(0)}$, $n_{12}^{(1)}$ and $n_{45}^{(1)}$) are associated with the contributions of this graph to particular cuts of the five point amplitude; in particular, the existence of non-vanishing $n_{12}^{(1)}$ and $n_{45}^{(1)}$ indicate the presence of a four-point contact. In particular theories, some of these functions can vanish entirely, corresponding to the absence of such contact operators; conversely, for theories with higher-derivative corrections of sufficient order, these functions can contain additional powers of their corresponding off-shell propagators.

There are three possible cuts of the corresponding amplitude $\mathcal{A}_5$ to which this particular graph dressing can contribute: the maximal cut $\mathcal{A}_3(1, 2, i)\mathcal{A}_3(-i, 3, j)\mathcal{A}_3(-j, 4, 5)$ and the two next-to-maximal cuts $\mathcal{A}_3(1, 2, i)\mathcal{A}_4(-i, 3, 4, 5)$ and $\mathcal{A}_4(1, 2, 3, i)\mathcal{A}_3(-i, 4, 5)$. We depict its contributions to such cuts as follows; expressions are found by applying relevant cut conditions to the original form of the numerator:

$$\mathcal{A}_3(1, 2, i)\mathcal{A}_3(-i, 3, j)\mathcal{A}_3(-j, 4, 5) : \quad n \circ \begin{array}{c} 2 \quad 3 \quad 4 \\ | \quad | \quad | \\ \hline | \quad | \quad | \\ 1 \quad \quad 5 \end{array} \quad \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \end{array} = n(12345)|_{\{s_{12}, s_{45}\} \rightarrow 0} \quad (4.14)$$

$$= n^{(0)} \quad (4.15)$$

$$\mathcal{A}_3(1, 2, i)\mathcal{A}_4(-i, 3, 4, 5) : \quad n \circ \begin{array}{c} 2 \quad 3 \quad 4 \\ | \quad | \quad | \\ \hline | \quad | \quad | \\ 1 \quad \quad 5 \end{array} \quad \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \end{array} = n(12345)|_{s_{12} \rightarrow 0} \quad (4.16)$$

$$= n^{(0)} + s_{45} n_{45}^{(1)} \quad (4.17)$$

$$\mathcal{A}_4(1, 2, 3, i) \mathcal{A}_3(-i, 4, 5) : \quad n \circ \begin{array}{c} 2 \quad 3 \quad 4 \\ | \quad | \quad | \\ \hline | \quad | \quad | \\ 1 \quad \quad 5 \end{array} = n(12345)|_{s_{45} \rightarrow 0} \quad (4.18)$$

$$= n^{(0)} + s_{12} n_{12}^{(1)} \quad (4.19)$$

where we indicate which cuts are responsible for the relevant kinematic limits. We see that all the kinematic information in the numerator except for the contact function $n_{12,45}^{(2)}$ contribute to these cuts! This is consistent with the removal of non-contact terms $\not{n}$ from the numerator n by carefully constructing combinations of cut conditions.

We can see clearly that simply adding all kinematic limits will result in overcounting of certain cut contributions. For example, each next-to-maximal cut contains data already specified by the maximal cut (namely, what we have parameterized as $n^{(0)}$). But as outlined in Equation (4.11) the approach of identifying each individual $n_P^{(|P|)}$ can occur recursively.

Starting with the s_{12} cut condition,

$$n \circ \begin{array}{c} 2 \quad 3 \quad 4 \\ | \quad | \quad | \\ \hline | \quad | \quad | \\ 1 \quad \quad 5 \end{array} = n^{(0)} + s_{45} n_{45}^{(1)} \quad (4.20)$$

and subtracting the information from the maximal cut isolates the $n_{45}^{(1)}$ term:

$$n \circ \begin{array}{c} 2 \quad 3 \quad 4 \\ | \quad | \quad | \\ \hline | \quad | \quad | \\ 1 \quad \quad 5 \end{array} - n \circ \begin{array}{c} 2 \quad 3 \quad 4 \\ | \quad | \quad | \\ \hline | \quad | \quad | \\ 1 \quad \quad 5 \end{array} = s_{45} n_{45}^{(1)} \quad (4.21)$$

We will denote this unique contribution by marking on the diagram the inverse propagator to which it is proportional, allowing us the diagrammatic illustration of the recursive definition of $s_{45} n_{45}^{(1)}$,

$$n \circ \begin{array}{c} 2 \quad 3 \quad 4 \\ | \quad | \quad | \\ \hline | \quad | \quad | \\ 1 \quad \quad 5 \end{array} \equiv n \circ \begin{array}{c} 2 \quad 3 \quad 4 \\ | \quad | \quad | \\ \hline | \quad | \quad | \\ 1 \quad \quad 5 \end{array} - n \circ \begin{array}{c} 2 \quad 3 \quad 4 \\ | \quad | \quad | \\ \hline | \quad | \quad | \\ 1 \quad \quad 5 \end{array} . \quad (4.22)$$

Similarly,

$$n \circ \begin{array}{c} 2 \quad 3 \quad 4 \\ | \quad | \quad | \\ \text{---} \\ | \quad | \quad | \\ 1 \quad 5 \end{array} \begin{array}{c} \bullet \\ \curvearrowright \\ \bullet \end{array} \begin{array}{c} \text{---} \\ | \\ \text{---} \end{array} \equiv n \circ \begin{array}{c} 2 \quad 3 \quad 4 \\ | \quad | \quad | \\ \text{---} \\ | \quad | \quad | \\ 1 \quad 5 \end{array} \begin{array}{c} \text{---} \\ | \\ \text{---} \end{array} - n \circ \begin{array}{c} 2 \quad 3 \quad 4 \\ | \quad | \quad | \\ \text{---} \\ | \quad | \quad | \\ 1 \quad 5 \end{array} \begin{array}{c} \text{---} \\ | \\ \text{---} \end{array} \begin{array}{c} \text{---} \\ | \\ \text{---} \end{array}. \quad (4.23)$$

And at last we come to the contact term $s_{12}s_{45}n_{12,45}^{(1)}$ whose recursive definition, following Equation (4.11), is simply:

$$n \circ \begin{array}{c} 2 \quad 3 \quad 4 \\ | \quad | \quad | \\ \text{---} \\ | \quad | \quad | \\ 1 \quad 5 \end{array} \begin{array}{c} \bullet \\ \curvearrowright \\ \bullet \end{array} \begin{array}{c} \bullet \\ \curvearrowright \\ \bullet \end{array} = n \circ \left(\begin{array}{c} 2 \quad 3 \quad 4 \\ | \quad | \quad | \\ \text{---} \\ | \quad | \quad | \\ 1 \quad 5 \end{array} - \begin{array}{c} 2 \quad 3 \quad 4 \\ | \quad | \quad | \\ \text{---} \\ | \quad | \quad | \\ 1 \quad 5 \end{array} \begin{array}{c} \bullet \\ \curvearrowright \\ \bullet \end{array} \begin{array}{c} \text{---} \\ | \\ \text{---} \end{array} - \begin{array}{c} 2 \quad 3 \quad 4 \\ | \quad | \quad | \\ \text{---} \\ | \quad | \quad | \\ 1 \quad 5 \end{array} \begin{array}{c} \text{---} \\ | \\ \text{---} \end{array} \begin{array}{c} \bullet \\ \curvearrowright \\ \bullet \end{array} - \begin{array}{c} 2 \quad 3 \quad 4 \\ | \quad | \quad | \\ \text{---} \\ | \quad | \quad | \\ 1 \quad 5 \end{array} \begin{array}{c} \text{---} \\ | \\ \text{---} \end{array} \begin{array}{c} \text{---} \\ | \\ \text{---} \end{array} \right). \quad (4.24)$$

$$n \circ \begin{array}{c} 2 \quad 3 \quad 4 \\ | \quad | \quad | \\ \text{---} \\ | \quad | \quad | \\ 1 \quad 5 \end{array} \begin{array}{c} \bullet \\ \curvearrowright \\ \bullet \end{array} \begin{array}{c} \bullet \\ \curvearrowright \\ \bullet \end{array} = n \circ \left(\begin{array}{c} 2 \quad 3 \quad 4 \\ | \quad | \quad | \\ \text{---} \\ | \quad | \quad | \\ 1 \quad 5 \end{array} - \begin{array}{c} 2 \quad 3 \quad 4 \\ | \quad | \quad | \\ \text{---} \\ | \quad | \quad | \\ 1 \quad 5 \end{array} \begin{array}{c} \bullet \\ \curvearrowright \\ \bullet \end{array} \begin{array}{c} \text{---} \\ | \\ \text{---} \end{array} - \begin{array}{c} 2 \quad 3 \quad 4 \\ | \quad | \quad | \\ \text{---} \\ | \quad | \quad | \\ 1 \quad 5 \end{array} \begin{array}{c} \text{---} \\ | \\ \text{---} \end{array} \begin{array}{c} \bullet \\ \curvearrowright \\ \bullet \end{array} - \begin{array}{c} 2 \quad 3 \quad 4 \\ | \quad | \quad | \\ \text{---} \\ | \quad | \quad | \\ 1 \quad 5 \end{array} \begin{array}{c} \text{---} \\ | \\ \text{---} \end{array} \begin{array}{c} \text{---} \\ | \\ \text{---} \end{array} \right).$$

4.3 Six point example

We continue now to six-point with:

$$n(123456) = n \circ \begin{array}{c} 2 \quad 3 \quad 4 \quad 5 \\ | \quad | \quad | \quad | \\ \hline | \quad \quad \quad | \\ 1 \quad \quad \quad 6 \end{array} \quad (4.25)$$

We again begin by parameterizing the numerator so that dependence on inverse propagators is made manifest:

$$\begin{aligned} n(123456) = & n^{(0)} + s_{12} n_{12}^{(1)} + s_{123} n_{123}^{(1)} + s_{56} n_{56}^{(1)} + s_{12} s_{123} n_{12,123}^{(2)} \\ & + s_{12} s_{56} n_{12,56}^{(2)} + s_{123} s_{56} n_{123,56}^{(2)} + s_{12} s_{123} s_{56} n_{12,123,56}^{(3)}. \end{aligned} \quad (4.26)$$

Using the graphical conventions explained in the five point example, this can be written,

$$\begin{aligned} n \circ \begin{array}{c} 2 \quad 3 \quad 4 \quad 5 \\ | \quad | \quad | \quad | \\ \hline | \quad \quad \quad | \\ 1 \quad \quad \quad 6 \end{array} &= n \circ \left(\begin{array}{c} 2 \quad 3 \quad 4 \quad 5 \\ | \quad | \quad | \quad | \\ \hline | \quad \quad \quad | \\ 1 \quad \quad \quad 6 \end{array} \right. \\ &+ \begin{array}{c} 2 \quad 3 \quad 4 \quad 5 \\ | \quad | \quad | \quad | \\ \hline | \quad \quad \quad | \\ 1 \quad \quad \quad 6 \end{array} \begin{array}{c} \text{green arc } (2,3) \\ \text{red dashed } (3,4) \\ \text{red dashed } (4,5) \end{array} \\ &+ \begin{array}{c} 2 \quad 3 \quad 4 \quad 5 \\ | \quad | \quad | \quad | \\ \hline | \quad \quad \quad | \\ 1 \quad \quad \quad 6 \end{array} \begin{array}{c} \text{red dashed } (2,3) \\ \text{green arc } (3,4) \\ \text{red dashed } (4,5) \end{array} \\ &+ \begin{array}{c} 2 \quad 3 \quad 4 \quad 5 \\ | \quad | \quad | \quad | \\ \hline | \quad \quad \quad | \\ 1 \quad \quad \quad 6 \end{array} \begin{array}{c} \text{red dashed } (2,3) \\ \text{red dashed } (3,4) \\ \text{green arc } (4,5) \end{array} \\ &+ \begin{array}{c} 2 \quad 3 \quad 4 \quad 5 \\ | \quad | \quad | \quad | \\ \hline | \quad \quad \quad | \\ 1 \quad \quad \quad 6 \end{array} \begin{array}{c} \text{red dashed } (2,3) \\ \text{green arc } (3,4) \\ \text{green arc } (4,5) \end{array} \\ &+ \begin{array}{c} 2 \quad 3 \quad 4 \quad 5 \\ | \quad | \quad | \quad | \\ \hline | \quad \quad \quad | \\ 1 \quad \quad \quad 6 \end{array} \begin{array}{c} \text{green arc } (2,3) \\ \text{red dashed } (3,4) \\ \text{green arc } (4,5) \end{array} \\ &+ \begin{array}{c} 2 \quad 3 \quad 4 \quad 5 \\ | \quad | \quad | \quad | \\ \hline | \quad \quad \quad | \\ 1 \quad \quad \quad 6 \end{array} \begin{array}{c} \text{green arc } (2,3) \\ \text{green arc } (3,4) \\ \text{red dashed } (4,5) \end{array} \\ &\left. + \begin{array}{c} 2 \quad 3 \quad 4 \quad 5 \\ | \quad | \quad | \quad | \\ \hline | \quad \quad \quad | \\ 1 \quad \quad \quad 6 \end{array} \begin{array}{c} \text{green arc } (2,3) \\ \text{green arc } (3,4) \\ \text{green arc } (4,5) \end{array} \right) \quad (4.27) \end{aligned}$$

The recursive definition of the contact is then simply given by:

$$\begin{aligned}
n \circ \left(\begin{array}{c} 2 \quad 3 \quad 4 \quad 5 \\ | \quad | \quad | \quad | \\ \text{---} \\ | \quad | \quad | \quad | \\ 1 \quad \quad \quad 6 \end{array} \right) &= n \circ \left(\begin{array}{c} 2 \quad 3 \quad 4 \quad 5 \\ | \quad | \quad | \quad | \\ \text{---} \\ | \quad | \quad | \quad | \\ 1 \quad \quad \quad 6 \end{array} \right) - \\
&- \left(\begin{array}{c} 2 \quad 3 \quad 4 \quad 5 \\ | \quad | \quad | \quad | \\ \text{---} \\ | \quad | \quad | \quad | \\ 1 \quad \quad \quad 6 \end{array} \right) - \left(\begin{array}{c} 2 \quad 3 \quad 4 \quad 5 \\ | \quad | \quad | \quad | \\ \text{---} \\ | \quad | \quad | \quad | \\ 1 \quad \quad \quad 6 \end{array} \right) - \left(\begin{array}{c} 2 \quad 3 \quad 4 \quad 5 \\ | \quad | \quad | \quad | \\ \text{---} \\ | \quad | \quad | \quad | \\ 1 \quad \quad \quad 6 \end{array} \right) \\
&- \left(\begin{array}{c} 2 \quad 3 \quad 4 \quad 5 \\ | \quad | \quad | \quad | \\ \text{---} \\ | \quad | \quad | \quad | \\ 1 \quad \quad \quad 6 \end{array} \right) - \left(\begin{array}{c} 2 \quad 3 \quad 4 \quad 5 \\ | \quad | \quad | \quad | \\ \text{---} \\ | \quad | \quad | \quad | \\ 1 \quad \quad \quad 6 \end{array} \right) - \left(\begin{array}{c} 2 \quad 3 \quad 4 \quad 5 \\ | \quad | \quad | \quad | \\ \text{---} \\ | \quad | \quad | \quad | \\ 1 \quad \quad \quad 6 \end{array} \right) \\
&- \left(\begin{array}{c} 2 \quad 3 \quad 4 \quad 5 \\ | \quad | \quad | \quad | \\ \text{---} \\ | \quad | \quad | \quad | \\ 1 \quad \quad \quad 6 \end{array} \right) \quad (4.28)
\end{aligned}$$

with e.g.

$$\begin{aligned}
\left(\begin{array}{c} 2 \quad 3 \quad 4 \quad 5 \\ | \quad | \quad | \quad | \\ \text{---} \\ | \quad | \quad | \quad | \\ 1 \quad \quad \quad 6 \end{array} \right) &= s_{12} s_{123} n_{12,123}^{(2)} \\
&= n \circ \left(\begin{array}{c} 2 \quad 3 \quad 4 \quad 5 \\ | \quad | \quad | \quad | \\ \text{---} \\ | \quad | \quad | \quad | \\ 1 \quad \quad \quad 6 \end{array} \right) - \left(\begin{array}{c} 2 \quad 3 \quad 4 \quad 5 \\ | \quad | \quad | \quad | \\ \text{---} \\ | \quad | \quad | \quad | \\ 1 \quad \quad \quad 6 \end{array} \right) \\
&- \left(\begin{array}{c} 2 \quad 3 \quad 4 \quad 5 \\ | \quad | \quad | \quad | \\ \text{---} \\ | \quad | \quad | \quad | \\ 1 \quad \quad \quad 6 \end{array} \right) - \left(\begin{array}{c} 2 \quad 3 \quad 4 \quad 5 \\ | \quad | \quad | \quad | \\ \text{---} \\ | \quad | \quad | \quad | \\ 1 \quad \quad \quad 6 \end{array} \right) \\
&= n(123456)|_{s_{56} \rightarrow 0} - \left(s_{12} n_{12}^{(1)} \equiv n(123456)|_{\{s_{123}, s_{56}\} \rightarrow 0} - n^{(0)} \right) \\
&- \left(s_{123} n_{123}^{(1)} \equiv n(123456)|_{\{s_{12}, s_{56}\} \rightarrow 0} - n^{(0)} \right) - n^{(0)}. \quad (4.29)
\end{aligned}$$

4.4 More on gravitational contact isolation

We take this as an opportunity to detail how all-multiplicity gravitational ($\sqrt{-g}R$) contact terms can appear via double copy. The basic mechanism is conceptually straightforward, but it is instructive to see it illustrated explicitly in the context of local color-dual representations of Yang–Mills theory.

We will depict the emergence of the gravitational contacts required for linearized diffeomorphism below, first noting how $n_{\text{YM}}(12345)$ behaves for the half-ladder with at most a 4-point contact contributing singularly per channel,

$$n_{\text{YM}}(12345) = n_{\text{YM}}^{(0)} + s_{12} n_{\text{YM}12}^{(1)} + s_{45} n_{\text{YM}45}^{(1)} \quad (4.30)$$

$$= \left(\begin{array}{c} 2 \quad 3 \quad 4 \\ | \quad | \quad | \\ \text{---} \\ | \quad | \quad | \\ 1 \quad 5 \end{array} \begin{array}{c} \text{red dashed} \\ \text{red dashed} \end{array} + \begin{array}{c} 2 \quad 3 \quad 4 \\ | \quad | \quad | \\ \text{---} \\ | \quad | \quad | \\ 1 \quad 5 \end{array} \begin{array}{c} \text{red dashed} \\ \text{green arc} \end{array} + \begin{array}{c} 2 \quad 3 \quad 4 \\ | \quad | \quad | \\ \text{---} \\ | \quad | \quad | \\ 1 \quad 5 \end{array} \begin{array}{c} \text{green arc} \\ \text{red dashed} \end{array} \right). \quad (4.31)$$

The double-copy numerator for gravity then is quite simply of the following form,

$$n_{\text{GR}} = n_{\text{YM}} \times \widetilde{n_{\text{YM}}} \quad (4.32)$$

$$= \left(\begin{array}{c} 2 \quad 3 \quad 4 \\ | \quad | \quad | \\ \text{---} \\ | \quad | \quad | \\ 1 \quad 5 \end{array} \begin{array}{c} \text{red dashed} \\ \text{red dashed} \end{array} + \begin{array}{c} 2 \quad 3 \quad 4 \\ | \quad | \quad | \\ \text{---} \\ | \quad | \quad | \\ 1 \quad 5 \end{array} \begin{array}{c} \text{red dashed} \\ \text{green arc} \end{array} + \begin{array}{c} 2 \quad 3 \quad 4 \\ | \quad | \quad | \\ \text{---} \\ | \quad | \quad | \\ 1 \quad 5 \end{array} \begin{array}{c} \text{green arc} \\ \text{red dashed} \end{array} \right) \times \quad (4.33)$$

$$\left(\begin{array}{c} 2 \quad 3 \quad 4 \\ | \quad | \quad | \\ \text{---} \\ | \quad | \quad | \\ 1 \quad 5 \end{array} \begin{array}{c} \text{red dashed} \\ \text{red dashed} \end{array} + \begin{array}{c} 2 \quad 3 \quad 4 \\ | \quad | \quad | \\ \text{---} \\ | \quad | \quad | \\ 1 \quad 5 \end{array} \begin{array}{c} \text{red dashed} \\ \text{green arc} \end{array} + \begin{array}{c} 2 \quad 3 \quad 4 \\ | \quad | \quad | \\ \text{---} \\ | \quad | \quad | \\ 1 \quad 5 \end{array} \begin{array}{c} \text{green arc} \\ \text{red dashed} \end{array} \right). \quad (4.34)$$

The novel five-point gravitational contact arises from the cross terms that in concert from both sides cancel all propagators,

$$n_{\text{GR}} \circ \begin{array}{c} 2 \quad 3 \quad 4 \\ | \quad | \quad | \\ \text{---} \\ | \quad | \quad | \\ 1 \quad 5 \end{array} \begin{array}{c} \text{green arc} \\ \text{green arc} \end{array} = n_{\text{GR}} - \not{n}_{\text{GR}} \quad (4.35)$$

$$\begin{aligned}
&= \begin{array}{c} \text{Diagram 1: A horizontal line with vertices 1, 5, 1, 5. Above the line are vertical lines labeled 2, 3, 4, 2, 3, 4. A red dashed line is between 2 and 3. A green arc connects 3 and 4. A wavy line is above the second 3.} \\ \text{Diagram 2: A horizontal line with vertices 1, 5, 1, 5. Above the line are vertical lines labeled 2, 3, 4, 2, 3, 4. A red dashed line is between 3 and 4. A green arc connects 2 and 3. A wavy line is above the second 3.} \end{array} \\
&+ \begin{array}{c} \text{Diagram 3: A horizontal line with vertices 1, 5, 1, 5. Above the line are vertical lines labeled 2, 3, 4, 2, 3, 4. A red dashed line is between 2 and 3. A green arc connects 3 and 4. A wavy line is above the second 3.} \\ \text{Diagram 4: A horizontal line with vertices 1, 5, 1, 5. Above the line are vertical lines labeled 2, 3, 4, 2, 3, 4. A red dashed line is between 3 and 4. A green arc connects 2 and 3. A wavy line is above the second 3.} \end{array} \quad (4.36)
\end{aligned}$$

$$= s_{12}s_{45} \left(n_{12}^{(1)} \tilde{n}_{45}^{(1)} + n_{45}^{(1)} \tilde{n}_{12}^{(1)} \right). \quad (4.37)$$

Similarly at six points, a color-dual Yang-Mills numerator can have the following form with at most two collapsed propagators at a time,

$$\begin{aligned}
n^{\text{YM}} \circ \begin{array}{c} \text{Diagram 1: A horizontal line with vertices 1, 6. Above the line are vertical lines labeled 2, 3, 4, 5.} \end{array} &= \left(\begin{array}{c} \text{Diagram 2: A horizontal line with vertices 1, 6. Above the line are vertical lines labeled 2, 3, 4, 5. Three red dashed lines are between 2-3, 3-4, and 4-5.} \\ \text{Diagram 3: A horizontal line with vertices 1, 6. Above the line are vertical lines labeled 2, 3, 4, 5. A green arc connects 2 and 3. Two red dashed lines are between 3-4 and 4-5.} \\ \text{Diagram 4: A horizontal line with vertices 1, 6. Above the line are vertical lines labeled 2, 3, 4, 5. A green arc connects 3 and 4. One red dashed line is between 2-3.} \\ \text{Diagram 5: A horizontal line with vertices 1, 6. Above the line are vertical lines labeled 2, 3, 4, 5. A green arc connects 4 and 5. No red dashed lines.} \end{array} \right. \\
&+ \begin{array}{c} \text{Diagram 6: A horizontal line with vertices 1, 6. Above the line are vertical lines labeled 2, 3, 4, 5. A green arc connects 2 and 3. Two red dashed lines are between 3-4 and 4-5.} \\ \text{Diagram 7: A horizontal line with vertices 1, 6. Above the line are vertical lines labeled 2, 3, 4, 5. A green arc connects 3 and 4. One red dashed line is between 2-3.} \\ \text{Diagram 8: A horizontal line with vertices 1, 6. Above the line are vertical lines labeled 2, 3, 4, 5. A green arc connects 4 and 5. No red dashed lines.} \end{array} \\
&+ \left(\begin{array}{c} \text{Diagram 9: A horizontal line with vertices 1, 6. Above the line are vertical lines labeled 2, 3, 4, 5. Two green arcs connect 2-3 and 3-4. One red dashed line is between 4-5.} \\ \text{Diagram 10: A horizontal line with vertices 1, 6. Above the line are vertical lines labeled 2, 3, 4, 5. Two green arcs connect 3-4 and 4-5. One red dashed line is between 2-3.} \\ \text{Diagram 11: A horizontal line with vertices 1, 6. Above the line are vertical lines labeled 2, 3, 4, 5. Two green arcs connect 2-3 and 4-5. No red dashed lines.} \end{array} \right) \quad (4.38)
\end{aligned}$$

Notice for example the appearance of what appear to be five-point contacts in say the last contact graph which collapses propagators for $(k_1 + k_2)^2$ and $(k_1 + k_2 + k_3)^2$. These are of course spurious and cancel against similar terms in the sum over graphs contributing to gauge-invariant objects such as the full-amplitude or even ordered cuts like $A(1234l)A(-l56)$. Such terms can contribute in cross-terms to the gravitational contact depending on the generalized gauge and so must in general still be considered in all topologies that contribute to the full amplitude.

4.5 Spurious non-locality

It is possible to formally dress kinematic numerators of cubic graphs with functions that may contain poles. A fine example would be the virtuous numerators of ref. [37]. Introducing spurious non-locality in numerator dressings is allowed as long as any gauge invariant construction involving the numerators, such as ordered amplitudes, has poles only where physical. Virtuous representations where external states are encoded in ordered amplitudes make this required property manifest. In such cases, however, the above individual numerator contact isolation procedure is not guaranteed to work. One can resort to method of maximal cuts as applied to local gauge-invariant blocks – ordered amplitudes for gauge theories, and full amplitudes otherwise. We hope it does not escape the reader that this is literally the equivalent of building up a local representation.

5 The Operator Promotion

We will now to establish a mapping from momentum space amplitudes to position space operators. More sharply, for some term n contributing to an amplitude $\mathcal{A}_m$, we will prescribe an operator promotion $n \rightarrow \hat{\mathcal{O}}(n) \equiv \hat{n}$ to an m -field operator $\hat{n}$. This operator will by construction yield a Feynman rule that dresses the m -point vertex, and thus first shows up as a contribution to the m -point contact diagram. We must now establish which contributions n to the amplitude we need to specify in order to describe the entire theory at said multiplicity.

Graph-organized double-copy amplitudes, as established, are written as sum over contributions from all relevant distinct cubic graphs Γ . We will write this contribution from each graph $g \in \Gamma$ as $Q(g)$, so $\mathcal{A} = \sum_{g \in \Gamma} Q(g)$. Such a contribution $Q(g)$ is generically a function of particle momenta k_i , external state wavefunctions³, and gauge group structure constants f^{abc} and generators T_{ij}^a . We can map this back to position space via Fourier transform, replacing momenta k^μ with derivatives ∂^μ and replacing wavefunctions with appropriately Lorentz indexed external fields, e.g. $\varepsilon_\mu \rightarrow A_\mu$.

While generically there are many distinct cubic graphs relevant to a given set of external states, we only need to consider the contribution of one graph per basis topology when writing down the operator. Schematically, we can understand this as follows.

Let us initially restrict ourselves to an amplitude with only a single cubic graph topology; for example, Yang Mills at four points. All three graphs $\Gamma = \{g_s, g_t, g_u\}$ can be considered permutations of the external leg labels on the first graph (chosen arbitrarily to be the s -channel). If we label g_s as $g(1234)$, we can identify g_t as $g(4123)$ and g_u as $g(4231)$. These are only 3 of 24 permutations of the labels (1234); all remaining permutations are isomorphic to one of the three channels. We can instead choose to write the full amplitude as a sum over all such permutations, modded out by the appropriate symmetry factor:

$$\mathcal{A}_4 = \frac{1}{8} \sum_{\sigma \in S_4} Q(g(\sigma)) \quad (5.1)$$

³*Exempli gratia*, polarizations $\varepsilon_i(k_i)$; spinors $u_i(k_i)$, $v_i(k_i)$; 1; etcetera

Consider the operator obtained from mapping just the s -channel contribution $Q(g(1234))$ to position space as specified above. This four-field operator will give rise to a four-point vertex Feynman rule $\mathcal{V}_4$. Feynman rules, at their core, are obtained by performing functional derivatives on momentum-space operators, which has the effect of summing over all permutations of labels on identical fields:

$$\mathcal{V}_4 \varepsilon_1 \varepsilon_2 \varepsilon_3 \varepsilon_4 = \sum_{\sigma \in S_4} Q(g(\sigma)) \quad (5.2)$$

Feynman rules inherently sum over all such permutations, which we can identify as precisely identical to the sum over all graphs in the full amplitude $\mathcal{A}_4$. Thus, we need only consider the contribution from *a single graph for each topology* in order to write an operator that encapsulates the contributions from *all graphs*.

Color-kinematics duality simplifies this description further. Color-kinematics Jacobi identities relate the numerator dressings of distinct cubic graph topologies; all such dressings can be written in terms of linear relations upon relabelings of those of the basis graphs. As a result, from a single graph dressing for each *basis* topology, we can write down the full operator that gives rise to the complete full amplitude $\mathcal{A}$.

In this approach, each multiplicity/set of external states is considered separately and gives rise to a (generically non-local) contact operator. This contrasts significantly with the typical approach, where there can be a limited number of low-point operators, and higher multiplicity amplitudes are generated by sewing together such vertices with internal propagators. By focusing on each multiplicity separately, we are able to preserve a manifestly color-dual structure for the operators. This allows us to see clearly that the operators can be double copied to write down operators for different theories.

5.1 Double-Copy Structure

How do we represent double-copy structure at the level of fields? We want both sides of the double copy to be on equal footing, so each conceptual copy is assigned field content. This decomposition is most easily understood through some examples for common double-copy theories, summarized in Table 1. In this notation, $\otimes$ signifies the conceptual double-copy product, and the fields on the right-hand side represent the constituent ‘single-copy’ or root structures.

For the graviton, this decomposition is of course very familiar. For something like the Yang-Mills gauge field, $A_\mu^a(x)$, our decomposition implies that the color information (carried by φ^a) is conceptually separated from the vector kinematic structure (carried by A_μ). The scalar fields φ^a will appear in the operator promotions of color factors c_g , while the colorless vector structures A_μ will correspond to kinematic numerators n_g . Similarly, the bi-adjoint scalar field $\phi^{a\tilde{a}}(x)$ is decomposed into two distinct color-carrying scalar structures. This ‘factorized’ field representation is advantageous as it allows us to systematically construct operators for a plethora of theories related by double copy by mixing and matching these constituent operator types, as long as one is diligent about the bookkeeping of symmetries

Table 1: Color-dual double-copy field decompositions for various theories. A_μ represents a generic vector (kinematic) structure, φ^a an adjoint scalar (color) structure, and π a generic scalar (kinematic, e.g., NLSM pion) structure. Tilded objects denote distinct copies if necessary.

Double-Copied Field	Symbol	Root 1	Root 2	Interpretation
Graviton	$h_{\mu\nu}(x)$	$A_\mu(x)$	$A_\nu(x)$	Vector $\otimes$ Vector
Gauge Boson (YM)	$A_\mu^a(x)$	$A_\mu(x)$	$\varphi^a(x)$	Vector $\otimes$ Scalar (color)
Bi-adjoint Scalar	$\phi^{a\tilde{a}}(x)$	$\varphi^a(x)$	$\tilde{\varphi}^{\tilde{a}}(x)$	Scalar ₁ $\otimes$ Scalar ₂
Pion (NLSM-like)	$\pi^a(x)$	$\varphi^a(x)$	$\pi(x)$	Scalar (∂^2) $\otimes$ Scalar (color)
Born-Infeld Photon	$A_\mu^{\text{BI}}(x)$	$A_\mu(x)$	$\pi(x)$	Vector $\otimes$ Scalar (∂^2)
Z-theory	$Z^{aA}(x)$	$\varphi^a(x)$	$z^A(x)$	Scalar (color) $\otimes$ α'^∞ -Scalar (color)
Open Superstring (vector)	$\text{OSS}_\mu^A(x)$	$A_\mu(x)$	$z^A(x)$	Vector $\otimes$ α'^∞ -Scalar (color)
(DF) ² +YM	$B_\mu^a(x)$	$B_\mu(x)$	$\varphi^a(x)$	Vector $\otimes$ Scalar (color)
Open Bosonic String (vector)	$\text{OBS}_\mu^A(x)$	$B_\mu(x)$	$z^A(x)$	Vector $\otimes$ α'^∞ -scalar (color)
Closed Superstring (graviton)	$\text{CSS}_{\mu\nu}$	$A_\mu(x)$	$(A)_\mu^{\text{sv}}(x)$	Vector $\otimes$ α'^∞ -Vector
Heterotic String (graviton)	$\text{HS}_{\mu\nu}$	$B_\mu(x)$	$(A)_\mu^{\text{sv}}(x)$	Vector $\otimes$ α'^∞ -Vector

and physical state projections. It would be an error to assign dilaton or antisymmetric 2-form operators to gravity – as it would be a mistake at the level of amplitudes. This places the burden of correct book-keeping on physical observables—where projection to desired physical states is part and parcel to unitarity methods. Some of the more exotic fields (α'^∞) involve non-local interactions (at least, infinite towers of higher-derivative operators), and yet we can still understand their tree-level amplitudes at least as field-theory double-copies, and can use the approach of this current paper to write operators to reproduce those scattering amplitudes. In the spirit of sheer pragmatism of writing down operators which generate the amplitudes at hand, we will defer addressing almost all subtleties of this construction to former work (validity of the theories as quantum field theories, unitarity, etc), and simply discuss examples of such constructions in section 6.2.

Let us address a critical point right here. It is perfectly acceptable to regard the above mapping as merely a creative way of labeling well known and familiar fields. That would be the most conservative stance and entirely appropriate if all one wished to do was have an elegant way of classifying and constructing operators. There is a path forward to a state-level understanding and appreciation of double-copy that is however now available to us, and one that we will address directly in section 7.

For Yang Mills theory, this means we will write down separate operators for the color factor c_g , the kinematic numerator dressing n_g , and the inverse propagator $1/d_g$, and then multiply these together to constitute the overall graph contribution $Q(g)$. For n_g , the steps to arriving at an operator are fairly familiar: convert polarizations to vector fields, and convert momenta to derivatives acting on the appropriate vector fields.

The color factor c_g is simply a contraction of structure constants for the relevant gauge group, so it can be tempting to identify the corresponding operator as simply this contraction,

as it needs no conversion from momentum to position space. However, we would like our operator promotions to be as theory-agnostic as possible: each numerator factor, no matter its nature, should have an associated set of fields. We will choose to write down the c_g operator promotion by inserting scalar fields carrying adjoint color indices. For example, at four points,

$$\hat{c}_g = f^{abe} f^{ecd} \varphi^a(x) \varphi^b(x) \varphi^c(x) \varphi^d(x) \quad (5.3)$$

Finally, we promote the inverse propagator by generically allowing the Lagrangian to be non-local. We want to be able to introduce an operator that targets propagator structure associated with a Feynman graph. For example, consider an s -channel, $(k_1 + k_2)^2$, four point graph. We will want a propagator operator $\frac{1}{\hat{s}}$ that yields a Feynman rule that generates propagator kinematics, schematically,

$$\frac{1}{8} \frac{\hat{c}_s \hat{n}_s}{\hat{s}} \rightarrow \frac{c_s n_s}{s} + \frac{c_t n_t}{t} + \frac{c_u n_u}{u}. \quad (5.4)$$

All such non-locality will be canceled upon subtraction of cut-terms, but it will be key to our organization to be able to target specific graphs. It turns out the writing down of such operators is greatly facilitated by introducing a redundancy in spacetime variables as we describe in the next section.

5.2 Operators can look like Amplitudes

As theories increase in complexity (especially as higher-derivative corrections are included), the process of mapping momenta in numerators to derivatives in operators becomes quite verbose. A key issue is that k_i^μ must be translated to ∂^μ acting specifically upon field ϕ_i . This can obfuscate the structure of the operator, as one must match up Lorentz indices of derivatives acting upon different fields to realize the corresponding Lorentz invariant dot product that will arise in the amplitude.

We introduce the following measure:

$$\boxed{\int \mathcal{D}_m \equiv \int \prod_i^m d^D x_i d^D \tilde{x}_i \delta^D(x - x_i) \delta^D(x - \tilde{x}_i)} \quad (5.5)$$

where we include both x_i and $\tilde{x}_i$ as auxiliary coordinates for the left and right sides of the double-copy inspired field decomposition.

We will also write:

$$\boxed{\phi_m^a(\mathbf{q}) \equiv \phi^{a_1}(q_1) \cdots \phi^{a_m}(q_m).} \quad (5.6)$$

As an example, let us consider the four point amplitude for the non-linear sigma model.

The pion fields carry color, so in our double-copy framework we will annotate this as:

$$\pi^a(x) \equiv (\phi^a \otimes \pi)(x) = \int \mathcal{D}_1 \phi^a(\tilde{x}_1) \otimes \pi(x_1) \quad (5.7)$$

The amplitude is given in terms of cubic graphs via some color-factor, kinematic numerator, and propagator as:

$$\mathcal{A}_4 = \frac{c_s n_s}{s} + \text{perms} . \quad (5.8)$$

We will see that we can write the position space operator that generates this amplitude as follows:

$$\mathcal{O}_4 = \int \mathcal{D}_4 \hat{\mathcal{A}}_4 \pi_4^a(x) . \quad (5.9)$$

with

$$\hat{\mathcal{A}}_4 \equiv \frac{1}{8} \frac{\hat{c}_s \hat{n}_s}{\hat{s}} . \quad (5.10)$$

To see what the hatted objects will mean let's first be very clear about the building blocks of the amplitude written in terms of cubic graphs. We have a color factor for the s -channel graph given by:

$$c_s \equiv f^{a_1 a_2 e} f^{e a_3 a_4} . \quad (5.11)$$

a propagator $s \equiv (k_1 + k_2)^2$, and a numerator factor

$$n_s = \frac{1}{3} s(u - t) \quad (5.12)$$

We note the presence of the inverse propagator s in the kinematic numerator means that we are really describing a four-point contact amplitude. Indeed the entire amplitude can be written as:

$$\mathcal{A}_4 = c_s(u - t) + c_t(u - s) + c_u(s - t) \quad (5.13)$$

$$= c_s(2u - s - t) + c_u(-u - t) \quad (5.14)$$

$$= c_s u + (s \leftrightarrow u) . \quad (5.15)$$

We will stick with our graph organization at the moment as it allows the duality between color and kinematics to be manifest, and will make it straightforward to double-copy to generate Born-Infeld shortly.

One could imagine writing a position space operator that generates something like n_s at four-points (recalling that $s_{ij} = (k_i + k_j)^2 = 2k_i \cdot k_j$ for on-shell $k_i^2 = 0$) as follows,

$$\begin{aligned} \mathcal{O} = 4 \times \frac{1}{3} [& (\partial_\mu \partial_\nu \pi^{a_1}(x)) (\partial^\mu \pi^{a_2}(x)) (\partial^\nu \pi^{a_3}(x)) (\pi^{a_4}(x)) \\ & - (\partial_\mu \partial_\nu \pi^{a_1}(x)) (\partial^\mu \pi^{a_2}(x)) (\pi^{a_3}(x)) (\partial^\nu \pi^{a_4}(x))] . \end{aligned} \quad (5.16)$$

Arguably this is far less pleasant than the amplitude. We can make the structure more clear

by borrowing a trick of ref. [38] and assigning each field its own auxiliary position variable:

$$\phi_i(x) = \int d^d x_i \delta^{(d)}(x - x_i) \phi_i(x_i). \quad (5.17)$$

As such, $k_{i,\mu}$ can be mapped to:

$$\partial_\mu \phi_i(x) = \int d^d x_i \delta^{(d)}(x - x_i) \left(\frac{\partial}{\partial x_i^\mu} \right) \phi_i(x_i). \quad (5.18)$$

We will define $\partial_{i,\mu} \equiv \partial/\partial x_i^\mu$ for notational convenience. So far, this is, of course, a trivial statement. However, the benefits of this become clear when we consider the full product of fields appearing in an operator. For the NLSM numerator, we can rewrite the operator promotion using these auxiliary spacetime coordinates as follows:

$$\mathcal{O} = \frac{4}{3} \int \prod_{i=1}^4 \left(d^d x_i \delta^{(d)}(x - x_i) \right) [\quad (5.19)$$

$$(\partial_{1,\mu} \partial_{1,\nu} \pi_1^{a_1}(x_1)) (\partial_2^\mu \pi_2^{a_2}(x_2)) (\partial_3^\nu \pi_3^{a_3}(x_3)) (\pi_4^{a_4}(x_4)) \quad (5.20)$$

$$- (\partial_{1,\mu} \partial_{1,\nu} \pi_1^{a_1}(x_1)) (\partial_2^\mu \pi_2^{a_2}(x_2)) (\pi_3^{a_3}(x_3)) (\partial_4^\nu \pi_4^{a_4}(x_4))] \quad (5.21)$$

Since $\partial_{i,\mu}$ only acts on the i^{th} field, we can factor all the partial derivatives out from the product of fields (and dot them together into Lorentz invariants):

$$\int \mathcal{D}_4 \hat{n} \pi_4^a(\mathbf{x}) = \frac{4}{3} \int \prod_{i=1}^4 \left(d^d x_i \delta^{(d)}(x - x_i) \right) (\partial_1 \cdot \partial_2) ((\partial_1 \cdot \partial_3) - (\partial_1 \cdot \partial_3)) \prod \pi_i^{a_i}(x_i) \quad (5.22)$$

Here, we can identify the $s(u - t)$ structure immediately by inspection! Indeed let's sharpen the game by introducing

$$\boxed{\hat{s}_{ij} \equiv (\partial_i + \partial_j)^2}, \quad (5.23)$$

with then $\hat{s} \equiv \hat{s}_{12}$, $\hat{t} \equiv \hat{s}_{23}$, $\hat{u} \equiv \hat{s}_{13}$, so that:

$$\hat{n} = \hat{s} (\hat{t} - \hat{u}). \quad (5.24)$$

With $\hat{c}_s^{\mathbf{a}} \equiv c_s^{\mathbf{a}}$, our amplitude is generated by an operator written:

$$\mathcal{O}_4 = \int \mathcal{D}_4 \frac{1}{8} \frac{\hat{c}_s^{\mathbf{a}} \hat{n}_s}{\hat{s}} \pi_4^a(\mathbf{x}) \quad (5.25)$$

$$= \int \mathcal{D}_4 \frac{1}{8} \hat{c}_s^{\mathbf{a}} (\hat{t} - \hat{u}) \pi_4^a(\mathbf{x}) \quad (5.26)$$

How should we maximize the ease of translating vector kinematic numerators to operators? Let vector fields A_μ carry their own Lorentz indices, so our m -field vector product,

$A_{\mu,m}^a(q)$, will be defined as:

$$\boxed{A_{\mu,m}^a(q) \equiv A_{\mu_1}^{a_1}(q_1) \cdots A_{\mu_m}^{a_m}(q_m)} . \quad (5.27)$$

Polarizations $\varepsilon_{\mu_i}(k_i)$ appearing in Lorentz invariants in scattering amplitude expressions are then simply removed to allow the metric contractions to indicate which field is being contracted,

$$\eta^{\mu_i\nu} \varepsilon_{\mu_i} \rightarrow \eta^{\mu_i\nu} \hat{\varepsilon}_{\mu_i} \equiv \eta^{\mu_i\nu} . \quad (5.28)$$

We often annotate the polarization associated with the j th external leg as ε_j when its Lorentz indices are suppressed in expressions, as per:

$$n_3 = [(\varepsilon_1 \cdot \varepsilon_2)(\varepsilon_3 \cdot (k_1 - k_2)) + \text{cyclic}] . \quad (5.29)$$

As such we would write the three-pointing Yang-Mills amplitude as

$$\mathcal{A}_3^{\text{YM}} = g f^{a_1 a_2 a_3} n_3 , \quad (5.30)$$

and the three-point operator associated with this cubic vertex interaction is simply given as:

$$\mathcal{O}_3^{\text{YM}} = g \int \mathcal{D}_3 \frac{1}{6} f^{a_1 a_2 a_3} \hat{n}_3^\mu \mathbf{A}_{\mu,3}^a(x) , \quad (5.31)$$

with the factor of 6 from the automorphic symmetry of the 3 vertex, and

$$\hat{n}_3^\mu(x) = [(\hat{\varepsilon}_1 \cdot \hat{\varepsilon}_2) (\hat{\varepsilon}_3 \cdot (\hat{k}_1 - \hat{k}_2)) + \text{cyclic}] \quad (5.32)$$

$$= [(\eta^{\mu_1 \mu_2}) (\eta^{\mu_3 \rho} (\partial_{\rho x_1} - \partial_{\rho x_2})) + \text{cyclic}] \quad (5.33)$$

$$= -2 [(\eta^{\mu_1 \mu_2}) (\eta^{\mu_3 \rho} \partial_{\rho x_2})] + \text{cyclic} . \quad (5.34)$$

It is not hard to see that this explicitly reproduces the cubic term from $\mathcal{L}^{\text{YM}} = -\frac{1}{4} F_{\mu\nu}^a F^{a\mu\nu}$,

$$\mathcal{L}_3^{\text{YM}} = \frac{1}{6} \int \mathcal{D}_3 g f^{a_1 a_2 a_3} \hat{n}_3^\mu \mathbf{A}_{\mu,3}^a(x) \quad (5.35)$$

$$= f^{a_1 a_2 a_3} \int \prod_i^3 d^d x_i \delta^{(d)}(x - x_i) \hat{n}_3^\mu \mathbf{A}_{\mu,3}^a(x) \quad (5.36)$$

$$= -2 \frac{g}{6} f^{a_1 a_2 a_3} [A^{a_1 \mu} \partial^\nu A_\mu^{a_2} A_\nu^{a_3} + \text{cyclic}] \quad (5.37)$$

$$= -g f^{a_1 a_2 a_3} A^{a_1 \mu} \partial^\nu A_\mu^{a_2} A_\nu^{a_3} . \quad (5.38)$$

Given the fact that graviton polarizations factorize, and that graviton graph numerators at tree-level can be written as the double-copy of color-dual Yang-Mills numerators we can

easily write the three-graviton interaction operator as simply:

$$\mathcal{O}_3^{\text{GR}} = \frac{\kappa}{2} \frac{1}{6} \int \mathcal{D}_3 (\hat{n}_3^\mu \hat{n}_3^\nu)_{\text{ST}} \mathbf{A}_\mu(\mathbf{x}) \otimes_{\text{ST}} \tilde{\mathbf{A}}_\nu(\tilde{\mathbf{x}}). \quad (5.39)$$

Now of course, written in our new form this it's a delight in contrast to manifestly position space operators. This is a feature shared in general with EFT operators, especially as number of derivatives increase. To see how to begin expose a more verbose form we begin to expand out:

$$(\hat{n}_3^\mu \hat{n}_3^\nu)_{\text{ST}} = 4 \left((\hat{\varepsilon}_1 \cdot \hat{\varepsilon}_2) (\hat{\varepsilon}_3 \cdot \hat{k}_2) + \text{cyclic} \right) \left((\hat{\varepsilon}_1 \cdot \hat{\varepsilon}_2) (\hat{\varepsilon}_3 \cdot \hat{k}_2) + \text{cyclic} \right) \Big|_{\text{sym}} \quad (5.40)$$

$$= 4 \left((\eta^{\mu_1 \mu_2}) (\eta^{\mu_3 \rho} \partial_{2\rho}) + \text{cyclic} \right) \left((\eta^{\nu_1 \nu_2}) (\eta^{\nu_3 \sigma} \partial_{2\sigma}) + \text{cyclic} \right) \Big|_{\text{sym}}. \quad (5.41)$$

A brute writing out of a more familiar form of the operator proceeds as the other cases

$$\begin{aligned} \mathcal{O}_3^{\text{GR}} &= \frac{\kappa}{2} \frac{1}{6} \int \mathcal{D}_3 (\hat{n}_3^\mu \hat{n}_3^\nu)_{\text{ST}} (\mathbf{h}_{\mu\nu,3}(\mathbf{x}) \equiv \mathbf{A}_\mu(\mathbf{x}) \otimes \tilde{\mathbf{A}}_\nu(\tilde{\mathbf{x}})). \quad (5.42) \\ &= \frac{\kappa}{2} \frac{4}{6} \int \mathcal{D}_3 \left((\eta^{\mu_1 \mu_2}) (\eta^{\mu_3 \rho} \partial_{2\rho}) + \text{cyclic} \right) \left((\eta^{\nu_1 \nu_2}) (\eta^{\nu_3 \sigma} \partial_{2\sigma}) + \text{cyclic} \right) \mathbf{h}_{\mu\nu,3}(\mathbf{x}) \\ &= \frac{\kappa}{4} \left(\eta^{\mu_1 \mu_2} \eta^{\mu_3 \rho} \eta^{\nu_1 \nu_3} \eta^{\nu_2 \sigma} \partial_\sigma h_{\mu_1 \nu_1} \partial_\rho h_{\mu_2 \nu_2} h_{\mu_3 \nu_3} + \eta^{\mu_1 \mu_2} \eta^{\mu_3 \rho} \eta^{\nu_1 \nu_2} \eta^{\nu_3 \sigma} h_{\mu_1 \nu_1} \partial_{\rho, \sigma} h_{\mu_2 \nu_2} h_{\mu_3 \nu_3} \right. \\ &\quad + \eta^{\mu_1 \mu_2} \eta^{\mu_3 \rho} \eta^{\nu_1 \sigma} \eta^{\nu_2 \nu_3} h_{\mu_1 \nu_1} \partial_\rho h_{\mu_2 \nu_2} \partial_\sigma h_{\mu_3 \nu_3} + \eta^{\mu_1 \mu_3} \eta^{\mu_2 \rho} \eta^{\nu_1 \nu_3} \eta^{\nu_2 \sigma} h_{\mu_2 \nu_2} h_{\mu_3 \nu_3} \partial_{\rho, \sigma} h_{\mu_1 \nu_1} \\ &\quad + \eta^{\mu_1 \mu_3} \eta^{\mu_2 \rho} \partial_\rho h_{\mu_1 \nu_1} (\eta^{\nu_3 \sigma} \eta^{\nu_1 \nu_2} h_{\mu_3 \nu_3} \partial_\sigma h_{\mu_2 \nu_2} + \eta^{\nu_1 \sigma} \eta^{\nu_2 \nu_3} h_{\mu_2 \nu_2} \partial_\sigma h_{\mu_3 \nu_3}) \\ &\quad + \eta^{\mu_2 \mu_3} \eta^{\mu_1 \rho} (\eta^{\nu_1 \nu_3} \eta^{\nu_2 \sigma} h_{\mu_2 \nu_2} \partial_\sigma h_{\mu_1 \nu_1} + \eta^{\nu_3 \sigma} \eta^{\nu_1 \nu_2} h_{\mu_1 \nu_1} \partial_\sigma h_{\mu_2 \nu_2}) \partial_\rho h_{\mu_3 \nu_3} \\ &\quad \left. + \eta^{\rho \mu_1} \eta^{\mu_2 \mu_3} \eta^{\sigma \nu_1} \eta^{\nu_2 \nu_3} h_{\mu_1 \nu_1} h_{\mu_2 \nu_2} \partial_{\rho, \sigma} h_{\mu_3 \nu_3} \right) \\ &= \frac{\kappa}{4} h^{\mu\nu} \left(\partial_{\nu'} h_{\mu' \nu} \partial_\mu h^{\mu' \nu'} + (\partial_{\mu'} h_{\mu \nu'} - \partial_\mu h_{\mu' \nu'}) \partial_{\nu'} h^{\mu' \nu'} \right) \quad (5.43) \end{aligned}$$

Arriving at the final expression simply applied naive integration by parts and relabeling. It is not hard to difficult to see this correctly reproduces the scattering amplitude. Even the final line one can write back in terms of A and $\tilde{A}$, via $h^{\mu\nu} \equiv A \otimes \tilde{A}$, appreciating that μ 's are only contracted with μ 's and ν 's are only contracted with ν 's.

5.3 Generalized gauge accounting in $\hat{\mathcal{O}}(\mathcal{A}_m)$

Imagine a generalized gauge scheme, call it amplitudes gauge, where application of Feynman rules for contacts strictly of multiplicity less than m to graph diagrammatics at m -points yielded precisely the form of $\mathcal{A}$ calculated from the desired form minus the necessary novel contact information. In such a case one can promote the novel contact term from amplitude data directly to the Lagrangian density without care for decisions made at earlier interactions and the gauge used for writing off-shell propagators. Naturally decisions made in all generality can affect the necessary form of the contact required to achieve the target form of the cubic graph dressing at each multiplicity. The functional automorphic form of kinematic numerators

we favor here protects us from many unintentional glitches, but it is a matter of certainty that one must verify that written contacts yield precisely the desired form of the cubic dressings given the conventions used in writing down earlier contacts. Let us make the following clarifying example: $(k_1 + k_2)^2 = 2k_1 \cdot k_2$ when legs one and two are on-shell. One will not have any problems ignoring this distinction when writing the contribution to a four-point contact as $\hat{s} = \hat{k}_1^2 + \hat{k}_2^2 + 2\hat{k}_1 \cdot \hat{k}_2$ as opposed to $\hat{s} = 2\hat{k}_1 \cdot \hat{k}_2$, but the Feynman rules will yield significantly different results in other contexts (at higher multiplicity or in the midst of loop-level diagrams) when k_1 and k_2 are off-shell.

As a matter of course for the m -point amplitude one should verify in any particular action that $\hat{\mathcal{O}}([\mathcal{A}_m]) = \hat{\mathcal{O}}([\mathcal{A}_m])_{\mathcal{L}}$. The latter term represents the explicit Feynman construction of the amplitude using only lower-multiplicity interactions as written in the theory so far as well as the particular choice of propagator. If it does not, reflecting features of previously made generalized gauge choices, then one must of course simply adjust the contact $\mathcal{L}_m$ by the difference accordingly. This recipe embodies the pragmatic recognition that $\mathcal{L}_m = \hat{\mathcal{O}}(\mathcal{A}_m) - \hat{\mathcal{O}}([\mathcal{A}_m])_{\mathcal{L}}$.

A fair question is whether one should bother identifying the contact at the level of the amplitude to begin with, i.e. why not simply promote the full amplitude and subtract the promotion of the partial amplitude generated from lower-multiplicity Feynman rules? One motivation that has found success in repeated aspects of the amplitudes program is to distinguish the universal – in this case the contact unambiguously identified from on-shell cut-data – from the particular, i.e. the form of the contact established via ones gauge choices and propagator structure. This may appear to be a minor or subtle technical point but it is a reality that in much of our lives we often find ourselves confronting in the present the repercussions of decisions we have made in the past.

This is a point that is well emphasized in all approaches that lift gauge amplitudes to gravity contacts, from ref. [9] through e.g. refs. [4, 5]

6 Applications and Examples

In this section, we illustrate the operator promotion framework across a range of theories. These examples are selected to emphasize different features: pedagogical clarity, connection to EFT classification, the power of resummation from string-inspired amplitudes, and gravitational double-copy consistency.

6.1 Pedagogical Case: Yang-Mills and Einstein-Hilbert Gravity

We begin simply with the Yang-Mills theory, where amplitudes are tractable and the color-kinematics structure is manifest on-shell for all low multiplicities requiring novel contact information. This serves as a clear entry point into the operator promotion method, highlighting how contact terms arise and how redundancies are automatically eliminated.

The traditional representation for Yang-Mills actions modulo ghosts and gauge fixing terms is compactly given in terms of the field strength:

$$\mathcal{L}^{\text{YM}} = -\frac{1}{4}\text{Tr}(F^2) \quad (6.1)$$

where we trace over the color indices. For perturbative calculation, this is often expanded out in terms of vector gauge fields, A , resulting in cubic and quartic contact interactions:

$$\mathcal{L}^{\text{YM}} = \mathcal{L}_2^{\text{YM}} + g \mathcal{L}_3^{\text{YM}} + g^2 \mathcal{L}_4^{\text{YM}} \quad (6.2)$$

where g is here taken to be the coupling g^{YM} , $\mathcal{L}_3$ can be written in position-space as proportional to $f^{abc} A^{a\mu} A^{b\nu} \partial_\mu A_\nu^c$, and $\mathcal{L}_4$ as proportional to $c_s(A^a \cdot A^c)(A^b \cdot A^d)$. This is incredibly compact, completely obscuring the tremendous factorial complexity in extracting predictions with traditional approaches. That can be seen perhaps as a boon, certainly both gauge and Lorentz invariance are manifest. One price we do pay is that this traditional form also obscures fundamental color-dual building blocks that, for example, Yang-Mills shares with gravitational theories. It also obscures the well-known fact that these interactions are forced on us at lowest mass-dimension when we require massless vector interactions to be gauge invariant.

Our approach as described in the previous two sections, is to uplift the predictive kinematic weights of m -point graphs $n(g)$ directly to m -field operators $n(g) \rightarrow \hat{n}(g)$ and similarly with the propagator structure $d(g^m) \rightarrow \hat{d}(g^m)$. As already discussed such a promotion involves an almost trivial transcription. External polarizations are replaced by contractions with explicit gauge fields and momentum invariants are appropriately Fourier-transformed and clarified by introducing auxiliary spacetime coordinates constrained via delta functions.

Indeed this simple trick of introducing auxiliary spacetime coordinates will make double-copy structure trivially manifest at the operator level. Carrying out the integration over those spatial delta-functions will obscure the relation to explicit prediction but can allow straightforward identification of these operators with more familiar representations. If we do not explicitly carry out the integration over the auxiliary space-time coordinates, then the reverse operation when generating amplitudes is equally trivial $\hat{n} \rightarrow n$. Having done the hard work of extracting predictions we can recycle that work for future excavation in the description of the theory itself.

Let us quote our earlier example 3 operators for Yang-Mills:

$$\hat{n}_3^\mu(x) = -2 [(\eta^{\mu_1\mu_2})(\eta^{\mu_3\rho}\partial_{\rho x_2})] + \text{cyclic} \quad (6.3)$$

This is the only non-trivial kinematics allowed at three-points between massless vectors at this mass-dimension. It is maximally antisymmetric which necessitates being dressed with an antisymmetric tensor, f^{abc} to make a Bose-invariant amplitude. As shown in eq. (5.38), the

traditional $\mathcal{L}_3^{\text{YM}}$ is completely equivalent to,

$$\mathcal{L}_3^{\text{YM}} = -\frac{g}{6} \int \mathcal{D}_3 f^{a_1 a_2 a_3} \hat{n}_3^\mu \mathbf{A}_{\mu,3}^a(\mathbf{x}). \quad (6.4)$$

$$= -g f^{a_1 a_2 a_3} A^{a_1 \mu} \partial^\nu A_\mu^{a_2} A_\nu^{a_3}. \quad (6.5)$$

Only slightly more involved than the very simple operator for three-point Yang-Mills is the operator associated with the four-point color-dual Yang-Mills numerator at tree-level. Assuming cubic dressings follow the isomorphism of graphs under relabeling there is a unique answer for how contacts are added in a way that satisfies linearized gauge invariance at four-points. These contacts can always be absorbed to the dressings of cubic graphs by multiplying and dividing by propagators. In other words if there is a contact proportional to c_s , say X_{c_s} . Then n_s can absorb the contact by taking some non-contact $n'_s \rightarrow n_s = n'_s + sX$. Gauge invariance can only occur if the color factors satisfy Jacobi relations, and any such representation which assigns contacts to cubic graphs must also satisfy the color-dual Jacobi relations.

The numerator associated with the s channel graph is simply given as the product of cubic numerators added then to a contact term:

$$n_s^{\text{YM}} = n_3(1, 2, -k_{12})n_3(3, 4, k_{12}) + \quad (6.6)$$

$$- s_{12} (\eta_{\alpha_1 \alpha_2} \eta_{\alpha_3 \alpha_4} + \eta_{\alpha_1 \alpha_3} \eta_{\alpha_2 \alpha_4} - \eta_{\alpha_1 \alpha_4} \eta_{\alpha_2 \alpha_3}) \varepsilon^{\alpha_1}(k_1) \varepsilon^{\alpha_2}(k_2) \varepsilon^{\alpha_3}(k_3) \varepsilon^{\alpha_4}(k_4) \quad (6.7)$$

with $k_{ij} = k_i + k_j$. The four-point amplitude is simply given as a sum over all three-factorization channel graphs: $\mathcal{A}_4 = \frac{c_s n_s}{s} + \frac{c_t n_t}{t} + \frac{c_u n_u}{u}$, where the additional channel dressings⁴ are simply related to the s channel by permuting labels.

The above numerator is promoted to an operator following the pervious section's conventions,

$$\hat{n}_s^{\text{YM},\mu} = \hat{n}_3^{\mu_1 \mu_2 \nu}(1, 2, -k_{12}) \cdot \hat{n}_3^{\mu_3 \mu_4 \nu}(3, 4, k_{12}) + \quad (6.8)$$

$$- \hat{s} (\eta^{\mu_1 \mu_2} \eta^{\mu_3 \mu_4} + \eta^{\mu_1 \mu_3} \eta^{\mu_2 \mu_4} - \eta^{\mu_1 \mu_4} \eta^{\mu_2 \mu_3}). \quad (6.9)$$

Locality of operators at higher multiplicity is made explicit by subtracting out the equivalent of cut-contributions as per the organization of the Method of Maximal cuts.

$$\hat{\mathcal{A}}_4 = \frac{1}{8} \int \mathcal{D}_4 \frac{c_s^a \hat{n}_s^{\text{YM},\mu}}{\hat{s}} \mathbf{A}_{\mu,4}(\mathbf{x}) \quad (6.10)$$

Where we subtract off everything but the contact contribution as described earlier. We

⁴One could note that the conventions chosen here follow $c_s = c_t + c_u$, rather than the more symmetric $c_s + c_t + c_u = 0$. This is just a matter of convention of permutation order around vertices used to define what is labeled by t and u graphs. It does not matter at all as long as the numerators follow the same symmetry convention as is required by Bose invariance.

unpack the cut notation we introduced earlier,

$$\hat{\mathcal{A}}_4 = \frac{1}{8} \int \mathcal{D}_4 \frac{c_s^a \hat{\eta}_s^{\text{YM},\mu}}{\hat{s}} A_{\mu,4}^a(x). \quad (6.11)$$

Here, following eq. (4.12) and identifying $n_s^{(1)}$ explicitly as the coefficient of s in n_s^{YM} , we find,

$$\hat{\eta}_s^{\text{YM},\mu} = \hat{n}_s^{\text{YM},\mu} - \hat{s} \hat{n}_s^{(1)\text{YM},\mu} \quad (6.12)$$

$$\begin{aligned} &= [\hat{n}_3^{\mu_1\mu_2\nu}(1, 2, -k_{12}) \cdot \hat{n}_3^{\mu_3\mu_4}{}_\nu(3, 4, k_{12}) - \hat{s} (\eta^{\mu_1\mu_2}\eta^{\mu_3\mu_4} + \eta^{\mu_1\mu_3}\eta^{\mu_2\mu_4} - \eta^{\mu_1\mu_4}\eta^{\mu_2\mu_3})] \\ &\quad + \hat{s} (\eta^{\mu_1\mu_2}\eta^{\mu_3\mu_4} + \eta^{\mu_1\mu_3}\eta^{\mu_2\mu_4} - \eta^{\mu_1\mu_4}\eta^{\mu_2\mu_3}) \\ &= \hat{n}_3^{\mu_1\mu_2\nu}(1, 2, -k_{12}) \cdot \hat{n}_3^{\mu_3\mu_4}{}_\nu(3, 4, k_{12}). \end{aligned} \quad (6.13)$$

Clearly, the only surviving contribution to $\mathcal{L}_4^{\text{YM}}$ after subtracting off cuts is the usual contact term, here written in a way that makes the amplitude's color-dual nature manifest.

$$\mathcal{L}_4^{\text{YM}} = \frac{g^2}{8} \int \mathcal{D}_4 \left(\frac{c_s^a \hat{n}_s^{\text{YM},\mu}}{\hat{s}} A_{\mu,4}^a(x) \right) - \hat{\mathcal{A}}_4 \quad (6.14)$$

$$= -\frac{g^2}{8} \int \mathcal{D}_4 \frac{c_s^a \hat{s}_{12} (\eta_{\alpha_1\alpha_2}\eta_{\alpha_3\alpha_4} + \eta_{\alpha_1\alpha_3}\eta_{\alpha_2\alpha_4} - \eta_{\alpha_1\alpha_4}\eta_{\alpha_2\alpha_3})}{\hat{s}_{12}} A_{\mu,4}^a(x) \quad (6.15)$$

$$= -\frac{g^2}{8} \int \mathcal{D}_4 c_s^a (\eta_{\alpha_1\alpha_2}\eta_{\alpha_3\alpha_4} + \eta_{\alpha_1\alpha_3}\eta_{\alpha_2\alpha_4} - \eta_{\alpha_1\alpha_4}\eta_{\alpha_2\alpha_3}) A_{\mu,4}^a(x) \quad (6.16)$$

$$= -\frac{g^2}{8} f^{a_1a_2b} f^{ba_3a_4} (\eta_{\alpha_1\alpha_2}\eta_{\alpha_3\alpha_4} + \eta_{\alpha_1\alpha_3}\eta_{\alpha_2\alpha_4} - \eta_{\alpha_1\alpha_4}\eta_{\alpha_2\alpha_3}) A_{\mu_1}^{a_1} A_{\mu_2}^{a_2} A_{\mu_3}^{a_3} A_{\mu_4}^{a_4} \quad (6.17)$$

$$= -\frac{g^2}{8} f^{a_1a_2b} f^{ba_3a_4} (A^{a_1} \cdot A^{a_2} A^{a_3} \cdot A^{a_4} + A^{a_1} \cdot A^{a_3} A^{a_2} \cdot A^{a_4} - A^{a_1} \cdot A^{a_4} A^{a_2} \cdot A^{a_3}) \quad (6.18)$$

$$= -\frac{g^2}{8} (f^{a_1a_2b} f^{ba_3a_4} - f^{a_1a_2b} f^{ba_4a_3}) (A^{a_1} \cdot A^{a_3} A^{a_2} \cdot A^{a_4}) \quad (6.19)$$

$$= -\frac{g^2}{4} (f^{a_1a_2b} f^{ba_3a_4}) (A^{a_1} \cdot A^{a_3} A^{a_2} \cdot A^{a_4}). \quad (6.20)$$

Note that in the third to last line the $A^{a_1} \cdot A^{a_2} A^{a_3} \cdot A^{a_4}$ term vanishes due to antisymmetry of color weights.

The same procedure can be used to make every color dual numerator manifest at higher multiplicity, while explicitly only adding 0 to the action.

We will carry out the five field contribution as an example, but perhaps first here it is important to emphasize the connection at four-points with the gravitational contact operator. Any term from the double-copy that survives the removal of cut terms contributes to the contact,

$$\mathcal{L}_4^{\text{GR}} = \left(\frac{\kappa}{2}\right)^2 \frac{1}{8} \int \mathcal{D}_4 \frac{\hat{n}_4^{\text{YM},\mu} \hat{n}_4^{\text{YM},\nu}}{\hat{s}_{12}} h_{\mu\nu,4}(x) - \hat{\mathcal{A}}_4^{\text{GR}} \quad (6.21)$$

$$= \left(\frac{\kappa}{2}\right)^2 \frac{1}{8} \int \mathcal{D}_4 \frac{\hat{n}_4^{\text{YM},\mu} \hat{n}_4^{\text{YM},\nu} - \text{cut}(\hat{n}_4^{\text{YM},\mu} \hat{n}_4^{\text{YM},\nu})}{\hat{s}_{12}} h_{\mu\nu,4}(x) \quad (6.22)$$

$$= \left(\frac{\kappa}{2}\right)^2 \frac{1}{8} \int \mathcal{D}_4 \frac{\hat{n}_4^{\text{YM},\mu} \hat{n}_4^{\text{YM},\nu} - \hat{\eta}_s^{\text{YM},\mu} \hat{\eta}_s^{\text{YM},\nu}}{\hat{s}_{12}} h_{\mu\nu,4}(x) \quad (6.23)$$

$$= \left(\frac{\kappa}{2}\right)^2 \frac{1}{8} \int \mathcal{D}_4 \frac{\hat{s}_{12} \left(\hat{\eta}_s^{\text{YM},\mu} \hat{n}_s^{(1)\text{YM},\nu} + \hat{n}_s^{(1)\text{YM},\mu} \hat{\eta}_s^{\text{YM},\nu} + \hat{s}_{12} \hat{n}_s^{(1)\text{YM},\mu} \hat{n}_s^{(1)\text{YM},\nu} \right)}{\hat{s}_{12}} h_{\mu\nu,4}(x) \quad (6.24)$$

$$= \left(\frac{\kappa}{2}\right)^2 \frac{1}{8} \int \mathcal{D}_4 \left(\hat{\eta}_s^{\text{YM},\mu} \hat{n}_s^{(1)\text{YM},\nu} + \hat{n}_s^{(1)\text{YM},\mu} \hat{\eta}_s^{\text{YM},\nu} + \hat{s}_{12} \hat{n}_s^{(1)\text{YM},\mu} \hat{n}_s^{(1)\text{YM},\nu} \right) h_{\mu\nu,4}(x). \quad (6.25)$$

It is of course straightforward to carry out the delta-point integration to localize to a familiar representation of contact terms, but already here it is sufficient to see the main point which is that in the product of $n_s^{\text{YM}} \hat{n}_s^{\text{YM}}$ that gives the gravitational numerator dressings is the contact term that arises from either side canceling the pole.

Continuing to higher multiplicity contacts is purely mechanical. Color-kinematics and factorization uniquely solves linearized diffeomorphism for Yang-Mills, and associated double-copy satisfies linearized diffeomorphism at each multiplicity. Five points is an instructive example to see how this yields both no additional contact terms for Yang-Mills but non-trivial contact for gravitation.

We can introduce the vanishing $\mathcal{L}_5^{\text{YM}}$, as follows.

$$\mathcal{L}_5^{\text{YM}} = \frac{1}{8} \int \mathcal{D}_5 \frac{c_5^a \hat{n}_5^{\text{YM},\mu}}{\hat{s}_{12} \hat{s}_{45}} A_{\mu,5}^a(x) - \hat{\mathcal{A}}_5^{\text{YM}} \quad (6.26)$$

Because there is no five-point contact for Yang-Mills, we will find that $\mathcal{L}_5^{\text{YM}} = 0$. Let us see how that works out. First we promote the known color-dual five-point Yang-Mills graph numerator to operator form as previous examples. We start by defining various (operator promotions of) sewings of lower multiplicity,

$$\hat{n}_5^{(3,3,3),\mu} \equiv \eta_{\nu\rho} \eta_{\sigma\tau} \hat{n}_3^{\mu_1\mu_2\nu}(1, 2, -k_{12}) \hat{n}_3^{\rho\mu_3\sigma}(k_{12}, 3, k_{45}) \hat{n}_3^{\tau\mu_4\mu_5}(-k_{45}, 4, 5) \quad (6.27)$$

$$\hat{n}_5^{(4,3),\mu} \equiv [\eta_{\nu\rho} \hat{n}_s^{\text{YM},\mu_1\mu_2\mu_3\nu}(1, 2, 3, k_{45}) \hat{n}_3^{\rho\mu_4\mu_5}(-k_{45}, 4, 5) - \hat{n}_5^{(3,3,3),\mu}] + \quad (6.28)$$

$$[\eta_{\nu\rho} \hat{n}_3^{\mu_1\mu_2\nu}(1, 2, -k_{12}) \hat{n}_s^{\rho\mu_3\mu_4\mu_5}(k_{12}, 3, 4, 5) - \hat{n}_5^{(3,3,3),\mu}], \quad (6.29)$$

and can then compactly write the color-dual five-point graph numerator operator as:

$$\hat{n}_5^{\text{YM},\mu} = \hat{n}_5^{(3,3,3),\mu} + \hat{n}_5^{(4,3),\mu} \quad (6.30)$$

$$+ [(s_{45} \eta^{\alpha_3\alpha_4} (\eta^{\alpha_1\alpha_2} k_1^{\alpha_5} - 2\eta^{\alpha_1\alpha_5} k_1^{\alpha_2} - (1 \leftrightarrow 2)) - (4 \leftrightarrow 5)) - \quad (6.31)$$

$$(\{45\} \leftrightarrow \{21\})] . \quad (6.32)$$

The cut contribution, following Equation (4.12), is given by

$$\hat{\mathcal{A}}_5^{\text{YM}} = \frac{1}{8} \int \mathcal{D}_5 \frac{c_5^a \hat{\eta}_5^{\text{YM},\mu}}{\hat{s}_{12}\hat{s}_{45}} \mathbf{A}_{\mu,5}^a(x) \quad (6.33)$$

with

$$\hat{\eta}_5^{\text{YM},\mu} = \hat{n}_5^{\text{YM},\mu} - \hat{s}_{12}\hat{s}_{45}\hat{n}_{\Delta,5}^{(2)}. \quad (6.34)$$

But as there is no term in $\hat{n}_5^{\text{YM},\mu}$ proportional to the product $\Delta = \hat{s}_{12}\hat{s}_{45}$, we have that

$$\hat{n}_{\Delta,5}^{(2)} = 0 \quad (6.35)$$

and so the cut contribution is the entire five-point numerator, and thus $\mathcal{L}_5^{\text{YM}}$ indeed vanishes.

We note that the analogous gravity five field contribution does not vanish:

$$\mathcal{L}_5^{\text{GR}} = \frac{1}{8} \int \mathcal{D}_5 \frac{\hat{n}_5^{\text{YM},\mu} \hat{n}_5^{\text{YM},\nu}}{\hat{s}_{12}\hat{s}_{45}} \mathbf{h}_{\mu\nu,5}(x) - \mathcal{O}(\hat{\mathcal{A}}_5^{\text{GR}})_{\mathcal{L}} \quad (6.36)$$

$$= \frac{1}{8} \int \mathcal{D}_5 \left(\hat{n}_{(12)}^{(1),\mu} \hat{n}_{(45)}^{(1),\nu} + \hat{n}_{(45)}^{(1),\mu} \hat{n}_{(12)}^{(1),\nu} \right) \mathbf{h}_{\mu\nu,5}(x) + \text{gauge dep. terms}, \quad (6.37)$$

as the product $\hat{n}_5^{\text{YM},\mu} \hat{n}_5^{\text{YM},\nu}$ does have terms proportional to $\Delta = \hat{s}_{12}\hat{s}_{45}$ as illustrated in section 4.4. We annotated the cut operator $\mathcal{O}(\hat{\mathcal{A}}_5^{\text{GR}})_{\mathcal{L}}$ to emphasize that the important difference is at the level of lower-multiplicity Feynman rules generated from operators already defined (yielding potential additional gauge-choice specific terms) as discussed in section 5.3. Through four-points we can be rather cavalier about this, but starting at five points lower multiplicity choices of operator feed into the form of the contact operator.

This mechanism of generating gravitational contact terms from products of lower-order effective structures within the constituent Yang-Mills numerators persists to all higher multiplicities. For instance, at six points, while the Yang-Mills action receives no new local 6-field operator, the double copy of 6-point color-dual Yang-Mills numerators (which must now include contributions from distinct cubic topologies beyond the half-ladder, related by Jacobi identities) will generate a non-vanishing $\mathcal{L}_6^{\text{GR}}$. The extraction of this term follows the same principles of promoting the full $n_6^{\text{YM}} \tilde{n}_6^{\text{YM}}$ numerator and systematically subtracting all contributions reconstructible from 3-, 4-, and 5-point gravitational interactions via our cut-based procedure on both half-ladder and trimenuous topologies at the level of Feynman rules.

Our primary goal, contrasting with previous efforts to make the duality between color and kinematics manifest at the level of the action, is the construction of a standard local effective action, achieved by isolating novel contact terms $\mathcal{L}_m = \hat{\mathcal{O}}(\mathcal{C}_m)$. However, it's noteworthy that the intermediate operators $\bar{\mathcal{L}}_m \equiv \hat{\mathcal{O}}(\mathcal{A}_m)$ (before subtracting the cut-constructible parts $\hat{\mathcal{O}}(\mathcal{A}_m)$) could themselves be viewed as an alternative specification of the predictions of theory. In such an approach, the m -point scattering amplitude would be generated solely by an m -point contact diagram dressed with the Feynman rule derived from $\bar{\mathcal{L}}_m$. This rule would,

by construction, reproduce the full $\mathcal{A}_m$, including all its poles and residues, and if $\mathcal{A}_m$ was provided in a color-dual form, this vertex would directly yield that structure. While these $\bar{\mathcal{L}}_m$ operators are generally non-local (as they encode propagator structures), this perspective offers a direct map from full m -point amplitudes to m -point effective vertices that generate them in their entirety, conceptually aligning with efforts to find Lagrangians or Feynman rules that directly manifest color-kinematics duality for all graph contributions. There is nothing surprising about this – as we have endeavored to emphasize we are simply writing operators to correctly produce amplitudes from known amplitude data.

6.2 Higher Derivative Operators from Pointlike Fields to String Theory

We now turn to one of the motivating points for our formalism – the canonical identification of higher derivative operators. As demonstrated in refs. [1, 2], the duality between color and kinematics allows for a particularly efficient way to construct and classify distinct amplitudes associated with higher derivative operators. As such we have now a sharp mechanism for promoting these amplitudes to construct and classify the higher derivative operators themselves while maintaining the color-dual structure.

Recall that via field redefinition, color-zeroes, kinematic-zeroes, and total derivatives there are an uncountably infinite number of ways of rewriting any given operator. Consider the symmetrized four-field gauge operator responsible for generating the abelian four-point Born-Infeld amplitude. The four-field operator is given simply as

$$\mathcal{O}_{\text{BI}} = \frac{g^2}{32} \left((F^2)^2 + (F\tilde{F})^2 \right). \quad (6.38)$$

As many readers may appreciate this same operator (suitably supersymmetrized) results in a superamplitude that is a universal state-encoding prefactor for four-point scattering for the maximally supersymmetric Yang-Mills at any loop order, and squared is the universal prefactor for four-point scattering in the maximally supersymmetric supergravity amplitude. The four point amplitude it produces is quite simply:

$$\mathcal{A}_{\text{BI}} = stA_4^{\text{YM}}(1234) \quad (6.39)$$

$$= n_s^{\text{YM}}t + n_t^{\text{YM}}s \quad (6.40)$$

$$= n_s^{\text{YM}}(t-u)/3 + n_t^{\text{YM}}(s-u)/3 + n_u^{\text{YM}}(t-s)/3 \quad (6.41)$$

$$= \frac{n_s^{\text{YM}}s(t-u)/3}{s} + \text{perms} \quad (6.42)$$

$$= \frac{n_s^{\text{YM}}n_s^\pi}{s} + \text{perms} \quad (6.43)$$

In the third line we used that all four-point cubic graph representations of tree-level Yang-Mills satisfy Jacobi $n_u = n_s - n_t$ as well as conservation of momenta. In the fifth line we identified the NLSM four-point kinematic numerator $n_s^\pi \propto s(t-u)$. Here we see the four-point demonstration that we can understand Born-Infeld as the double-copy between Yang-

Mills and NLSM pions – a pattern that explicitly holds to all multiplicity. This extends to supersymmetric Dirac-Born-Infeld-Volkov-Akulov (DBI-VA) where the scalar and fermionic states are carried from supersymmetry on the Yang-Mills single-copy.

We can clearly therefore write the Born-Infeld four-point operator.

$$\mathcal{O}_{\text{BI}} = \int \mathcal{D}_4 \frac{1}{8} \frac{\hat{n}_s^\pi \hat{n}_s^{\text{YM}, \mu}}{\hat{s}} \mathbf{A}_{\mu, 4}^{\text{BI}}(x) \quad (6.44)$$

The seeming non-locality is immediately canceled with the inverse-propagator present in the pion numerator.

As fantastic as this is for exposing the double-copy structure of BI operator by operator, our approach comes into its own in that the entire tower of higher derivative gluonic operators contributing to the supersymmetric open string, whose predictions are described in ref. [1], are simply given by promoting the higher-derivative color-weights of Equation (2.12) to operators. Let's remind ourselves what the structure looks like. To build these types of higher derivative operators we are considering all higher derivative modifications to the color-weight of while respecting adjoint-type algebraic relations. In general this means considering all color-dressings $o(g)$ that are adjoint compositions between graph dressings:

$$o(g) = (c \circledast n)(g) \quad (6.45)$$

Where we differentiate between pure-scalar kinematic $n(g)$ and color weights, $c(g)$ that may contain higher-derivative contributions. Clearly the mass-dimension of the LHS is the sum of the mass-dimensions of both $c(g)$ and $n(g)$. As a matter of book-keeping we mod out any permutation invariants such that the contribution to a full amplitude of a particular mass dimension will always be given by:

$$o(g) = \sum_{ij} \alpha^i w_{j\vec{k}[j]} o_j(g) \left(\prod_l p_l(g)^{k[j]_l} \right) \quad (6.46)$$

where α^i tracks mass-dimension, w_X are Wilson coefficients, and the sum of mass-dimensions of each permutation invariant p_l weighted by $k[j]_l$ is equal to i minus the mass-dimension of o_j^{HD} , i.e.,

$$i = \sum_l k_l \times [p_l] + [o_j^{\text{HD}}]. \quad (6.47)$$

At four-points this is particularly simple as the number of permutation invariants for massless kinematics is incredibly small:

$$\sigma_2 = s^2 + t^2 + u^2 \quad (6.48)$$

$$\sigma_3 = stu \quad (6.49)$$

Where $s + t + u = 0$, $s = (k_1 + k_2)^2$, and $t = (k_2 + k_3)^2$. It turns out that every permutation

invariant of (s, t, u) of higher mass-dimension m can be represented by:

$$p_m = \sum_{ij} w_{ij} \sigma_2^i \sigma_3^j \quad (6.50)$$

with constant w_{ij} and $2i + 3j = m$. If we restrict ourselves to four-point $o_j^{\text{HD}}(g)$ that satisfy antisymmetry and Jacobi we have a very small number of building blocks then as per ref. [1],:

$$o_1(g_s) = c_s \quad (6.51)$$

$$o_2(g_s) = (c \textcircled{a} n^{\text{ss}})_s = c_t(s - u) - c_u(t - s) \quad (6.52)$$

$$o_3(g_s) = d^{abcd} n_s^\pi. \quad (6.53)$$

The first requires little discussion – higher-derivative contributions play a part only via the trivial product of permutation invariants that have been modded out. The second arises from composing color weights with the covariantized free-scalar kinematic numerator. The third’s Jacobi satisfying properties arise from the color-dual pion kinematic weight $n_s^\pi \propto s(t - u)$, and the color-weights contribute through the four-point color permutation invariant. Note the mass dimension of $[o_j] = 2(j - 1)$ for these three graph dressings.

So for the four-point case we have simply that all higher-derivative adjoint four-field color-weights are spanned by:

$$o(g) = \sum_{i=1}^{\infty} \alpha^i \sum_{j=1}^3 \sum_{k_2, k_3} w_{jk_2 k_3} \sigma_2^{k_2} \sigma_3^{k_3} o_j(g) \quad (6.54)$$

where $((j - 1) + 2k_2 + 3k_3) = i$. For $o(g)$ to be the result of a local (non-factorizing) operator we require that $k_3 \geq 1$ for o_1 and o_2 . The promotion to an operator is straightforward using:

$$\hat{o}_1^a(g_s) = c_s \quad (6.55)$$

$$\hat{o}_2^a(g_s) = c_t(\hat{s} - \hat{u}) - c_u(\hat{t} - \hat{s}) \quad (6.56)$$

$$\hat{o}_3^a(g_s) = d^{a_1 a_2 a_3 a_4} \hat{n}_s^\pi. \quad (6.57)$$

$$\hat{\sigma}_2 = \hat{s}^2 + \hat{t}^2 + \hat{u}^2 \quad (6.58)$$

$$\hat{\sigma}_3 = \hat{s} \hat{t} \hat{u}. \quad (6.59)$$

Indeed we can write any four-field operator contributing to the gluonic sector of the α' expansion of the open superstring as:

$$\mathcal{O}^{\text{OSS}} = \int \mathcal{D}_4 \frac{1}{8} \frac{\hat{o}_s^a \hat{n}_s^{\text{YM}, \mu}}{\hat{s}} A_{\mu, 4}^a(x). \quad (6.60)$$

Recall that the all multiplicity open superstring was notably expressed in terms of m -point super-Yang-Mills amplitudes [39, 40]. Remarkably the authors of ref. [41], isolated the doubly-ordered disk integrals which carry all orders in α' corrections and double-copy with ordered

Table 2: Values of the $w_{j,k_2,k_3}^{[MD]/2}$ coefficients appearing in $\hat{o}$ that match $\mathcal{O}^{\text{OSS}}$ to the low energy expansion of the open superstring amplitude through mass dimension 26.

$ \begin{aligned} w_{1,0,0}^{[0]} &= 1 \\ w_{1,0,1}^{[3]} &= -\zeta_3 \\ w_{3,1,0}^{[4]} &= \frac{5\zeta_4}{4} \\ w_{1,1,1}^{[5]} &= -\frac{\zeta_5}{2} \\ w_{3,2,0}^{[6]} &= \frac{21\zeta_6}{32} \\ w_{2,0,2}^{[7]} &= -\frac{1}{4}\zeta_3\zeta_4 \\ w_{1,2,1}^{[7]} &= -\frac{\zeta_7}{4} \\ w_{3,3,0}^{[8]} &= \frac{85\zeta_8}{256} \\ w_{1,1,2}^{[8]} &= \frac{\zeta_3\zeta_5}{2} + \frac{49\zeta_8}{128} \\ w_{3,2,1}^{[9]} &= \frac{1}{32}(-20\zeta_4\zeta_5 - 21\zeta_3\zeta_6 - 16\zeta_2\zeta_7) \\ w_{1,3,1}^{[9]} &= -\frac{\zeta_9}{8} \\ w_{3,4,0}^{[10]} &= \frac{341\zeta_{10}}{2048} \\ w_{3,1,2}^{[10]} &= \frac{5}{8}\zeta_3^2\zeta_4 + \zeta_2\zeta_3\zeta_5 + \frac{737\zeta_{10}}{320} \\ w_{2,2,2}^{[11]} &= \frac{1}{256}(-20\zeta_5\zeta_6 - 16\zeta_4\zeta_7 - 21\zeta_3\zeta_8) \\ w_{3,3,1}^{[11]} &= \frac{1}{256}(-84\zeta_5\zeta_6 - 80\zeta_4\zeta_7 - 85\zeta_3\zeta_8 - 64\zeta_2\zeta_9) \\ w_{1,4,1}^{[11]} &= -\frac{\zeta_{11}}{16} \\ w_{3,5,0}^{[12]} &= \frac{1365\zeta_{12}}{16384} \\ w_{3,2,2}^{[12]} &= \frac{1}{4}\zeta_2\zeta_5^2 + \frac{21}{64}\zeta_3^2\zeta_6 + \frac{1}{8}\zeta_3(5\zeta_4\zeta_5 + 4\zeta_2\zeta_7) + \frac{3162705\zeta_{12}}{1415168} \\ w_{1,3,2}^{[12]} &= \frac{1}{8}(\zeta_5\zeta_7 + \zeta_3\zeta_9) + \frac{1793\zeta_{12}}{8192} \\ w_{2,0,4}^{[13]} &= -\frac{1}{24}\zeta_4(\zeta_3^3 + 2\zeta_9) - \frac{35\zeta_3\zeta_{10}}{256} \\ w_{3,4,1}^{[13]} &= \frac{-336\zeta_6\zeta_7 - 340\zeta_5\zeta_8 - 320\zeta_4\zeta_9 - 341\zeta_3\zeta_{10} - 256\zeta_2\zeta_{11}}{2048} \\ w_{1,5,1}^{[13]} &= -\frac{\zeta_{13}}{32} \end{aligned} $	$ \begin{aligned} w_{3,0,0}^{[2]} &= 2\zeta_2 \\ w_{2,0,1}^{[4]} &= \frac{\zeta_4}{4} \\ w_{3,0,1}^{[5]} &= -2\zeta_2\zeta_3 \\ w_{2,1,1}^{[6]} &= \frac{5\zeta_6}{32} \\ w_{1,0,2}^{[6]} &= \frac{1}{16}(8\zeta_3^2 + 5\zeta_6) \\ w_{3,1,1}^{[7]} &= -\frac{5}{4}\zeta_3\zeta_4 - \zeta_2\zeta_5 \\ w_{2,2,1}^{[8]} &= \frac{21\zeta_8}{256} \\ w_{3,0,2}^{[8]} &= \zeta_2\zeta_3^2 + \frac{155\zeta_8}{96} \\ w_{2,1,2}^{[9]} &= \frac{1}{32}(-4\zeta_4\zeta_5 - 5\zeta_3\zeta_6) \\ w_{1,0,3}^{[9]} &= \frac{1}{48}(-8\zeta_3^3 - 15\zeta_3\zeta_6 - 16\zeta_9) \\ w_{2,3,1}^{[10]} &= \frac{85\zeta_{10}}{2048} \\ w_{2,0,3}^{[10]} &= \frac{1}{8}\zeta_3^2\zeta_4 + \frac{35\zeta_{10}}{256} \\ w_{1,2,2}^{[10]} &= \frac{1}{8}(\zeta_5^2 + 2\zeta_3\zeta_7) + \frac{321\zeta_{10}}{1024} \\ w_{3,0,3}^{[11]} &= -\frac{155}{96}\zeta_3\zeta_8 - \frac{1}{3}\zeta_2(\zeta_3^3 + 2\zeta_9) \\ w_{1,1,3}^{[11]} &= \frac{1}{128}(-4\zeta_5(8\zeta_3^2 + 5\zeta_6) - 49\zeta_3\zeta_8 - 64\zeta_{11}) \\ w_{2,4,1}^{[12]} &= \frac{341\zeta_{12}}{16384} \\ w_{2,1,3}^{[12]} &= \frac{1}{64}\zeta_3(8\zeta_4\zeta_5 + 5\zeta_3\zeta_6) + \frac{152911\zeta_{12}}{707584} \\ w_{1,0,4}^{[12]} &= \frac{1}{96}\zeta_3(4\zeta_3^3 + 15\zeta_3\zeta_6 + 32\zeta_9) + \frac{199881\zeta_{12}}{707584} \\ w_{2,3,2}^{[13]} &= \frac{-80\zeta_6\zeta_7 - 84\zeta_5\zeta_8 - 64\zeta_4\zeta_9 - 85\zeta_3\zeta_{10}}{2048} \\ w_{3,1,3}^{[13]} &= \frac{-5(40\zeta_3^3\zeta_4 + 96\zeta_2\zeta_3^2\zeta_5 + 155\zeta_5\zeta_8 + 80\zeta_4\zeta_9) - 2211\zeta_3\zeta_{10} - \zeta_2\zeta_{11}}{960} \\ w_{1,2,3}^{[13]} &= \frac{-4(32\zeta_3\zeta_5^2 + 32\zeta_3^2\zeta_7 + 20\zeta_6\zeta_7 + 49\zeta_5\zeta_8) - 321\zeta_3\zeta_{10} - 512\zeta_{13}}{1024} \end{aligned} $
--	---

Yang-Mills amplitudes to form color-ordered string theory amplitudes. These integrals were later interpreted as doubly-ordered amplitudes in a very special bi-colored all-order effective field theory called Z -theory [42–44]. One ordering crucially obeys field theory KK and BCJ relations, and the other ordering obeys string monodromy relations. When the monodromy ordering is dressed with Chan-Paton factors then we find color-dual field theory amplitudes. The duality between color and kinematics appears to all orders in α' is precisely due to the all-order mixing of color and scalar kinematic weights as was clarified in [1, 2].

Indeed, from the structure of eq. (6.60) one immediately recognizes that that the four-point Z -theory α' expanded amplitude is encoded in

$$\mathcal{O}_4^Z = \int \mathcal{D}_4 \frac{1}{8} \frac{\hat{o}_s^a c_s^{\tilde{a}}}{\hat{s}} \phi_4^{a\tilde{a}}(x). \quad (6.61)$$

By construction this doubly-colored theory obeys field theory relations on its $\tilde{a}$ color-ordered amplitudes. Actual Z theory amplitudes are reproduced with appropriate choice of Wilson coefficients which also imposes string theory monodromy relations on its a color-ordered amplitudes. We include in Tab. 2 the necessary Wilson coefficients through mass dimension 26 that reproduce the low energy expansions of Z and open super-string amplitudes.

Given the ease of the series expansion one can wonder if it is possible to write the resummed open string and Z theory operators. At four-points this is straightforward to do

in closed form as the disc integrals are known in terms of Euler Gamma functions. One can simply write the s-channel Z -theory numerator as,

$$n_s^Z = \frac{1}{3\sigma_3} s(t-u) [stA^Z(s,t)] \quad (6.62)$$

where the Chan-Paton dressed permutation invariant $stA^Z(s,t)$ is given in ref. [1], and we quote here,

$$[stA^Z(s,t)] = \frac{\pi^2}{\alpha'} \frac{\csc(\pi\alpha's) \csc(\pi\alpha't) \csc(\pi\alpha'u)}{\Gamma(-\alpha's)\Gamma(-\alpha't)\Gamma(-\alpha'u)} \times \\ (c_s z_s + c_t z_t + c_u z_u + d^{a_1 a_2 a_3 a_4} 2 [\sin(\pi\alpha's) + \sin(\pi\alpha't) + \sin(\pi\alpha'u)]) , \quad (6.63)$$

where $z_s = (\sin(\pi\alpha'u) - \sin(\pi\alpha't))/3$, $c_s = f^{a_1 a_2 e} f^{e a_3 a_4}$, and the other channel c_g follow from simple relabeling. Note the z_g satisfy $z_s = z_t + z_u$ in concordance with $c_s = c_t + c_u$.

The promotion of kinematic invariants s_{ij} to operators $\hat{s}_{ij}$ is applied directly within these expressions. This defines operators that compactly encode the all-orders α' behavior. It is worth noting that the resulting position-space operators, such as $\Gamma(-\alpha'\hat{s})$, are necessarily non-polynomial in derivatives and reflect the rich analytic structure of the underlying string amplitudes. While their formal properties as differential operators can be intricate, their definition and utility within our framework are anchored by their ability to generate the correct, known string S-matrix elements. This S-matrix-centric perspective allows us to construct action-level counterparts for theories whose amplitudes may not conform to the strictest notions of point-particle locality or perturbative unitarity (as is the case for full string theory or certain exotic theories like conformal supergravity, respectively). The primary goal is a faithful operator encoding of the on-shell physics.

If one is happy to promote unevaluated disk integrals one can use the technology of virtuous trees to write Jacobi satisfying numerators at any multiplicity. The m -point Chan-Paton dressed, but field theory ordered Z -theory amplitude is given by summing the doubly-ordered Z integrals over the integration domain weighted by Chan-Paton traces.

$$Z(q_1, \dots, q_m) = \alpha'^{m-3} \sum_{\rho \in S_{m-1}(2, \dots, m)} \text{Tr}[1\rho] \int_{D(1\rho)} \frac{dz_1 dz_2 \dots dz_m}{\text{vol}(SL(2, \mathbb{R}))} \frac{\prod_{i < j}^m |z_{ij}|^{\alpha' s_{ij}}}{z_{q_1 q_2} z_{q_2 q_3} \dots z_{q_m q_1}} \quad (6.64)$$

Where the trace $\text{Tr}[1\rho]$ is shorthand for the appropriate color trace on Chan-Paton indices $\text{Tr}[T^{a_1} T^{a_{\rho_2}} \dots T^{a_{\rho_m}}]$. Notice all kinematics appear in terms of Mandelstam invariants which are trivially promoted to differential operators as we will demonstrate. One simply needs to map from ordered amplitudes to functional kinematic numerators, which can be accomplished via the virtuous tree representation introduced by Broedel and Carrasco in [37] and can be constructively built at any multiplicity by symmetrizing over the KLT kernel as per Fu, Du, and Feng in ref. [45] and Naculich in ref. [46].

As an example let us consider the five-point Chan-paton dressed, field theory ordered

Z -theory amplitude which satisfies the KK and BCJ relations on the field-theory ordering,

$$Z(12345) = \alpha'^2 \sum_{\rho \in S_4(2, \dots, 5)} \text{Tr}[1\rho] \int_{D(1\rho)} \frac{dz_1 dz_2 \cdots dz_5}{\text{vol}(SL(2, \mathbb{R}))} \frac{\prod_{i < j}^5 |z_{ij}|^{\alpha' s_{ij}}}{z_{q_1 q_2} z_{q_2 q_3} \cdots z_{q_5 q_1}} \quad (6.65)$$

Conceptually, one could imagine promoting $s_{ij} \rightarrow \hat{s}_{ij}$ directly within the Koba-Nielsen integrand, e.g., in terms like $|z_{ab}|^{\alpha' s_{ab}} \rightarrow |z_{ab}|^{\alpha' \hat{s}_{ab}}$. Such an object would formally define an operator whose coefficients are given by integrals of operator-valued functions over the string worldsheet coordinates z_i . While a full exploration of such operators is beyond our present scope, their formal series expansion in α' would correspond to an infinite tower of local higher-derivative operators whose coefficients are the standard string integrals. If this bothers discerning analytic readers of taste, feel free to read the prescription as promoting $s_{ij} \rightarrow \hat{s}_{ij}$ in the expressions for numerators that are already expressed as functions of Mandelstam invariants (which may themselves be the result of evaluating string integrals or their expansions).

The low energy expansion of Equation (6.65) has been verified [2] to be spanned by a constructive compositional ansatz through α'^9 (mass dimension 18). Let us see how we can build a color-dual numerator in terms of the unexpanded disk integrals. A manifestly Jacobi satisfying, functional, color-dual numerator of the five-point half-ladder graph [37] is given by:

$$\begin{aligned} n_5^Z = \frac{1}{30} & \left(\left[s_{12} s_{45} (Z_{12345} - Z_{12354} - Z_{21345} + Z_{21354}) \right] \right. \\ & + \left[s_{12} (s_{34} - s_{35}) (Z_{14352} + Z_{15342}) \right. \\ & \quad \left. + s_{45} (s_{13} - s_{23}) (Z_{51324} - Z_{41325}) \right] \\ & + \left[(s_{12} s_{34} - s_{12} s_{35}) Z_{14352} + (s_{12} s_{34} - s_{12} s_{35}) Z_{15342} \right. \\ & \quad + (-s_{15} s_{23} - s_{25} s_{34}) Z_{14325} + (s_{14} s_{23} + s_{24} s_{35}) Z_{15324} \\ & \quad \left. + (s_{13} s_{24} + s_{14} s_{25}) Z_{41352} + (-s_{13} s_{25} - s_{15} s_{24}) Z_{51342} \right] \Big). \quad (6.66) \end{aligned}$$

Here we use the shorthand $Z_X = Z(X)$. The s_{ij} are trivially promoted to differential operators $\hat{s}_{ij}$, including in the numerator of the disk integrals $|z_{ij}|^{s_{ij}}$, yielding a five-point operator which reproduces the entire gluonic sector of the tree-level open superstring by:

$$\mathcal{O}_5^{\text{OSS}} = \int \mathcal{D}_5 \frac{\hat{n}_5^{Z, a} n_s^{\text{YM}, \mu}}{\hat{s}} \mathbf{A}_{\mu, 5}^a(x) \dots \quad (6.67)$$

Promoting n^{YM} to its supernumerator using on-shell superspace naturally reproduces the entire multiplet.

The virtuous (tree-encoded, color-dual, functionally symmetric) numerator for the half-ladder graph labeled $\sigma_1 \dots \sigma_m$ is given in closed form in terms of virtuous trace-kinematic graph weights τ as follows.

$$n_\sigma = \tau_{\sigma_1[\sigma_2, [\sigma_3, [\dots, m] \dots]]} . \quad (6.68)$$

Here the brackets in the kinematic trace τ signify an antisymmetric combination, following BernDennen. i.e.

$$\tau_{1[2,[3,4]]} = \tau_{1234} - \tau_{1243} - (\tau_{1342} - \tau_{1432}) . \quad (6.69)$$

The approach to generating such τ is perhaps obvious from KLT construction, they come easily from ordered amplitudes. Given that full amplitudes are bose-symmetric, one can average the KLT expression over all permutations of leg labels and simply read off the cycle τ ,

$$\tau_\sigma = -\frac{\partial}{\partial \tilde{A}(\sigma)} \frac{1}{m!} \sum_{S_m} \frac{1}{2} \left(\sum_{\rho, \tau \in S_{m-3}} \tilde{A}(1, \rho, m, m-1) S_{\rho|\tau} A(1, \tau, m-1, m) + \tilde{A} \leftrightarrow A \right) \quad (6.70)$$

where the outer sum is over all permutations of leg labels. The derivative is taken with the understanding that all orderings of trees are cyclically identified. Amusingly the definition of the antisymmetric Jacobi-satisfying n_σ in terms of shuffle operations on τ can be seen as equivalent of re-expressing the permutation summed KLT expression in a Kleiss-Kujif basis for the $\tilde{A}$,

$$n_\sigma = -\frac{\partial}{\partial \tilde{A}(\sigma)} \left(\frac{1}{m!} \sum_{S_m} \left[\frac{1}{2} \sum_{\rho, \tau \in S_{m-3}} \tilde{A}(1, \rho, m, m-1) S_{\rho|\tau} A(1, \tau, m-1, m) + \tilde{A} \leftrightarrow A \right] \right) \Big|_{\tilde{A}(\sigma_1 \beta \sigma_n)} \quad (6.71)$$

We see here an all-order constructive form of color-dual numerators. If the ordered amplitudes are Yang-Mills then the double-copy is to gravitation, and the novel contact information of the n -point graviton is encoded as the expression in the double-copy that survives subtraction of all cuts.

7 Path to Quantum Gravity via Double Copy

Having established the operator promotion framework, we now explore its implications for understanding quantum gravity as a double copy of Yang-Mills theory, both at the level of the effective action and the quantum states.

7.1 From Yang-Mills amplitudes to the Einstein-Hilbert action

Our framework offers a constructive route⁵ to the full expansion of the Einstein-Hilbert action, $\sqrt{-g}R$, written in terms of the fluctuation $h_{\mu\nu} = g_{\mu\nu} - \eta_{\mu\nu}$. While it is well known that

⁵We should emphasize, as noted in our introduction, that we are not the first to do make this constructive argument. Notably using only the existence of KLT relations – Bern and Grant first demonstrated this

gravity amplitudes arise as the double copy of gauge theory, we go further: we construct each individual interaction term in $\sqrt{-g}R$ from local operator contributions derived directly from gauge theory amplitudes.

This relies on two established facts:

1. Gravitational amplitudes are a double copy of Yang-Mills. At tree level and all multiplicities, the gravitational m -point amplitude takes the form

$$\mathcal{A}_m^{\text{GR}} = \sum_{g \in \Gamma_3^m} \frac{n_m^{\text{YM}}(g) \tilde{n}_m^{\text{YM}}(g)}{d_g}, \quad (7.1)$$

where $n_m^{\text{YM}}(g)$ and $\tilde{n}_m^{\text{YM}}(g)$ are color-dual numerators for each cubic graph g , and d_g is the product of graph propagators.

2. Color-dual numerators exist at all multiplicity. A variety of constructions provide explicit expressions for $n_m^{\text{YM}}(g)$ that satisfy the kinematic Jacobi relations and generate the full Yang-Mills S-matrix.

Given these facts, our formalism promotes each gravitational amplitude to a local operator in the action. We extract the novel m -point contact term by subtracting contributions reconstructible from lower-point factorizations,

$$\mathcal{C}_m^{\text{GR}} = \mathcal{A}_m^{\text{GR}} - \mathcal{A}_m^{\text{GR}}, \quad (7.2)$$

and define the corresponding operator as

$$\mathcal{L}_m^{\text{GR}} = \hat{\mathcal{O}}(\mathcal{C}_m^{\text{GR}}). \quad (7.3)$$

Each $\mathcal{L}_m$ is manifestly local and required for the interacting action to be invariant under linearized diffeomorphism invariance and arises entirely from graph-level double-copy structure.

Crucially, Yang-Mills theory contains no fundamental contact terms beyond four points. In our framework, this means $\mathcal{L}_m^{\text{YM}} = 0$ for $m > 4$, and all higher contact structure in gravity must arise from cross terms in the double copy. That is, products of pole-canceling terms from each gauge-theory copy can combine to produce contact terms in gravity. We make this mechanism explicit at five points in Section 6.1 and more generally in Section 4.4.

For example, the Yang-Mills half-ladder at five points contains separate terms one proportional to s_{12} and another proportional to s_{45} . Upon double copy between two such numerator factors these terms appear together yielding a contact term proportional to $s_{12}s_{45}$, precisely the form required for the five-point expansion of $\sqrt{-g}R$. This structure persists at six points and beyond. The gravitational contact terms $\mathcal{L}_m^{\text{GR}}$ are nonvanishing for all $m \geq 3$, and for $m > 4$ are not physically present in either gauge-theory factor alone.

was towards generating spectacularly compact representations of higher-multiplicity gravitational operators in ref. [9].

Summing over all multiplicities gives the complete expansion of the gravitational action,

$$\mathcal{S}_{\text{tree}}^{\text{GR}} = \int d^D x \sum_{m \geq 3} \mathcal{L}_m^{\text{GR}}. \quad (7.4)$$

Each term $\mathcal{L}_m$ generates the correct m -point amplitude, and thus the sum matches the standard expansion of the Einstein-Hilbert action up to field redefinitions and total derivatives. This provides a direct, structured map from gauge theory to gravity that makes manifest the double-copy origin of each interaction.

7.2 State level description

We have shown how local operators in double-copy theories can be expressed in terms of factorized operators acting on constituent gauge-theory-like copies. Let us now outline the path toward a state-level encoding of this structure, focusing on quantum gravity emerging from Yang-Mills theory.

To begin, recall the conceptual path from Maxwell theory to Yang-Mills. Free photon states encode linearized gauge invariance via $D - 2$ physical polarizations. On-shell three-vector interactions, constrained by mass dimension and gauge invariance, uniquely fix the kinematic numerator to be the maximally antisymmetric structure:

$$n_3(k_1, k_2, k_3) = (\varepsilon_1 \cdot \varepsilon_2)(\varepsilon_3 \cdot (k_1 - k_2)) + \text{cyclic}. \quad (7.5)$$

Consistency with Bose symmetry then demands an accompanying maximally antisymmetric color factor, the first manifestation of the duality between color and kinematics. Imposing gauge invariance at four points necessitates a local contact term such that both color and kinematic numerators (assigned to cubic graphs) obey Jacobi identities. Promoting this amplitude structure to field-level operators yields Yang-Mills theory.

In constructing gravity, we proceed in parallel. Free graviton states must be consistent with spin-2 gauge symmetry, i.e., invariance under linearized diffeomorphisms: $h_{\mu\nu} \rightarrow h_{\mu\nu} + \partial_\mu \xi_\nu + \partial_\nu \xi_\mu$. In our amplitude-centric formulation, $h_{\mu\nu}$ (defined as $g_{\mu\nu} - \eta_{\mu\nu}$ without restriction on magnitude) primarily encodes asymptotic on-shell data. Each free graviton state corresponds to a symmetric-traceless (ST) polarization tensor, which precisely captures the $D(D - 3)/2$ physical polarizations. This ST structure is realized as a double copy of two gauge-theory polarization vectors:

$$\varepsilon_{\text{GR}}^{\mu\nu} = \varepsilon^{(\mu} \tilde{\varepsilon}^{\nu)} - \frac{1}{D} \eta^{\mu\nu} (\varepsilon \cdot \tilde{\varepsilon}). \quad (7.6)$$

This definition inherently enforces the requirements of linearized diffeomorphism invariance at the level of asymptotic states.

Interactions must preserve this symmetry at every multiplicity. This is guaranteed when gravitational amplitudes are expressed as a double copy of color-dual Yang-Mills amplitudes. The operator promotion framework developed in this paper provides an all-multiplicity algo-

rithm for expressing every gravitational operator in a factorized form, $\hat{\mathcal{O}}_{GR} \sim \hat{\mathcal{O}}_{YM} \otimes \hat{\mathcal{O}}_{\widetilde{YM}}$, suitable for acting on a correspondingly structured Hilbert space.

Let us make this state-level structure concrete. A free Yang-Mills state of momentum k , helicity λ , and adjoint color a is:

$$|\text{YM}(k)_\lambda^a\rangle = |k\rangle \otimes |\varepsilon_\lambda\rangle \otimes |a\rangle_G. \quad (7.7)$$

A free graviton state, built as a double copy, then takes the form:

$$|\text{GR}(k)_{\lambda,\tilde{\lambda}}\rangle = |k\rangle \otimes |\varepsilon_{\text{GR};\lambda,\tilde{\lambda}}\rangle, \quad (7.8)$$

where $\varepsilon_{\text{GR};\lambda,\tilde{\lambda}}^{\mu\nu}(k)$ can be expressed as the symmetric traceless (ST) tensor from Eq. (7.6) formed from $\varepsilon_\lambda(k)$ and $\tilde{\varepsilon}_{\tilde{\lambda}}(k)$. This defines a state-level double copy:

$$|\text{GR}(k)_{\lambda,\tilde{\lambda}}\rangle = \left(|\text{YM}(k)_\lambda^a\rangle \otimes_{\text{ST}} |\widetilde{\text{YM}(k)}_{\tilde{\lambda}}^{\tilde{b}}\rangle \right)_{\text{color-singlet}}. \quad (7.9)$$

The slashed color indices $\not{a}, \not{\tilde{b}}$ signify that color degrees of freedom are traced over, yielding a color-singlet state essential for diffeomorphism invariance of observables. The $\otimes_{\text{ST}}$ denotes that the tensor product of the kinematic parts (momentum and polarization kets) is projected onto the symmetric-traceless representation for the graviton. More explicitly, one can construct this state from constituent polarization kets as:

$$|\text{GR}(k)_{\lambda,\tilde{\lambda}}\rangle = \frac{1}{\sqrt{N_c N_{\tilde{c}}}} \sum_{c,\tilde{c}} \int d\tilde{k} \delta^D(k-\tilde{k}) \left[|k\rangle \otimes P_{ST}(|\varepsilon_\lambda\rangle \otimes |\tilde{\varepsilon}_{\tilde{\lambda}}\rangle) \otimes |c\rangle_G \otimes |\tilde{c}\rangle_{\tilde{G}} \right]_{\text{projected to singlet}}. \quad (7.10)$$

Here, the sum and normalization factor ensure a color-singlet state if $|c\rangle_G$ and $|\tilde{c}\rangle_{\tilde{G}}$ are from identified color spaces; for distinct groups, the singlet projection is trivial.

Representation theory dictates the decomposition of the tensor product of two vector (**1**) representations: $\mathbf{1} \otimes \mathbf{1} = \mathbf{2}_{\text{ST}} \oplus \mathbf{A}_{\text{AS}} \oplus \mathbf{S}_{\text{Tr}}$, corresponding to the spin-2 graviton (symmetric-traceless), a 2-form (antisymmetric), and a dilaton (scalar trace), respectively. Explicitly, if $|T^{\mu\nu}\rangle \equiv |\varepsilon^\mu\rangle \otimes |\tilde{\varepsilon}^\nu\rangle$, then $|T^{\mu\nu}\rangle = |h_{\text{ST}}^{\mu\nu}\rangle + |B_{\text{AS}}^{\mu\nu}\rangle + \eta^{\mu\nu} |\phi_{\text{Tr}}\rangle$. Our construction of Einstein-Hilbert gravity, by promoting amplitudes that solely describe interacting gravitons, effectively projects onto the $|h_{\text{ST}}^{\mu\nu}\rangle$ sector. This projection is crucial⁶, particularly when considering interactions with matter or loop-level effects, where unphysical propagation of the full **A** or **S** sectors must be avoided for pure gravity. The method of maximal cuts, as employed in our contact term extraction and in advanced loop computations [48], provides a systematic way to enforce such projections, ensuring that only the desired physical degrees of freedom contribute.

The operators $\mathcal{L}_m = \hat{\mathcal{O}}(\mathcal{A}_m) - \hat{\mathcal{O}}(\mathcal{A}_m)$ derived in this paper manifestly factorize into

⁶The alternative lands on the states of the fat graviton [47] which include states of the Kalb-Ramond antisymmetric two-form and dilaton.

gauge-theory-like operators. These are designed to act on the constituent kets within the tensor product structure of states like Eq. (7.9), ensuring that interactions respect the double-copy inheritance. The construction of all such gravitational operators from their gauge theory counterparts via our systematic procedure is thus, in principle, an algorithmic task to all multiplicities.

At the level of quantum states, we remain agnostic about whether the color trace implies a true partial trace over entangled subsystems in a quantum information sense, or if it is a formal projection onto the $G \times \tilde{G}$ invariant subspace of $\mathcal{H}_{\text{YM}} \otimes \tilde{\mathcal{H}}_{\text{YM}}$. What is essential is that the resulting gravitational Hilbert space sector is built from these gauge-invariant (color-singlet) tensor factor combinations, correctly capturing the physical graviton polarizations and their interactions.

While this state-level construction is most transparent in flat spacetime, it is not fundamentally limited. If the background geometry itself arises from a double copy of classical gauge field configurations (as in Kerr-Schild metrics or generalized classical double-copy solutions), graviton fluctuations $\delta h_{\mu\nu}$ can still be expressed as ST tensor products of gauge-field fluctuations $\delta A_\mu \otimes \delta \tilde{A}_\nu$ around these backgrounds. The projection to physical degrees of freedom remains locally well-defined, extending the applicability of this framework to a broad class of curved spacetimes.

Our construction of single-graviton states as specific projections of tensor products of Yang-Mills state structures, $|\text{GR}\rangle \sim (|\text{YM}\rangle \otimes_{\text{ST}} |\widetilde{\text{YM}}\rangle)_{\text{color-singlet}}$, provides a Fock-space basis for gravitational theories built from double-copy principles. This perspective, where the double copy is realized at the level of individual particle states and the operators that act upon them, complements other approaches that explore state-level manifestations of the double copy, notably c.f. the work of Cheung and Remmen [49], which explored N -graviton dynamics by introducing an entanglement ansatz between $2N$ -gluon states under the same $\text{SU}(N_c)$ gauge group.

8 Outlook

In this paper we introduced a method to promote color-dual amplitudes directly to operators at the level of the action which make manifest their double-copy structure. We introduce a generalization of the method of maximal cuts that lets us uniquely and algorithmically identify contact contributions,

$$\mathcal{L}_m = \hat{\mathcal{O}}(\mathcal{A}_m - \mathcal{A}_m), \quad (8.1)$$

and apply this to modding out redundancy that can occur when writing entire amplitudes as non-local operators. We introduce a transparent operator promotion that allows for quantum field operators to look like the momentum-space amplitude expressions they generate. We demonstrate the utility by for the first time, giving an action level expression that correctly generates the massless vector contributions to the open superstring theory at five-points – encoding an infinite number of higher-dimensional operators as a disk integral that can be lifted directly from the predictions of the two-dimensional CFT.

This has immediate applications. For EFT construction it offers an alternative to ansatz methods, especially for complex high-derivative operators or theories with many fields. The procedure is straightforward and there is non-trivial potential for automation and database engagement with physical content defining both actions and predictions simultaneously.

It should be noted that the efficacy of the double-copy prescription, evident in scattering amplitudes across a wide range of theories, has naturally spurred investigations into its underlying mathematical foundations. Significant progress has been made in understanding these structures from the perspective of homotopy algebras, such as L_∞ -algebras and the Batalin-Vilkovisky formalism (e.g., [50] and references therein). These approaches aim to provide a rigorous algebraic basis for color-kinematics duality and the double-copy operation itself, often by identifying kinematic algebras as part of these sophisticated mathematical structures.

Our work, while starting from the S-matrix and focusing on a direct, constructive path to actions, is complementary to these formal algebraic investigations. By providing an explicit ‘amplitude-to-action’ map that preserves double-copy structures, we offer a concrete realization of these principles at the level of effective Lagrangians and quantum states. It would be interesting future work to explore the precise connections between our constructed operators and their counterparts in the language of homotopy algebras.

The construction presented here explicitly applies the method of maximal cuts to identify and isolate the novel contribution of every local gravitational operator. This framework is not restricted to tree level: it builds upon the same unitarity-compatible bookkeeping used in multiloop computations, where maximal-cut techniques have been extensively employed to project out unwanted scalar and two-form contributions. In fact, this approach was recently formalized precisely as a maximal-cut-based operation in [48], where it was used to extract the Einstein–Hilbert predictions at loop level from the gravitational double copy.

Crucially, the same method also resolves subtleties that arise already at tree level when coupling to massive matter. In such cases, intermediate states in the double copy can include unphysical scalar components unless projected out via symmetric-traceless projections. The method of maximal cuts provides a systematic and physically meaningful way to enforce this projection channel-by-channel, ensuring that only the gravitational degrees of freedom propagate. As such, the operator construction presented here is not only valid at the level of free states but is fully compatible with loop-level unitarity and matter couplings at tree level.

More broadly, the double-copy principle extends far beyond field-theoretic S-matrix constructions, with both open and closed string theories admitting fully consistent, all-orders-in- α' double-copy formulations. The open superstring (OSS) amplitude, for instance, can be expressed as a field-theory-level double copy involving Z-theory and super Yang-Mills (sYM) [41–44]. We exploited this earlier to give the open-superstring operator at tree-level for five-points. Z-theory acts as a single-copy theory capturing all α' dependence and satisfying string monodromy relations (associated here with capital Latin indices, e.g., A, B). When its Chan-Paton stripped, ordered amplitudes Z_{Ab} (where b is a field-theory-like index) are appropriately combined with ordered sYM amplitudes $\text{SYM}_{\bar{b}}$, they yield the complete

ordered OSS amplitude, denoted $\text{OSS}_A = Z_a \cdot \text{sYM} \equiv Z_{Ab} \otimes^{b\tilde{b}} \text{sYM}_{\tilde{b}}$, not just its low-energy expansion. Here, lower-case indices obey standard field-theory KK and BCJ relations.

The closed superstring (CSS), in turn, emerges from a KLT-like relation where this string KLT kernel, $\otimes^{\alpha' AB}$ (which implements a single-valued projection at the level of multiple zeta values), acts on these OSS structures:

$$\text{CSS} = \text{OSS}_A \otimes^{\alpha' AB} \text{OSS}_B \quad (8.2)$$

$$= (\text{sYM} \cdot Z_A) \otimes^{\alpha' AB} (\text{sYM} \cdot Z_B) \quad (8.3)$$

$$= \text{sYM} \cdot \left((Z_A \otimes^{\alpha' AB} Z_B) \cdot \text{sYM} \right) \quad (8.4)$$

$$\equiv \text{sYM} \cdot \text{sv}(\text{sYM}) . \quad (8.5)$$

Critically note both open and closed superstring theory amplitudes are expressed ultimately in terms of field-theory double-copies in terms of objects that respect field theory rules (color-kinematics duality / KK-BCJ relations), extending naturally to bosonic and heterotic string amplitudes [51, 52].

A potential concern regarding operator constructions based on the duality between color and kinematics is whether such representations exist at arbitrary multiplicity and loop order. However, from an S-matrix forward perspective this is a red herring: all multiloop integrands can be systematically constructed from products of tree-level amplitudes via generalized unitarity. If color-kinematics duality holds at tree level (the bare minimum required to play in the web of color-dual theories) then any loop-level prediction can be built from tree data that already satisfies the required duality. This point was emphasized to spectacular effect in the constructive work of ref. [53]. In this sense, the universality of the operator construction is inherited from the tree-level data, and no additional off-shell extension of color-kinematics duality is required.

Moreover, the method of maximal cuts explicitly selects the contributing operator structures from on-shell configurations, which are fully determined by consistent Jacobi-satisfying tree-level numerators. The operator mapping defined in this work thus provides a constructive algorithm for building a gravitational operator basis compatible with the double-copy structure at all multiplicity and loop order, without requiring an off-shell or Lagrangian-level duality.

One of the defining strengths of the double-copy framework is its constructive nature: it does not presuppose a simple or perturbative gravitational background. Even highly non-trivial gravitational configurations — including those far from asymptotic flatness — can be realized as double copies of appropriately structured gauge-theory field configurations. In this sense, the complexity of the gravitational field $h_{\mu\nu}$ is not a limitation but a challenge to be matched by equally intricate single-copy data. This flexible, bottom-up construction is a feature, not a flaw, of the double-copy program.

The true promise of establishing the double copy as an action-and-state-level duality extends far beyond perturbative calculations. It offers a concrete strategy for confronting notoriously difficult gravitational phenomena by reframing them in the language of Yang-Mills theory. While non-perturbative Yang-Mills dynamics, such as those governing instanton-mediated processes, are themselves challenging, they are fundamentally rooted in a well-understood, unitary quantum field theory on flat spacetime.

This perspective invites a paradigm shift: questions about Hawking radiation, the black hole information paradox, or even eternal inflation, need not be solely pursued within the often conceptually fraught arena of (semi-classical) quantum gravity on curved backgrounds. Instead, the double copy provides a map to potentially more tractable (though still intricate) problems in their dual gauge theories. The path to understanding these profound gravitational puzzles may therefore lie in leveraging our robust toolkit for flat-space Yang-Mills theory. Indeed, with structure that simplifies the S-matrix, and with explicit operator connections now established, the challenge becomes one of technical execution within these “simpler” copies. We hope that this doubled motivation can inspire new innovations around the still-significant challenges within Yang-Mills theory, such as confinement, with the promise that their resolution will also pay dividends in our understanding of the evolution of time and space.

Acknowledgments

We especially thank Zvi Bern, Marco Chiodaroli, Radu Roiban, and Henrik Johansson for conversations, comments, and related collaboration. We would like to also thank members of the Amplitudes and Insight group at Northwestern University, Sai Sasank Chava, Yaxi Cheng, Alex Edison, Nic Pavao, Aslan Seifi, Cong Shen, Teddy Wexler, and John Zhang, for important conversations and feedback on an earlier versions. SZ is grateful for the support from Northwestern University’s Presidential Fellowship. This work was supported by the DOE under contract DE-SC0015910, by the Alfred P. Sloan Foundation, and by Northwestern University via the Amplitudes and Insight Group, Department of Physics and Astronomy, and Weinberg College of Arts and Sciences.

References

- [1] 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](#)].
- [2] J.J.M. Carrasco, L. Rodina and S. Zekioglu, *Composing effective prediction at five points*, *JHEP* **06** (2021) 169 [[2104.08370](#)].
- [3] J.J.M. Carrasco and N.H. Pavao, *UV massive resonance from IR double copy consistency*, *Phys. Rev. D* **109** (2024) 065006 [[2310.06316](#)].
- [4] Z. Bern, T. Dennen, Y.-t. Huang and M. Kiermaier, *Gravity as the Square of Gauge Theory*, *Phys. Rev. D* **82** (2010) 065003 [[1004.0693](#)].

- [5] M. Tolotti and S. Weinzierl, *Construction of an effective Yang-Mills Lagrangian with manifest BCJ duality*, *JHEP* **07** (2013) 111 [[1306.2975](#)].
- [6] C. Cheung and C.-H. Shen, *Symmetry for Flavor-Kinematics Duality from an Action*, *Phys. Rev. Lett.* **118** (2017) 121601 [[1612.00868](#)].
- [7] M. Ben-Shahar and M. Guillen, *10D super-Yang-Mills scattering amplitudes from its pure spinor action*, *JHEP* **12** (2021) 014 [[2108.11708](#)].
- [8] M. Ben-Shahar and H. Johansson, *Off-shell color-kinematics duality for Chern-Simons*, *JHEP* **08** (2022) 035 [[2112.11452](#)].
- [9] Z. Bern and A.K. Grant, *Perturbative gravity from QCD amplitudes*, *Phys. Lett. B* **457** (1999) 23 [[hep-th/9904026](#)].
- [10] 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.
- [11] Z. Bern, L.J. Dixon, M. Perelstein and J.S. Rozowsky, *Multileg one loop gravity amplitudes from gauge theory*, *Nucl. Phys. B* **546** (1999) 423 [[hep-th/9811140](#)].
- [12] C. Cheung and G.N. Remmen, *Twofold Symmetries of the Pure Gravity Action*, *JHEP* **01** (2017) 104 [[1612.03927](#)].
- [13] Z. Bern, J.J.M. Carrasco and H. Johansson, *New Relations for Gauge-Theory Amplitudes*, *Phys. Rev. D* **78** (2008) 085011 [[0805.3993](#)].
- [14] 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](#)].
- [15] R. Monteiro, D. O’Connell and C.D. White, *Black holes and the double copy*, *JHEP* **12** (2014) 056 [[1410.0239](#)].
- [16] A. Luna, R. Monteiro, D. O’Connell and C.D. White, *The classical double copy for Taub–NUT spacetime*, *Phys. Lett. B* **750** (2015) 272 [[1507.01869](#)].
- [17] A. Luna, R. Monteiro, I. Nicholson, D. O’Connell and C.D. White, *The double copy: Bremsstrahlung and accelerating black holes*, *JHEP* **06** (2016) 023 [[1603.05737](#)].
- [18] D.A. Kosower, B. Maybee and D. O’Connell, *Amplitudes, observables, and classical scattering*, *JHEP* **02** (2019) 137 [[1811.10950](#)].
- [19] H. Elvang and Y.-t. Huang, *Scattering Amplitudes*, [1308.1697](#).
- [20] J.J.M. Carrasco, *Gauge and Gravity Amplitude Relations*, in *Theoretical Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Frontier: Amplitudes for Colliders*, pp. 477–557, WSP, 2015, DOI [[1506.00974](#)].
- [21] Z. Bern, J.J. Carrasco, M. Chiodaroli, H. Johansson and R. Roiban, *The duality between color and kinematics and its applications*, *J. Phys. A* **57** (2024) 333002 [[1909.01358](#)].
- [22] L. Borsten, *Gravity as the square of gauge theory: a review*, *Riv. Nuovo Cim.* **43** (2020) 97.
- [23] T. Adamo, J.J.M. Carrasco, M. Carrillo-González, M. Chiodaroli, H. Elvang, H. Johansson et al., *Snowmass White Paper: the Double Copy and its Applications*, in *Snowmass 2021*, 4, 2022 [[2204.06547](#)].

- [24] Z. Bern, J.J. Carrasco, M. Chiodaroli, H. Johansson and R. Roiban, *The SAGEX review on scattering amplitudes Chapter 2: An invitation to color-kinematics duality and the double copy*, *J. Phys. A* **55** (2022) 443003 [[2203.13013](#)].
- [25] Z. Bern, J.J.M. Carrasco, M. Chiodaroli, H. Johansson and R. Roiban, *Supergravity Amplitudes, the Double Copy, and Ultraviolet Behavior*, (2023), DOI [[2304.07392](#)].
- [26] J.J.M. Carrasco and N.H. Pavao, *Virtues of a symmetric-structure double copy*, *Phys. Rev. D* **107** (2023) 065005 [[2211.04431](#)].
- [27] K. Rafi and J. Tao, *The diameter of the thick part of moduli space and simultaneous whitehead moves*, *Duke Math. J.* **162** (2013) 1833.
- [28] Z. Bern, J.J.M. Carrasco, H. Johansson and D.A. Kosower, *Maximally supersymmetric planar Yang-Mills amplitudes at five loops*, *Phys. Rev. D* **76** (2007) 125020 [[0705.1864](#)].
- [29] Z. Bern, L.J. Dixon, D.C. Dunbar and D.A. Kosower, *One loop n point gauge theory amplitudes, unitarity and collinear limits*, *Nucl. Phys. B* **425** (1994) 217 [[hep-ph/9403226](#)].
- [30] Z. Bern, L.J. Dixon, D.C. Dunbar and D.A. Kosower, *Fusing gauge theory tree amplitudes into loop amplitudes*, *Nucl. Phys. B* **435** (1995) 59 [[hep-ph/9409265](#)].
- [31] Z. Bern and A.G. Morgan, *Massive loop amplitudes from unitarity*, *Nucl. Phys. B* **467** (1996) 479 [[hep-ph/9511336](#)].
- [32] Z. Bern, L.J. Dixon and D.A. Kosower, *One loop amplitudes for $e^+ e^-$ to four partons*, *Nucl. Phys. B* **513** (1998) 3 [[hep-ph/9708239](#)].
- [33] R. Britto, F. Cachazo and B. Feng, *Generalized unitarity and one-loop amplitudes in $N=4$ super-Yang-Mills*, *Nucl. Phys. B* **725** (2005) 275 [[hep-th/0412103](#)].
- [34] Z. Bern, J.J.M. Carrasco, L.J. Dixon, H. Johansson and R. Roiban, *Manifest Ultraviolet Behavior for the Three-Loop Four-Point Amplitude of $N=8$ Supergravity*, *Phys. Rev. D* **78** (2008) 105019 [[0808.4112](#)].
- [35] Z. Bern, J.J.M. Carrasco, L.J. Dixon, H. Johansson and R. Roiban, *The Complete Four-Loop Four-Point Amplitude in $N=4$ Super-Yang-Mills Theory*, *Phys. Rev. D* **82** (2010) 125040 [[1008.3327](#)].
- [36] J.J.M. Carrasco and H. Johansson, *Generic multiloop methods and application to $N=4$ super-Yang-Mills*, *J. Phys. A* **44** (2011) 454004 [[1103.3298](#)].
- [37] J. Broedel and J.J.M. Carrasco, *Virtuous Trees at Five and Six Points for Yang-Mills and Gravity*, *Phys. Rev. D* **84** (2011) 085009 [[1107.4802](#)].
- [38] M. de Roo and M.G.C. Eenink, *The Effective action for the four point functions in Abelian open superstring theory*, *JHEP* **08** (2003) 036 [[hep-th/0307211](#)].
- [39] C.R. Mafra, O. Schlotterer and S. Stieberger, *Complete N -Point Superstring Disk Amplitude I. Pure Spinor Computation*, *Nucl. Phys. B* **873** (2013) 419 [[1106.2645](#)].
- [40] C.R. Mafra, O. Schlotterer and S. Stieberger, *Complete N -Point Superstring Disk Amplitude II. Amplitude and Hypergeometric Function Structure*, *Nucl. Phys. B* **873** (2013) 461 [[1106.2646](#)].
- [41] J. Broedel, O. Schlotterer and S. Stieberger, *Polylogarithms, Multiple Zeta Values and Superstring Amplitudes*, *Fortsch. Phys.* **61** (2013) 812 [[1304.7267](#)].

- [42] J.J.M. Carrasco, C.R. Mafra and O. Schlotterer, *Abelian Z-theory: NLSM amplitudes and α' -corrections from the open string*, *JHEP* **06** (2017) 093 [[1608.02569](#)].
- [43] C.R. Mafra and O. Schlotterer, *Non-abelian Z-theory: Berends-Giele recursion for the α' -expansion of disk integrals*, *JHEP* **01** (2017) 031 [[1609.07078](#)].
- [44] J.J.M. Carrasco, C.R. Mafra and O. Schlotterer, *Semi-abelian Z-theory: NLSM+ ϕ^3 from the open string*, *JHEP* **08** (2017) 135 [[1612.06446](#)].
- [45] C.-H. Fu, Y.-J. Du and B. Feng, *Note on symmetric BCJ numerator*, *JHEP* **08** (2014) 098 [[1403.6262](#)].
- [46] S.G. Naculich, *Scattering equations and virtuous kinematic numerators and dual-trace functions*, *JHEP* **07** (2014) 143 [[1404.7141](#)].
- [47] A. Luna, R. Monteiro, I. Nicholson, A. Ochirov, D. O’Connell, N. Westerberg et al., *Perturbative spacetimes from Yang-Mills theory*, *JHEP* **04** (2017) 069 [[1611.07508](#)].
- [48] J.J.M. Carrasco and I.A. Vazquez-Holm, *Extracting Einstein from the loop-level double-copy*, *JHEP* **11** (2021) 088 [[2108.06798](#)].
- [49] C. Cheung and G.N. Remmen, *Entanglement and the double copy*, *JHEP* **05** (2020) 100 [[2002.10470](#)].
- [50] L. Borsten, B. Jurco, H. Kim, T. Macrelli, C. Saemann and M. Wolf, *Double Copy From Tensor Products of Metric BV $\square$ -Algebras*, *Fortsch. Phys.* **73** (2025) 2300270 [[2307.02563](#)].
- [51] T. Azevedo, M. Chiodaroli, H. Johansson and O. Schlotterer, *Heterotic and bosonic string amplitudes via field theory*, *JHEP* **10** (2018) 012 [[1803.05452](#)].
- [52] J.J.M. Carrasco, M. Lewandowski and N.H. Pavao, *Color-Dual Fates of $F3$, $R3$, and $N=4$ Supergravity*, *Phys. Rev. Lett.* **131** (2023) 051601 [[2203.03592](#)].
- [53] Z. Bern, E. Herrmann, R. Roiban, M.S. Ruf and M. Zeng, *Global bases for nonplanar loop integrands, generalized unitarity, and the double copy to all loop orders*, *JHEP* **06** (2025) 115 [[2408.06686](#)].